

Facultad de Ciencias Humanas y Sociales  
Departamento de Traducción y Comunicación



# **Estrategias de búsqueda onomasiológica en la actividad de traducción.**

**Una ayuda al diseño de diccionarios terminológicos**

**TESIS DOCTORAL**

Presentada por:

Verónica Pastor Enríquez

Dirigida por:

Dra. Amparo Alcina Caudet

Castellón de la Plana, 2013



## AGRADECIMIENTOS

Son muchas las personas e instituciones sin las que esta tesis no habría sido posible. Por ello, es justo que este trabajo comience expresándoles a todos mi más sincera gratitud.

Me gustaría agradecer, en primer lugar, a mi directora Amparo Alcina que me propusiera formar parte del grupo de investigación TecnoLeTTra y se ofreciera a ser la tutora de mi beca de investigación y directora de este trabajo. Sus consejos, sugerencias y correcciones han permitido el progreso de esta investigación.

Quisiera reconocer igualmente la calidad investigadora y humana de mis compañeras del grupo de investigación TecnoLeTTra. Gracias a todas por vuestro apoyo y, sobre todo, por vuestra amistad.

Mi agradecimiento también a todos los miembros del Departamento de Traducción y Comunicación de la Universitat Jaume I, incluidos el personal administrativo y técnico, por el maravilloso trato que he recibido. A los becarios y becarias por todos los momentos que hemos compartido.

Gracias a los miembros del grupo CRTT de la Université Lumière Lyon 2 por su acogida y asesoramiento durante mi estancia de investigación de tres meses. A los miembros del grupo OLST de la Université de Montréal por los seminarios y las reuniones informales donde hemos intercambiado ideas de gran valor para esta investigación. También a Gerardo Sierra por enviarme amablemente su tesis doctoral sobre diccionarios onomasiológicos y a Paul Bogaards por su artículo sobre el uso de diccionarios de lenguas extranjeras.

A la revista *International Journal of Lexicography*, la fuente original de publicación de los resultados de mi análisis empírico sobre diccionarios electrónicos, por autorizarme a reproducir dicho análisis en esta tesis doctoral y por los valiosos comentarios de sus revisores que han contribuido a la mejora del capítulo tercero de este trabajo.

A la Generalitat Valenciana por mi beca FPI y por financiar también mi estancia, asistencia y participación en congresos y cursos. Igualmente a los proyectos ONTODIC: *Metodología y tecnologías de elaboración de diccionarios terminológicos onomasiológicos basados en ontologías. Recursos terminológicos para la e-Traducción* (TSI2006-01911), ONTODIC II: *Metodología y tecnologías de elaboración de diccionarios terminológicos de colocaciones basados en ontologías. Recursos terminológicos para la e-Traducción* (TIN2009-07690), ambos financiados por el

Gobierno de España, y al proyecto DiCoInfo-ES: *Las funciones léxicas de la Lexicología Explicativa y Combinatoria (LEC) aplicadas a la terminología de la informática en español* (P1-1B2008-57), financiado por la Fundación Bancaja-Universitat Jaume I.

Esta tesis no habría sido posible sin la participación de los 64 traductores profesionales anónimos que nos han aportado su valiosa experiencia en el estudio piloto y el estudio empírico sobre estrategias de búsqueda onomasiológica en traducción.

Por último, gracias a mi familia, en especial a mis padres, por escucharme y animarme a continuar. A todos mis amigos, por estar a mi lado en los buenos y en los malos momentos. Pilar y Jorge, merecéis una mención especial en estos agradecimientos por el soporte informático. A mi marido Jorge, por su paciencia y su apoyo incondicional.

A todos los que han contribuido con su ayuda, experiencia y comprensión a este trabajo, muchas gracias.



# ÍNDICE

<b>RESUMEN .....</b>	<b>1</b>
<b>0 INTRODUCCIÓN.....</b>	<b>3</b>
0.1 HIPÓTESIS.....	4
0.2 OBJETIVOS.....	5
0.3 ESTRUCTURA DEL TRABAJO Y METODOLOGÍA.....	5
<b>ABSTRACT .....</b>	<b>11</b>
<b>0 INTRODUCTION.....</b>	<b>13</b>
0.1 HYPOTHESIS.....	14
0.2 OBJECTIVES .....	14
0.3 STRUCTURE AND METHODOLOGY.....	15
<b>PARTE I. ANTECEDENTES.....</b>	<b>21</b>
<b>1 LA BÚSQUEDA ONOMASIOLÓGICA (DEL CONCEPTO A LA DENOMINACIÓN). LOS     DICCIONARIOS ONOMASIOLÓGICOS .....</b>	<b>23</b>
1.1 ANTECEDENTES DE DICCIONARIOS ONOMASIOLÓGICOS .....	23
1.1.1 Roget's International Thesaurus of English Words and Phrases .....	25
1.1.2 Diccionarios inversos.....	26
1.1.3 Diccionario ideológico de la lengua española.....	28
1.1.4 Diccionarios combinatorios.....	31
1.2 PROYECTOS DE INVESTIGACIÓN SOBRE DICCIONARIOS ONOMASIOLÓGICOS ELECTRÓNICOS .....	32
1.2.1 Diccionario Electrónico para la Búsqueda Onomasiológica.....	32
1.3 RECURSOS LINGÜÍSTICOS BASADOS EN ONTOLOGÍAS .....	34
1.4 PERSPECTIVAS PARA EL FUTURO DE LOS DICCIONARIOS ONOMASIOLÓGICOS .....	34
1.5 RECAPITULACIÓN.....	35
<b>2 EL PROCESO DE TRADUCCIÓN, LA COMPETENCIA TERMINOLÓGICA DEL TRADUCTOR Y     LA RELACIÓN ENTRE TRADUCCIÓN Y TERMINOLOGÍA .....</b>	<b>37</b>
2.1 EL PROCESO DE TRADUCCIÓN Y LOS PROBLEMAS DE TRADUCCIÓN.....	37
2.1.1 Investigaciones empíricas sobre el proceso de traducción.....	39
2.1.2 Resultados de las investigaciones empíricas sobre problemas y estrategias de traducción.....	40
2.2 LA COMPETENCIA TRADUCTORA Y LA COMPETENCIA DOCUMENTAL Y TERMINOLÓGICA .....	45
2.3 LA RELACIÓN ENTRE TRADUCCIÓN Y TERMINOLOGÍA.....	53

2.4	HERRAMIENTAS TERMINOLÓGICAS EN TRADUCCIÓN .....	59
2.5	RECAPITULACIÓN.....	62
<b>PARTE II. TÉCNICAS DE BÚSQUEDA .....</b>		<b>63</b>
<b>3</b>	<b>TÉCNICAS DE BÚSQUEDA EN LOS RECURSOS Y NUESTRA PROPUESTA DE CLASIFICACIÓN .....</b>	<b>65</b>
3.1	ELEMENTOS DE LA CLASIFICACIÓN DE TÉCNICAS DE BÚSQUEDA .....	66
3.2	DICCIONARIOS ELECTRÓNICOS .....	68
3.2.1	Utilidad del diccionario como recurso de búsqueda, según la bibliografía .....	69
3.2.2	Las formas de búsqueda en diccionarios electrónicos, según la bibliografía .....	70
3.2.3	Análisis empírico de técnicas de búsqueda en diccionarios electrónicos .....	87
3.2.4	Clasificación de técnicas de búsqueda en diccionarios electrónicos.....	104
3.2.4.1	El primer elemento: la consulta .....	105
3.2.4.1.1	Búsqueda mediante una palabra exacta .....	106
3.2.4.1.2	Búsqueda mediante una palabra truncada .....	106
3.2.4.1.3	Búsqueda mediante una expresión aproximada.....	107
3.2.4.1.3.1	Búsqueda de una forma flexiva.....	108
3.2.4.1.3.2	Búsqueda de una expresión con similitud ortográfica o fonética a otra palabra.....	108
3.2.4.1.4	Búsqueda mediante un anagrama.....	110
3.2.4.1.5	Búsqueda mediante combinación de dos o más palabras.....	111
3.2.4.1.5.1	Presencia de todas las palabras introducidas.....	111
3.2.4.1.5.2	Presencia de alguna de las palabras introducidas .....	112
3.2.4.1.5.3	Presencia de una palabra y ausencia de otra.....	114
3.2.4.1.5.4	Presencia de la secuencia exacta de palabras introducida .....	115
3.2.4.1.5.5	Búsqueda mediante consultas en lenguaje natural.....	115
3.2.4.1.6	Búsqueda mediante el historial de consultas .....	117
3.2.4.1.7	Filtros.....	117
3.2.4.1.7.1	Filtro de categoría gramatical.....	117
3.2.4.1.7.2	Filtro de área temática .....	118
3.2.4.1.7.3	Filtro de idioma.....	119
3.2.4.2	El segundo elemento: el recurso o instrumento.....	119
3.2.4.2.1	Búsqueda en las entradas.....	120
3.2.4.2.2	Búsqueda en los campos de contenido.....	121
3.2.4.2.2.1	Búsqueda en los campos de definiciones.....	121
3.2.4.2.2.2	Búsqueda en los campos de relaciones léxicas.....	122
3.2.4.2.3	Búsqueda en un foro interno .....	131
3.2.4.2.4	Búsqueda en un corpus interno .....	132
3.2.4.2.5	Búsqueda en el índice temático .....	134

3.2.4.2.6	Búsqueda en conexiones externas .....	136
3.2.4.3	El tercer elemento: el resultado .....	136
3.2.4.3.1	Contextos.....	137
3.2.4.3.2	Colocaciones y palabras relacionadas.....	138
3.2.4.3.3	Representación gráfica de las relaciones de una palabra.....	139
3.2.4.3.4	Denominación o lista de denominaciones .....	140
3.2.4.3.5	Imagen o lista de imágenes .....	140
3.2.4.3.6	Archivos de audio .....	141
3.2.5	Recapitulación de nuestra propuesta de clasificación de técnicas de búsqueda en diccionarios electrónicos.....	142
3.3	CORPUS.....	143
3.3.1	Utilidad de los corpus como recurso de búsqueda terminológica, según la bibliografía.....	143
3.3.2	Tipología de corpus y las posibilidades de búsqueda, según la bibliografía.....	146
3.3.2.1	Según la compilación y la naturaleza puntual o estable .....	147
3.3.2.2	Según el marcado.....	151
3.3.2.3	Según el número de lenguas y la relación que se establece entre los textos .....	153
3.3.3	Análisis empírico de técnicas de búsqueda en corpus.....	158
3.3.3.1	Corpus en línea que incorporan sus propias interfaces de búsqueda .....	159
3.3.3.2	Programas de gestión de corpus .....	164
3.3.3.3	Programas de traducción asistida por ordenador (TAO) .....	171
3.3.4	Clasificación de técnicas de búsqueda en corpus.....	175
3.3.4.1	El primer elemento: la consulta .....	176
3.3.4.1.1	Búsqueda mediante expresiones léxicas.....	176
3.3.4.1.2	Búsqueda mediante expresiones gramaticales.....	177
3.3.4.1.3	Búsqueda mediante un número.....	177
3.3.4.1.4	Búsqueda mediante dos o más expresiones .....	178
3.3.4.1.4.1	Búsqueda mediante una combinación continua de expresiones.....	178
3.3.4.1.4.2	Búsqueda mediante una combinación discontinua de expresiones (con delimitación de proximidad).....	181
3.3.4.1.5	Búsqueda mediante el historial de consultas .....	183
3.3.4.1.6	Filtros.....	183
3.3.4.2	El segundo elemento: el recurso o instrumento.....	184
3.3.4.2.1	Búsqueda en un corpus monolingüe .....	185
3.3.4.2.2	Búsqueda en un corpus bilingüe o multilingüe: corpus comparable, corpus paralelo y corpus alineado.....	185
3.3.4.2.3	Búsqueda en un corpus etiquetado.....	187
3.3.4.3	El tercer elemento: el resultado .....	187
3.3.4.3.1	Lista de concordancias.....	188

3.3.4.3.1.1	Lista de concordancias monolingües.....	188
3.3.4.3.1.2	Lista de concordancias bilingües.....	189
3.3.4.3.2	Lista de palabras.....	190
3.3.4.3.3	Lista de agrupaciones de palabras.....	191
3.3.4.3.4	Lista de colocaciones.....	193
3.3.4.3.5	Lista de palabras relacionadas semánticamente.....	195
3.3.4.3.6	Lista de palabras relacionadas sintácticamente.....	198
3.3.4.3.7	Lista de categorías gramaticales.....	200
3.3.4.3.8	Representación gráfica de las relaciones de una palabra.....	201
3.3.4.3.9	Información estadística.....	202
3.3.5	Recapitulación de nuestra propuesta de clasificación de técnicas de búsqueda en corpus.....	203
3.4	INTERNET.....	204
3.4.1	Utilidad de Internet como recurso de búsqueda terminológica, según la bibliografía.....	205
3.4.2	La evolución de la Web y su influencia en las posibilidades de búsqueda.....	211
3.4.3	Análisis empírico de técnicas de búsqueda en Internet.....	216
3.4.3.1	Buscadores.....	217
3.4.3.2	Herramientas de análisis de la Web como corpus.....	224
3.4.4	Clasificación de técnicas de búsqueda en Internet.....	236
3.4.4.1	El primer elemento: la consulta.....	237
3.4.4.1.1	Búsqueda mediante una expresión léxica.....	237
3.4.4.1.2	Búsqueda mediante un número.....	238
3.4.4.1.3	Búsqueda mediante dos o más expresiones.....	239
3.4.4.1.4	Búsqueda mediante el historial de consultas.....	242
3.4.4.1.5	Filtros.....	243
3.4.4.2	El segundo elemento: el recurso o instrumento.....	244
3.4.4.3	El tercer elemento: el resultado.....	244
3.4.4.3.1	Lista de páginas web y lista de concordancias.....	244
3.4.4.3.2	Lista de palabras.....	247
3.4.4.3.3	Lista de agrupaciones de palabras.....	248
3.4.4.3.4	Lista de palabras relacionadas semánticamente.....	248
3.4.4.3.5	Lista de colocaciones.....	249
3.4.4.3.6	Representación gráfica de las relaciones de una palabra.....	250
3.4.4.3.7	Lista de imágenes.....	250
3.4.4.3.8	Archivos multimedia.....	251
3.4.4.3.9	Información estadística.....	252
3.4.5	Recapitulación de nuestra propuesta de clasificación de técnicas de búsqueda en Internet.....	253

3.5	PROPUESTA CONJUNTA DE CLASIFICACIÓN DE TÉCNICAS DE BÚSQUEDA EN DICCIONARIOS, CORPUS E INTERNET.....	254
<b>PARTE III. ESTRATEGIAS DE BÚSQUEDA .....</b>		<b>261</b>
<b>4</b>	<b>ESTRATEGIAS DE BÚSQUEDA EN TRADUCCIÓN Y NUESTRA PROPUESTA DE CLASIFICACIÓN .....</b>	<b>263</b>
4.1	ENFOQUE ORIENTADO A LOS USUARIOS.....	264
4.1.1	Recuperación de la información orientada a los usuarios.....	264
4.1.2	Lexicografía orientada a los usuarios .....	265
4.1.2.1	Teoría general de la lexicografía .....	268
4.1.2.2	Teoría moderna de las funciones lexicográficas.....	271
4.2	ESTUDIOS SOBRE ESTRATEGIAS DE BÚSQUEDA DE LOS TRADUCTORES EN DICCIONARIOS .....	290
4.2.1	Meyer (1988) .....	292
4.2.2	Tomaszczyk (1989).....	295
4.2.3	Roberts (1990) .....	296
4.2.4	Atkins y Varantola (1997; 1998b) .....	298
4.2.5	Varantola (1998; 2003b).....	300
4.2.6	Rogers y Ahmad (1998).....	302
4.2.7	Mackintosh (1998).....	304
4.2.8	Fraser (1999).....	307
4.2.9	Livbjerg y Mees (1999; 2003) .....	308
4.2.10	Corpas Pastor et al. (2001).....	311
4.2.11	Tarp (2002) .....	314
4.2.12	De la Rosa Martín (2003).....	316
4.2.13	Sánchez Ramos (2004b).....	318
4.2.14	Prieto Velasco y López Rodríguez (2009: 206-208) .....	319
4.2.15	Jousse et al. (2011).....	320
4.2.16	Resumen de estrategias de búsqueda en diccionarios detectadas en los estudios .....	321
4.3	ESTUDIOS SOBRE ESTRATEGIAS DE BÚSQUEDA DE LOS TRADUCTORES EN CORPUS.....	327
4.3.1	Estudios cualitativos sobre estrategias de búsqueda en corpus de distinto tipo ...	330
4.3.1.1	Aston (1999; 2000).....	330
4.3.1.2	Varantola (2000; 2002a; 2003a; 2006) .....	333
4.3.1.3	Kübler (2003) .....	336
4.3.1.4	Moraes (2006) .....	338
4.3.1.5	Roberts y Bossé-Andrieu (2006).....	339
4.3.1.6	Rodríguez Inés (2009).....	342

4.3.2	Estudios cualitativos sobre estrategias de búsqueda en corpus bilingües paralelos y comparables .....	342
4.3.2.1	Danielsson y Ridings (1996) .....	342
4.3.2.2	Teubert (1996) .....	343
4.3.2.3	Schäffner (1998) .....	344
4.3.2.4	Zanettin (1998; 2001b; 2002a; 2002b) .....	345
4.3.2.5	Gavioli y Zanettin (2000) .....	349
4.3.2.6	Bowker y Pearson (2002) .....	351
4.3.2.7	Hansen y Teich (2002) .....	352
4.3.2.8	Vargas Sierra (2002) .....	352
4.3.2.9	Pearson (2003) .....	353
4.3.2.10	Sánchez-Gijón (2003a; 2004a; 2004b) .....	355
4.3.2.11	Lagoudaki (2006) .....	359
4.3.2.12	Corpas Pastor y Seghiri (2009) .....	361
4.3.2.13	Philip (2009) .....	363
4.3.3	Estudios cualitativos sobre estrategias de búsqueda en corpus monolingües .....	364
4.3.3.1	Pearson (1996a; 1996c) .....	364
4.3.3.2	Bowker (1998a; 1998b; 2000; 2003a) .....	367
4.3.3.3	Friedbichler y Friedbichler (2000) .....	378
4.3.3.4	Stewart (2000) .....	379
4.3.3.5	Bowker y Pearson (2002) .....	380
4.3.3.6	Coffey (2002) .....	383
4.3.3.7	Wilkinson (2005a; 2005b; 2006a; 2006b; 2007; 2010) .....	383
4.3.3.8	Montero Martínez y Faber (2008) .....	389
4.3.3.9	Sánchez-Gijón (2009) .....	390
4.3.4	Resumen de estrategias de búsqueda en corpus detectadas en los estudios .....	391
4.4	ESTUDIOS SOBRE ESTRATEGIAS DE BÚSQUEDA DE LOS TRADUCTORES EN INTERNET .....	396
4.4.1	Alanen (1996) .....	399
4.4.2	Bergeron y Van Steenberg (2000) .....	399
4.4.3	Palomares Perraut y Pinto (2000) .....	402
4.4.4	Austermühl (2001) .....	403
4.4.5	Gambín (2004) .....	405
4.4.6	Gile (2004) .....	406
4.4.7	Walter (2006) .....	407
4.4.8	Maniez (2007) .....	408
4.4.9	Pinto y Sales (2007) .....	410
4.4.10	Montero Martínez y Faber (2008) .....	411
4.4.11	Enríquez Raído (2011) .....	413

4.4.12	Resumen de estrategias de búsqueda en Internet detectadas en los estudios.....	419
4.5	RECAPITULACIÓN.....	424
4.5.1	Métodos de análisis de las estrategias de búsqueda .....	424
4.5.2	Propuesta preliminar de clasificación de estrategias de búsqueda en traducción.....	430
4.5.2.1	Elementos que intervienen en una estrategia de búsqueda.....	431
4.5.2.2	Estrategias de búsqueda detectadas en los estudios.....	432
4.5.2.2.1	Las necesidades o problemas.....	432
4.5.2.2.2	Las acciones de búsqueda.....	434
4.5.2.2.2.1	Acciones para la resolución de problemas en LO.....	434
4.5.2.2.2.2	Acciones para la resolución de problemas en LM .....	437
4.5.2.3	Valoración sobre las estrategias de búsqueda identificadas en los estudios.....	448
<b>5</b>	<b>ESTUDIO EMPÍRICO SOBRE ESTRATEGIAS DE BÚSQUEDA ONOMASIOLOGICA EN TRADUCCIÓN.....</b>	<b>451</b>
5.1	ESTUDIO PILOTO: ENCUESTA A TRADUCTORES SOBRE TÉCNICAS Y ESTRATEGIAS DE BÚSQUEDA ONOMASIOLOGICA.....	454
5.1.1	Objetivos del estudio piloto.....	454
5.1.2	Metodología.....	455
5.1.2.1	Diseño del cuestionario .....	455
5.1.2.2	Sujetos.....	456
5.1.2.3	Recogida y análisis de la información .....	457
5.1.3	Resultados .....	457
5.1.4	Discusión.....	461
5.2	ESTUDIO PRINCIPAL: ENTREVISTA SOBRE ESTRATEGIAS DE BÚSQUEDA ONOMASIOLOGICA EN TRADUCCIÓN .....	464
5.2.1	Objetivos del estudio empírico .....	464
5.2.2	Metodología.....	465
5.2.2.1	Teoría consultada sobre cómo realizar entrevistas .....	466
5.2.2.2	Diseño de la entrevista.....	470
5.2.2.2.1	El guion y las preguntas de la entrevista.....	470
5.2.2.3	Sujetos.....	472
5.2.2.3.1	Selección de informantes.....	472
5.2.2.3.2	Acceso y contacto.....	473
5.2.2.4	Realización de las entrevistas .....	474
5.2.2.5	Análisis de la información.....	475
5.2.2.5.1	Recursos utilizados para el análisis .....	475
5.2.2.5.1.1	Programas utilizados para la transcripción: F4 y Dragon Naturally Speaking .....	476
5.2.2.5.1.2	Programa utilizado para la codificación: Atlas.ti.....	480

5.2.2.5.2	La fase de transcripción: preparación de las entrevistas para el análisis.....	483
5.2.2.5.3	La fase de codificación: análisis de las entrevistas.....	484
5.2.3	Resultados .....	494
5.2.3.1	Resultados sobre las características de los entrevistados.....	495
5.2.3.2	Resultados sobre las necesidades en traducción. Situaciones y frecuencia de aparición de las necesidades onomasiológicas .....	500
5.2.3.3	Resultados sobre las estrategias de búsqueda de los traductores y las estrategias de búsqueda onomasiológica .....	502
5.2.3.3.1	Estrategias de búsqueda para la resolución de problemas de recepción (LO.a).....	504
5.2.3.3.2	Estrategias de búsqueda para la resolución de problemas de transferencia (LM.b) .....	507
5.2.3.3.3	Estrategias de búsqueda para la resolución de problemas de verificación .....	521
5.2.3.3.4	Estrategias de búsqueda para la resolución de problemas de producción (LM.c) .....	523
5.2.3.4	Resultados sobre percepciones y expectativas de los usuarios respecto de los recursos..	524
5.2.3.4.1	Ventajas y desventajas observadas en los recursos .....	525
5.2.3.4.1.1	Ventajas y desventajas observadas en los diccionarios .....	525
5.2.3.4.1.2	Ventajas y desventajas observadas en los programas de traducción asistida.....	530
5.2.3.4.1.3	Ventajas y desventajas observadas en los corpus.....	532
5.2.3.4.1.4	Ventajas y desventajas observadas en Internet .....	532
5.2.3.4.1.5	Ventajas y desventajas observadas en los foros.....	534
5.2.3.4.2	Demandas en los recursos.....	534
5.2.3.4.2.1	Demandas en los diccionarios.....	535
5.2.3.4.2.2	Demandas en los programas de traducción asistida .....	538
5.2.3.4.2.3	Demandas en los corpus .....	538
5.2.3.4.2.4	Demandas en Internet.....	539
5.2.4	Discusión.....	540
5.2.4.1	Discusión sobre las necesidades onomasiológicas de los traductores.....	540
5.2.4.2	Discusión sobre las estrategias de búsqueda de los traductores y las estrategias de búsqueda onomasiológica .....	541
5.2.4.2.1	Discusión sobre las estrategias de búsqueda para la resolución de problemas de recepción (LO.a).....	541
5.2.4.2.2	Discusión sobre las estrategias de búsqueda para la resolución de problemas de transferencia (LM.b) .....	543
5.2.4.2.3	Discusión sobre las estrategias de búsqueda para la resolución de problemas de verificación .....	550
5.2.4.2.4	Discusión sobre las estrategias de búsqueda para la resolución de problemas de producción (LM.c) .....	551
5.2.4.2.5	Clasificación de estrategias de búsqueda en traducción donde se incluyen las estrategias de búsqueda onomasiológica detectadas.....	552
5.2.4.3	Discusión sobre percepciones y expectativas de los usuarios respecto de los recursos.....	561



5.2.4.4	Propuestas para la elaboración de recursos orientados a los traductores .....	564
<b>6</b>	<b>CONCLUSIÓN .....</b>	<b>569</b>
6.1	CONCLUSIÓN SOBRE ESTRATEGIAS DE BÚSQUEDA ONOMASIOLOGICA EN TRADUCCIÓN .....	572
6.2	CONCLUSIÓN SOBRE LA ELABORACIÓN DE RECURSOS ORIENTADOS A LOS TRADUCTORES .....	575
6.3	TRABAJOS FUTUROS .....	581
<b>6</b>	<b>CONCLUSION .....</b>	<b>582</b>
6.1	CONCLUSION ON ONOMASIOLOGICAL SEARCH STRATEGIES IN TRANSLATION .....	584
6.2	CONCLUSION ON THE DEVELOPMENT OF TRANSLATOR-ORIENTED RESOURCES .....	587
6.3	FUTURE WORK.....	592
	<b>BIBLIOGRAFÍA.....</b>	<b>595</b>
	<b>RECURSOS ANALIZADOS.....</b>	<b>632</b>

## ÍNDICE DE TABLAS

Tabla 1. Categorías de la parte sinóptica del <i>Diccionario ideológico de la lengua española</i> .....	29
Tabla 2. Capacidad para documentarse del traductor de textos especializados (Gonzalo García 2004: 278).....	50
Tabla 3. Características textuales de los textos técnicos y competencias requeridas para la traducción técnica (Gamero 2001: 48).....	50
Tabla 4. Tipología de problemas suscitados por unidades de significación especializada en la traducción especializada (Sánchez-Gijón 2003a: 355). ....	57
Tabla 5. Comparación de las fases del proceso traductor con y sin ayuda de herramientas informáticas (Alcina 2008: 92).....	61
Tabla 6. Sueños sobre accesibilidad parte 1: comunicarse con la base de datos (De Schryver 2003: 174).....	75
Tabla 7. Sueños sobre accesibilidad parte 2: búsqueda en toda la base de datos (De Schryver 2003: 176).....	76
Tabla 8. Sueños sobre accesibilidad parte 3: acceso a la macroestructura (De Schryver 2003: 177). 77	
Tabla 9. Sueños sobre accesibilidad parte 4: acceso a la microestructura (De Schryver 2003: 179)... 77	
Tabla 10. Sueños sobre accesibilidad parte 5: acceso a la medioestructura (De Schryver 2003: 181). ....	78
Tabla 11. Onomasiological search matrix, según Oppentocht y Schutz (2003: 222, adaptado de Geeraerts 2000: 82).....	82
Tabla 12. Preguntas y respuestas de un diccionario convencional (Sager 1993: 235). ....	84
Tabla 13. Tipo de consulta y de preguntas en un banco de términos informatizado moderno (Sager 1993: 236).....	84
Tabla 14. Resumen de la clasificación de técnicas de búsqueda en diccionarios electrónicos. ....	142
Tabla 15. Contextos definitorios de <i>PCR</i> extraídos del corpus de medicina e informática en inglés de <i>Bwananet</i> .....	181
Tabla 16. Resultados de la búsqueda de las expresiones <i>Cytomegalovirus</i> y <i>is a</i> hasta 5 posiciones de distancia en el <i>BNC</i> .....	182
Tabla 17. Resultados de la búsqueda del número 100 hasta 5 posiciones de la forma <i>metros</i> en el <i>Corpus del español</i> (BYU).....	182
Tabla 18. Resultados de búsqueda de la consulta <i>run</i> en el corpus paralelo alineado <i>COMPARA</i> ...186	
Tabla 19. Resultados de búsqueda de las consultas de la forma <i>run</i> en inglés y el lema <i>correr</i> en portugués en el corpus paralelo alineado <i>COMPARA</i> .....	186
Tabla 20. Resultado de concordancias del lema <i>arrow</i> a máximo 9 posiciones de distancia de un sustantivo. ....	188

Tabla 21. Lista de concordancias de la palabra <i>play</i> con sus equivalentes en portugués (énfasis añadido).....	190
Tabla 22. Resumen de la clasificación de técnicas de búsqueda en corpus.....	204
Tabla 23. Resumen de la clasificación de técnicas de búsqueda en Internet.....	254
Tabla 24. Resumen y códigos de nuestra clasificación conjunta de técnicas de búsqueda (en diccionarios, corpus e Internet). .....	258
Tabla 25. Resumen de las particularidades de las acciones de uso (de un diccionario) y la terminología utilizada para referirse a ellas (Wiegand 1998: 292, traducción propia del alemán).....	269
Tabla 26. Tipología de necesidades lingüísticas y conceptuales de distintos grupos de usuarios de diccionarios de lenguas para fines específicos (según Fuertes-Olivera y Arribas-Baño 2008: 3). .....	276
Tabla 27. Estrategias para resolver problemas de comprensión (Enríquez Raído 2011: 258). .....	415
Tabla 28. Estrategias para resolver problemas de formulación (Enríquez Raído 2011: 261). .....	415
Tabla 29. Resumen de las ventajas e inconvenientes de los instrumentos de recogida de datos. ....	428
Tabla 30. Ejemplo de estrategia de búsqueda para resolver un problema de recepción en LO.....	437
Tabla 31. Ejemplo de estrategia de búsqueda de un término original simple en un diccionario bilingüe + verificación del equivalente encontrado en un corpus comparable. ....	440
Tabla 32. Ejemplo de estrategia de búsqueda a partir de una hipótesis basada en la intuición + verificación (aceptación o rechazo del equivalente en LM). .....	441
Tabla 33. Ejemplo de búsqueda a partir de una o más hipótesis basadas en información relacionada del contexto + verificación (aceptación o rechazo del equivalente en LM). .....	444
Tabla 34. Muestra del resumen de la clasificación preliminar de estrategias de búsqueda del anexo 3.....	448
Tabla 35. Esquema del cuestionario. ....	456
Tabla 36. Prioridad dada por los traductores a los recursos según su utilidad para realizar una búsqueda onomasiológica .....	458
Tabla 37. Resumen de nuestra clasificación de estrategias de búsqueda en traducción, que incluye a las estrategias de búsqueda de tipo onomasiológico detectadas.....	553
Tabla 38. Propuesta para la elaboración de recursos orientados a los traductores.....	565

## ÍNDICE DE ILUSTRACIONES

Ilustración 1. Representación de la conexión que une el concepto con la palabra en sus dos direcciones C→P y P→C (Casares 1941: 51).....	24
Ilustración 2. Porción de la entrada de <i>food</i> relacionada con <i>fruit</i> en el <i>Roget's Thesaurus</i> .....	26
Ilustración 3. Entrada de la palabra <i>acid</i> y tabla relacionada en la página 14 del <i>Illustrated Reverse Dictionary</i> .....	27
Ilustración 4. Índice de tablas e ilustraciones en el <i>Illustrated Reverse Dictionary</i> .....	27
Ilustración 5. Entrada de la palabra <i>agreement</i> en el <i>Oxford Reverse Dictionary</i> .....	28
Ilustración 6. Entrada de la palabra <i>arrepentimiento</i> en la parte analógica del <i>Diccionario ideológico de la lengua española</i> .....	30
Ilustración 7. Entrada de la palabra <i>anillo</i> en el <i>Diccionario combinatorio práctico del español contemporáneo</i> .....	31
Ilustración 8. Ejemplo de búsqueda en el diccionario <i>DEBO</i> (Sierra 1996).....	33
Ilustración 9. Herramientas electrónicas del traductor en el proceso de traducción (Austermühl 2001: 13).....	39
Ilustración 10. Modelo de proceso de traducción, según Krings (1986a: 269).....	42
Ilustración 11. Modelos actuales de acceso al léxico en una traducción, según Séguinot (1996: 86).....	43
Ilustración 12. Modelo de competencia traductora, según PACTE (2001: 41).....	46
Ilustración 13. Modelo de competencia traductora, según PACTE (2003: 60).....	47
Ilustración 14. Implicaciones cognitivas en secuencias de acciones, según PACTE (2009: 223).....	48
Ilustración 15. Modelo de competencia traductora en el estudio TransComp (Göpferich 2009: 20).....	49
Ilustración 16. Representación de técnica de búsqueda en los recursos.....	67
Ilustración 17. Clasificación de diccionarios electrónicos según Lehr (1996: 315, traducida del alemán al inglés en De Schryver 2003:148).....	71
Ilustración 18. Representación de dos técnicas de búsqueda en un diccionario.....	105
Ilustración 19. Búsqueda de una expresión con similitud ortográfica en el diccionario <i>DUE</i> .....	109
Ilustración 20. Búsqueda de una expresión con similitud fonética en el <i>TLFi</i> .....	109
Ilustración 21. Presencia de todas las palabras en <i>OneLook Reverse Dictionary</i> .....	112
Ilustración 22. Presencia de alguna de las palabras en el diccionario <i>Wordsmyth</i> .....	113
Ilustración 23. Presencia de una palabra y ausencia de otra palabra utilizando operadores en el <i>CED</i> .....	114
Ilustración 24. Presencia de una secuencia exacta de palabras utilizando operadores en el <i>DUE</i> .....	115
Ilustración 25. Búsqueda de consultas en lenguaje natural en el <i>OneLook Reverse Dictionary</i> .....	116
Ilustración 26. Uso de filtros de categoría gramatical en el diccionario <i>Lexical Freenet</i> .....	117
Ilustración 27. Uso de filtros de área temática en <i>Cercaterm</i> .....	118

Ilustración 28. Uso de filtros de idioma en <i>IATE</i> .	119
Ilustración 29. Búsqueda alfabética inversa en la lista de entradas del <i>DRAE</i> .	121
Ilustración 30. Ejemplo de búsqueda en los campos de definiciones en <i>Termium</i> .	122
Ilustración 31. Navegación en las relaciones semánticas en <i>WordNet</i> .	123
Ilustración 32. Relación de hiponimia del concepto GENOME en <i>Genoma</i> .	124
Ilustración 33. Relaciones conceptuales en <i>OncoTerm</i> .	125
Ilustración 34. Búsqueda directa de relaciones semánticas en el diccionario <i>Wordsmyth</i> .	126
Ilustración 35. Búsqueda directa de relaciones semánticas en <i>Lexical Freenet</i> .	127
Ilustración 36. Búsqueda de relaciones sintagmáticas (colocaciones) en el diccionario <i>DiCoInfo</i> .	128
Ilustración 37. Colocaciones clasificadas por significados en el <i>DiCoInfo</i> .	129
Ilustración 38. Búsqueda de relaciones sintagmáticas (colocaciones) en <i>Just The Word</i> .	130
Ilustración 39. Búsqueda de alternativas a relaciones sintagmáticas (colocaciones) en <i>Just The Word</i> .	131
Ilustración 40. Búsqueda en el foro a través de la entrada de <i>agreement</i> en <i>WordReference</i> .	132
Ilustración 41. Búsqueda en un corpus interno en la base de datos <i>EOHS Term</i> .	133
Ilustración 42. Búsqueda en un corpus interno en <i>Just The Word</i> .	134
Ilustración 43. Búsqueda por exploración en el índice temático del <i>DRAE</i> .	135
Ilustración 44. Consulta directa con palabras clave en el índice del diccionario visual <i>Merriam Webster</i> .	135
Ilustración 45. Búsqueda en conexiones externas a otros diccionarios en línea en el <i>OneLook Reverse Dictionary</i> .	136
Ilustración 46. Resultados de uso en contexto en <i>FrameNet</i> .	138
Ilustración 47. Representación gráfica de las relaciones de <i>sequía</i> en <i>EcoLexicon</i> .	139
Ilustración 48. Resultado de una imagen en el diccionario visual de <i>Merriam Webster</i> .	141
Ilustración 49. Representación de técnica de búsqueda en un corpus.	176
Ilustración 50. Consulta de una combinación continua de formas, <i>Cámara de los Comunes</i> , en <i>CREA</i> .	179
Ilustración 51. Consulta de una combinación continua de expresiones mixtas en <i>Scientext</i> .	180
Ilustración 52. Búsqueda de una palabra con filtro de restricción de categoría gramatical en <i>Bwananet</i> .	183
Ilustración 53. Búsqueda de una palabra con filtro de restricción de categoría gramatical en <i>Bwananet</i> .	183
Ilustración 54. Resultado de concordancias ordenadas por el primer adjetivo a la izquierda de <i>rain</i> en <i>BNCweb</i> .	189
Ilustración 55. Generación y validación de una lista de términos en <i>TerminoWeb</i> .	191
Ilustración 56. Resultado de colocaciones de <i>software</i> en <i>WordSmith Tools 4</i> .	194

Ilustración 57. Resultados de la búsqueda de patrones de colocaciones de la palabra <i>software</i> en <i>WordSmith Tools 4</i> .....	194
Ilustración 58. Lista de sinónimos de <i>lluvia</i> en el contexto de sinónimos de <i>fuerte</i> ; concordancias de <i>aguacero</i> en el <i>Corpus del español</i> (de BYU). .....	196
Ilustración 59. Patrones de relaciones semánticas en <i>TerminoWeb</i> .....	197
Ilustración 60. Resultado de la búsqueda de patrones de similitud y sinonimia en un corpus virtual sobre cerámica en <i>TerminoWeb</i> .....	197
Ilustración 61. Lista de palabras modificadas por <i>college</i> y <i>university</i> en el <i>BNC</i> con la función <i>Sketch-Diff</i> de <i>Sketch Engine</i> .....	199
Ilustración 62. Resultado de relaciones sintácticas en el corpus <i>Scientext</i> .....	199
Ilustración 63. Resultado de lista de categorías gramaticales en <i>Phrases in English</i> . .....	200
Ilustración 64. Resultado de grafo generado por <i>Jaguar</i> a partir de un corpus sobre <i>cambio climático</i> .....	202
Ilustración 65. Resultados de información estadística de la palabra <i>ordenador</i> en <i>CREA</i> .....	203
Ilustración 66. Semantics-based characterization of metadata, según Kashyap et al. (2008: 30). ....	212
Ilustración 67. Pasos de evolución en la búsqueda (Vossen y Hagemann 2007: 268). .....	215
Ilustración 68. Resultados de la <i>Rueda de búsquedas</i> a partir de la palabra <i>sequía</i> en <i>Google</i> .....	220
Ilustración 69. Distribución geográfica de <i>computadora(s)</i> en <i>Diatopix</i> .....	222
Ilustración 70. Interfaz de búsqueda de <i>Copernic Agent Personal</i> . .....	223
Ilustración 71. Opciones de búsqueda avanzada de <i>WebCorp</i> .....	225
Ilustración 72. Concordancias generadas con <i>WebCorp</i> . .....	226
Ilustración 73. Interfaz principal de búsqueda de <i>WebCONC</i> .....	227
Ilustración 74. Restricción del corpus de búsqueda en <i>WebCONC</i> . .....	227
Ilustración 75. Introducción de términos de búsqueda con <i>KWiCFinder</i> . .....	229
Ilustración 76. Búsqueda avanzada en <i>Web Concordancer</i> . .....	231
Ilustración 77. Interfaz de búsqueda de <i>WWW Search Interfaces for Translators</i> .....	232
Ilustración 78. Interfaz de búsqueda de <i>Term-minator</i> .....	233
Ilustración 79. Búsqueda de <i>efecto invernadero</i> en la herramienta <i>Linguee</i> .....	234
Ilustración 80. Búsqueda de <i>greenhouse effect</i> en <i>Keybot</i> . .....	234
Ilustración 81. Búsqueda de <i>divisa</i> en el sitio web del Banco Central Europeo con <i>WeBiText</i> . .....	236
Ilustración 82. Búsqueda de combinaciones de formas exactas en <i>KWiCFinder</i> .....	240
Ilustración 83. Búsqueda de <i>Cytomegalovirus NEAR "is a"</i> en <i>Yahoo</i> .....	242
Ilustración 84. Comparación de resultados en <i>Google</i> y <i>KWiCFinder</i> . .....	246
Ilustración 85. Lista de palabras de un corpus sobre cerámica recuperado desde Internet con <i>WordSmith Tools 4</i> . .....	247
Ilustración 86. Resultado de colocaciones con <i>WebCorp</i> .....	249

Ilustración 87. Resultado de imágenes de <i>cigüeñal</i> en <i>Google</i> .....	251
Ilustración 88. Resultado de videos de <i>cigüeñal</i> en <i>Google</i> .....	251
Ilustración 89. Resultados estadísticos sobre <i>ordenador y computadora</i> en <i>Google trends</i> .....	252
Ilustración 90. Modelo de comunicación según Bergenholtz y Tarp (2003: 174).....	273
Ilustración 91. Clasificación de los diccionarios según las funciones lexicográficas (Bergenholtz y Tarp 2003: 182). .....	275
Ilustración 92. Modelo de proceso de traducción (Fata 2010: 88), según Muráth (2000: 159), basado en Gerzymisch-Arbogast (1996: 245). .....	277
Ilustración 93. Subcampos de la lexicografía en la intersección de la investigación en diccionarios y de los estudios de traducción (Fata 2010: 86). .....	277
Ilustración 94. El triángulo lexicográfico según Verlinde et al. (2010: 3).....	285
Ilustración 95. El continuo de necesidades de información (Varantola 2003b: 339).....	302
Ilustración 96. Modelo de uso del diccionario según Hartmann (1989c: 105). .....	323
Ilustración 97. Modelo de uso del diccionario según Bogaards (1993: 25; 2003: 29).....	324
Ilustración 98. Esquema de nuestro modelo de estrategia de búsqueda. ....	432
Ilustración 99. Representación de las necesidades o problemas.....	433
Ilustración 100. Transcripción en el programa <i>F4</i> con códigos de tiempo.....	477
Ilustración 101. Sincronización de las transcripciones y los archivos de sonido originales en <i>Atlas.ti</i> . .....	481
Ilustración 102. Asignación en el análisis con <i>Atlas.ti</i> del código LM.b.7.1 a la idea de consulta en un foro.....	488
Ilustración 103. Agrupaciones en familias de algunos códigos de nuestro análisis en <i>Atlas.ti</i> . .....	489
Ilustración 104. Fragmento de la red conceptual sobre las ventajas, desventajas y demandas en los recursos (nodo sobre las demandas de los traductores en Internet).....	492
Ilustración 105. Búsqueda de códigos y citas de nuestro análisis con la herramienta de búsqueda de <i>Atlas.ti</i> . .....	494
Ilustración 106. Distribución de la muestra según los años de experiencia. ....	496
Ilustración 107. Distribución de la muestra según la situación de empleo.....	496
Ilustración 108. Distribución de la muestra según la situación de los traductores que trabajan en plantilla.....	496
Ilustración 109. Distribución de la muestra según las especializaciones. ....	497
Ilustración 110. Distribución de la muestra según agrupaciones de las lenguas de trabajo de los traductores.....	499





## RESUMEN

En el grupo de investigación TecnoLeTTra estamos trabajando en el proyecto *ONTODIC: Metodología y tecnologías para la elaboración de diccionarios onomasiológicos basados en ontologías. Recursos terminológicos para la e-traducción*. Dicho proyecto consiste en elaborar una metodología sistemática para la creación de diccionarios terminológicos onomasiológicos, es decir, diccionarios que contengan terminología especializada en un ámbito de conocimiento y que permitan al usuario realizar consultas a partir del significado, y no solo a partir de la denominación.

Con la tesis doctoral *Estrategias de búsqueda onomasiológica en la actividad de traducción. Una ayuda al diseño de diccionarios terminológicos* pretendemos averiguar qué mecanismos de búsqueda resultarían útiles a los traductores para la consulta en un diccionario onomasiológico. La idea sería optimizar los resultados del diccionario onomasiológico mediante el diseño de un sistema de consulta que permita combinar en una sola herramienta todas las estrategias de búsqueda que use el traductor, lo que facilitaría la resolución de determinados problemas terminológicos.

Nuestro estudio empírico sobre estrategias de búsqueda onomasiológica en traducción se justifica de los análisis efectuados a lo largo de este trabajo en los que hemos detectado que la función onomasiológica de búsqueda en los diccionarios y las estrategias de búsqueda de tipo onomasiológico de los traductores no han recibido suficiente atención, en comparación con otros tipos búsquedas. Asimismo, la mayoría de los estudios analizados utilizan como sujetos a estudiantes de traducción, no a traductores profesionales, por lo que no podemos extrapolar directamente los resultados a la traducción profesional. No obstante, sí que vemos que algunos autores sugieren que los traductores tienen necesidades de naturaleza onomasiológica para acceder a las denominaciones que desconocen; también proponen que se precisa facilitar a los usuarios el acceso onomasiológico en los recursos.

La realización de entrevistas orales a un grupo de 48 traductores profesionales nos ha permitido aproximarnos a las estrategias de búsqueda onomasiológica utilizadas por los traductores en distintos recursos y las demandas o sugerencias propuestas por estos usuarios para adaptar los recursos a sus necesidades. A partir de esta información proporcionada por los traductores, hemos elaborado una clasificación de estrategias de búsqueda en traducción, en la que hemos identificado y ubicado a las estrategias de búsqueda onomasiológica. Finalmente, hemos aportado una serie de propuestas para la elaboración de recursos (onomasiológicos) orientados a los traductores.



## 0 Introducción

La terminología constituye uno de los elementos principales de estudio del trabajo de los traductores. El rápido avance de las nuevas tecnologías en los últimos años ha transformado las necesidades de traducción. La especialidad de los textos que se traducen aumenta considerablemente, así como la variedad temática de los mismos. Estas exigencias del mercado justifican la demanda por parte de los traductores de herramientas de trabajo que les resuelvan en el menor tiempo posible las dudas terminológicas que puedan plantearseles (Alcina 2008: 81).

Sin embargo, la falta de diccionarios o su pobre calidad obliga al traductor a ocuparse de la tarea terminológica con las consiguientes pérdidas de tiempo y eficacia en la traducción (Alcina 2009a: 33). Esta ausencia de diccionarios de calidad se debe en parte al alto coste que tienen las tareas de elaboración de un diccionario, especialmente en aquellas fases del trabajo que no están automatizadas como es el caso del análisis de conceptos y elaboración de definiciones.

En el grupo de investigación TecnoLeTTra<sup>1</sup> estamos trabajando en el proyecto *ONTODIC: Metodología y tecnologías para la elaboración de diccionarios onomasiológicos basados en ontologías. Recursos terminológicos para la e-traducción*<sup>2</sup>. Dicho proyecto consiste en elaborar una metodología sistemática para la creación de diccionarios terminológicos onomasiológicos, es decir, diccionarios que contengan terminología especializada en un ámbito de conocimiento y que permitan al usuario realizar consultas a partir del significado, y no solo a partir de la denominación.

Asimismo, estamos desarrollando el proyecto *ONTODIC II: Metodología y técnicas para elaborar diccionarios de colocaciones basados en ontologías. Recursos terminológicos para la e-Traducción*<sup>3</sup>.

La metodología y las técnicas desarrolladas se aplicarán al diseño de un diccionario onomasiológico de la cerámica industrial. Sin embargo, trataremos de automatizar al máximo las tareas para que puedan utilizarse en la elaboración de diccionarios de cualquier área temática.

---

<sup>1</sup> Grupo TecnoLeTTra: Tecnologías del lenguaje, la terminología y la traducción (<http://tecnolettra.uji.es>).

<sup>2</sup> Dirigido por la Dra. Alcina y financiado por el Gobierno de España (TSI2006-01911).

<sup>3</sup> Dirigido por la Dra. Alcina y financiado por el Gobierno de España (TIN2009-07690, subprograma TSI).

Entre los objetivos del proyecto ONTODIC se encuentran, por una parte, los relacionados con el desarrollo del análisis de conceptos terminológicos y su implementación, nos basábamos principalmente en las relaciones paradigmáticas entre los términos; asimismo, en el proyecto ONTODIC II estudiamos cómo representar en el diccionario las relaciones sintagmáticas, es decir, las colocaciones. Por otra parte, para que este nuevo modelo de diccionario resulte útil al mediador lingüístico se hace necesario analizar cuáles son sus necesidades y de qué modo un nuevo diccionario onomasiológico debe configurar los modos de acceso a la información contenida para responder a las demandas planteadas habitualmente u otras que sean requeridas.

Dentro de este segundo objetivo se enmarca el trabajo de la tesis doctoral *Estrategias de búsqueda onomasiológica en la actividad de traducción. Una ayuda al diseño de diccionarios terminológicos* con la que pretendemos averiguar qué mecanismos de búsqueda resultarían útiles a los traductores para la consulta en un diccionario onomasiológico. La idea sería optimizar los resultados del diccionario onomasiológico mediante el diseño de un sistema de consulta que permita combinar en una sola herramienta todas las estrategias de búsqueda que use el traductor, lo que facilitaría la resolución de determinados problemas terminológicos.

## **0.1 Hipótesis**

Partimos de la hipótesis de que, por un lado, los traductores necesitan hacer consultas onomasiológicas y, por otro, los recursos están dispersos y no puede accederse a ellos para hacer consultas onomasiológicas en un mismo entorno. Se parte de la base de que, hasta el momento, estos mediadores han trabajado principalmente con recursos semasiológicos para resolver sus dudas terminológicas, pero que utilizan estrategias de búsqueda de tipo onomasiológico. Los traductores tienen que diseñar determinadas estrategias para conseguir informaciones que los recursos no ofrecen directamente. Por ejemplo, una de las técnicas empleadas por los traductores es la consulta en corpus de especialidad o textos de Internet de términos, contextos y colocaciones.

La elaboración de diccionarios onomasiológicos haría posible al traductor realizar consultas a partir del significado para que el diccionario proporcionara automáticamente la denominación que representa a un determinado concepto.

## 0.2 Objetivos

En esta tesis nos hemos propuesto cumplir con los siguientes objetivos.

1) Confirmar si los traductores tienen necesidades que requieran la realización de búsquedas onomasiológicas, para acceder a las denominaciones a partir de los significados en los distintos recursos.

2) Detectar de qué manera realizan los traductores las búsquedas onomasiológicas en estos momentos. Por un lado, qué recursos y qué funciones se están implementando en los recursos (*técnicas de búsqueda*). Por otro lado, qué *estrategias de búsqueda* utilizan los traductores para recuperar la terminología que desconocen en los recursos.

3) Agrupar o clasificar las funciones de los recursos y las estrategias que resultan útiles a los traductores para realizar búsquedas onomasiológicas (*estrategias de búsqueda onomasiológica*). Creemos que este tipo de clasificaciones pueden ayudar en la elaboración de recursos diseñados para la búsqueda onomasiológica en traducción.

## 0.3 Estructura del trabajo y metodología

Para la consecución de los objetivos que acabamos de plantear, esta tesis doctoral consta de 3 partes y 5 capítulos. Además, incluimos 7 anexos en formato electrónico en el disco que adjuntamos con el trabajo.

En la **primera parte** presentamos los **antecedentes** del trabajo, que se componen de dos capítulos teóricos.

En el **primer capítulo**, *La búsqueda onomasiológica (del concepto a la denominación)*. *Los diccionarios onomasiológicos*, hacemos una breve revisión de los antecedentes de los diccionarios onomasiológicos, desde los inicios hasta nuestros días. Vemos que la búsqueda onomasiológica (del concepto a la denominación) en los diccionarios ha recibido un tratamiento deficiente, mientras que se ha favorecido la elaboración de diccionarios alfabéticos que permiten principalmente una búsqueda semasiológica de una palabra para acceder a su significado. Sin embargo, en este capítulo también observamos que las perspectivas de futuro sugieren que la transición de los diccionarios a formatos electrónicos debería facilitar la incorporación de búsquedas más avanzadas, en las que podrían incluirse las búsquedas onomasiológicas, además debería conseguirse que estos diccionarios onomasiológicos resulten útiles a los usuarios.

En el **capítulo dos**, *El proceso de traducción, la competencia terminológica del traductor y la relación entre traducción y terminología*, tratamos de indagar en qué

medida existe una necesidad onomasiológica en la actividad de traducción y hasta qué punto los traductores precisan la realización de labores de búsqueda en los recursos para resolver este y otros problemas. Observamos que el proceso de traducción se compone principalmente de dos fases, una de comprensión y otra de expresión. En la fase de *comprensión* la actividad del traductor es *semasiológica* porque el traductor busca el significado de determinados elementos del texto original; mientras que en la fase de *expresión* la actividad es *onomasiológica* porque el traductor busca en la lengua meta (LM) palabras o expresiones que reproduzcan el significado de la lengua origen (LO).

En este sentido, también observamos que en traducción existe una necesidad creciente de terminología, especialmente en el campo de la traducción especializada. Examinamos la aparición progresiva de competencias terminológicas dentro de la competencia traductora. Finalmente, estudiamos la necesidad del traductor de conocer los recursos y dominar las búsquedas de terminología en dichos recursos para llevar a cabo una búsqueda eficaz, que facilite la tarea de traducción.

A partir de estos antecedentes nos planteamos lo siguiente. Si los traductores tienen necesidades de distinto tipo, es de suponer que los recursos también deberán ser diferentes para satisfacer estas necesidades o problemas. Por tanto, presentamos una investigación en dos sentidos. Por un lado, debemos conocer qué funciones (o *técnicas de búsqueda*) ofrecen actualmente los recursos (tema que abordamos en la parte dos de este trabajo) y, por otro lado, de qué forma utilizan los traductores estas funciones (qué *estrategias de búsqueda* emplean) para satisfacer con éxito cada una de sus necesidades (en la tercera parte del trabajo).

La **segunda parte** del trabajo sobre **técnicas de búsqueda** se compone de un solo capítulo en el que se combina teoría y análisis empírico.

En el **capítulo tercero, *Técnicas de búsqueda en los recursos y nuestra propuesta de clasificación***, nos adentramos en el análisis de las funciones u opciones de búsqueda que permiten los recursos, lo que en esta tesis denominamos *técnicas de búsqueda*. Dividimos el capítulo en tres tipos de recursos: diccionarios, corpus e Internet. Por cada uno de estos recursos, abordamos primero estudios que han contribuido a la descripción de las características y las posibilidades en materia de búsquedas y, por ende, al desarrollo de diccionarios electrónicos, corpus e Internet. Si bien es cierto que en la bibliografía se han utilizado diferentes criterios de clasificación. En ocasiones también se han mezclado criterios y se han adoptado distintos puntos de vista para describir a estos recursos. Por ello, observamos una necesidad de profundizar respecto a cómo son las formas de

búsqueda en los recursos y cómo se pueden organizar. Analizamos recursos electrónicos, ya que como veremos incorporan las búsquedas más novedosas, frente a los tradicionales recursos en papel, en los que la búsqueda es más limitada.

Así, realizamos análisis empíricos de un conjunto de diccionarios electrónicos (Pastor y Alcina 2010, en prensa), corpus electrónicos (Pastor y Alcina 2009) y herramientas de búsqueda en Internet (Pastor y Alcina 2011) y clasificamos las técnicas de búsqueda que ofrecen siguiendo los mismos criterios de clasificación. Esto nos permite presentar al final del capítulo tres una propuesta conjunta de clasificación de técnicas de búsqueda en diccionarios, corpus e Internet.

Nuestra clasificación de técnicas de búsqueda se basa en tres ejes: la *consulta*, el *instrumento* y el *resultado*, que cubren las principales búsquedas que ofrecen los diccionarios, los corpus e Internet y las sistematiza. Es una clasificación que permite ser aplicada a la valoración de los recursos en general y de cualquier recurso en particular. Además es flexible, pues permite la incorporación de nuevos elementos.

La **tercera parte** del trabajo sobre **estrategias de búsqueda**, se compone de dos capítulos.

En el **capítulo cuatro**, *Estrategias de búsqueda en traducción y nuestra propuesta de clasificación*, también combinamos parte de teoría y análisis empírico. Mientras que en el capítulo tres nos centrábamos en el análisis de los recursos desde un punto de vista técnico (las técnicas de búsqueda), en el capítulo cuatro incorporamos nuevos factores que intervienen en una búsqueda y que no dependen tanto del recurso, sino del usuario. La combinación de todos estos elementos en una búsqueda es lo que denominamos *estrategia de búsqueda*.

Al principio de este capítulo, tratamos una serie de corrientes de investigación que abogan por un enfoque orientado al usuario. A continuación, examinamos un conjunto de estudios sobre las estrategias de búsqueda que realizan los traductores en distintos recursos, este análisis lo dividimos en diccionarios, corpus e Internet (en paralelo con el capítulo anterior), aunque, como veremos, en una misma estrategia de búsqueda pueden intervenir recursos de distinto tipo. En cada uno de los apartados hacemos primero una introducción a las estrategias de búsqueda de los traductores en cada uno de estos recursos y, después, en los subapartados resumimos uno a uno distintos estudios sobre estrategias de búsqueda de los traductores haciendo hincapié en los objetivos, los sujetos, las metodologías empleadas y sobre todo los resultados obtenidos. Al final de cada uno de los apartados sobre estrategias de búsqueda de los traductores en diccionarios, corpus e

Internet, ofrecemos recapitulaciones más breves que recogen las ideas más importantes del análisis.

De forma paralela al análisis de la bibliografía sobre estrategias de búsqueda de los traductores en diccionarios, corpus e Internet, sintetizamos los resultados del análisis de estos estudios (en el anexo 2), de manera que detectamos una serie de tendencias en las estrategias de búsqueda de los traductores que se repiten en la bibliografía. Finalmente, de la síntesis y ordenación de estas estrategias de búsqueda detectadas en los estudios, elaboramos una propuesta preliminar de clasificación de estrategias de búsqueda en traducción, que presentamos al final del capítulo cuatro.

Con todo ello, en el capítulo cuatro vemos, por un lado, en qué punto se encuentra la investigación en materia de estrategias de búsqueda en traducción y, por otro, clasificamos una serie de estrategias de búsqueda que se han detectado en la bibliografía, con lo que extraemos conclusiones que tomamos como punto de partida para el diseño de nuestra investigación empírica sobre estrategias de búsqueda onomasiológica en traducción.

En este punto nos gustaría aclarar al lector que los capítulos tres y cuatro de este trabajo no están pensados para ser leídos necesariamente de principio a fin. Los análisis y las descripciones de las clasificaciones que presentamos sobre *técnicas de búsqueda* y *estrategias de búsqueda* son bastante exhaustivos y detallados, ya que pretendíamos recopilar en un mismo trabajo todo el panorama sobre técnicas y estrategias de búsqueda que hemos observado. Asimismo, al final de cada uno de los apartados sobre el análisis de técnicas y estrategias de búsqueda en diccionarios, corpus e Internet incluimos recapitulaciones que resumen el análisis efectuado y, al final de cada uno de los capítulos, se encuentran nuestras propuestas de clasificación, que resumen los hallazgos encontrados en los análisis. Por tanto, en un principio, bastará con acceder a las recapitulaciones que resumen nuestros análisis, así como a nuestras dos propuestas de clasificación de técnicas de búsqueda y estrategias de búsqueda. No obstante, si el lector está interesado en profundizar en algunas de las técnicas o estrategias de búsqueda indicadas podrá también encontrar todas las descripciones en los apartados de análisis y explicación de técnicas y estrategias de búsqueda de estos capítulos.

Finalmente, en el **quinto capítulo**, presentamos nuestro *Estudio empírico sobre estrategias de búsqueda onomasiológica en traducción*. En los análisis efectuados a lo largo de este trabajo detectamos que la función onomasiológica de búsqueda en los diccionarios y las estrategias de búsqueda de tipo onomasiológico de los traductores no han recibido suficiente atención, en comparación con otros tipos búsquedas. Asimismo, la



mayoría de los estudios analizados utilizan como sujetos a estudiantes de traducción, no a traductores profesionales, por lo que no podemos extrapolar directamente los resultados a la traducción profesional. No obstante, sí que vemos que algunos autores sugieren que los traductores tienen necesidades de naturaleza onomasiológica para acceder a las denominaciones que desconocen; y también proponen que es necesario que se desarrollen recursos para facilitar a los usuarios el acceso onomasiológico a la terminología.

Estas premisas justifican que nos hayamos planteado un estudio empírico con el que pretendemos realizar una aportación a la investigación sobre las estrategias de búsqueda de los traductores, centrándonos en un tipo concreto de estrategias, las *estrategias de búsqueda onomasiológica*. En este estudio partimos de la base de que los traductores a veces se encuentran en la situación de que comprenden una idea o concepto de la LO, pero desconocen su denominación o equivalente en la LM. En este caso, no pueden buscar alfabéticamente la denominación en un diccionario y habrán tenido que recurrir a una serie de estrategias de búsqueda para encontrar la denominación. En nuestro estudio queremos contar con la experiencia de los traductores, que posiblemente han necesitado realizar este tipo de búsquedas, que hemos llamado *estrategias de búsqueda onomasiológica* (porque van del significado a la denominación).

Nuestra investigación sobre las estrategias de búsqueda onomasiológica de los traductores consta de dos fases. Primero, en un estudio piloto (Pastor 2008a, 2008b, 2010) distribuimos una encuesta a 15 traductores profesionales para evaluar qué preguntas y qué instrumento de medida pueden resultarnos útiles para recabar la información que necesitamos, además de realizar una primera aproximación a las estrategias de búsqueda onomasiológica. Segundo, presentamos nuestro estudio principal consistente en una entrevista oral que realizamos a un total de 48 traductores profesionales, con la que ahondamos en las estrategias de búsqueda onomasiológica utilizadas por los traductores en distintos recursos y las demandas o sugerencias propuestas por estos usuarios para adaptar los recursos a sus necesidades. A partir de la información proporcionada por los traductores, elaboramos una nueva clasificación de estrategias de búsqueda en traducción, en la que identificamos y ubicamos a las estrategias de búsqueda onomasiológica. Finalmente, hacemos una propuesta para la elaboración de recursos (onomasiológicos) orientados a los traductores.



## ABSTRACT

The TecnoLeTTra Team is working on the *ONTODIC Project: Methodology and technologies for the production of onomasiological dictionaries based on ontologies. Terminological resources for e-translation*. This project aims to develop a systematic methodology for creating onomasiological terminological dictionaries, namely, dictionaries that contain specialized terminology in a given field of knowledge and that allow users to query from a meaning to obtain a word, and not just from a word to obtain a meaning.

Our thesis entitled *Onomasiological Search Strategies in Translation Activity. An Aid for the Design of Terminological Dictionaries* intends to find out what search mechanisms would be useful to translators for querying an onomasiological dictionary. The idea is to improve the onomasiological dictionary results by designing a query system that allows the combination in one tool of all the search strategies that may be useful for translators and which would facilitate the solution of certain terminological problems.

Our empirical study on onomasiological search strategies in translation is justified by a previous analysis, which we carried out throughout this work. We observe that the onomasiological search functions in dictionaries and translators' onomasiological search strategies have not attracted enough attention in the literature as compared to other types of search strategies. Furthermore, most reviewed studies were conducted with translation students as subjects, not professional translators, so we cannot directly extrapolate the results to professional translation. However, some authors suggest that translators have onomasiological search needs to access unknown words from their meaning, while others also claim to provide users with onomasiological queries in resources.

Our oral interviews with 48 professional translators allowed us to collect the onomasiological search strategies used by translators in different resources, and to know their demands or suggestions in order to adapt the resources to translators' needs. From the information provided by the translators, we developed a classification of search strategies in translation, in which we also identify and place onomasiological search strategies. Finally, we make some proposals for the creation of (onomasiological) translation-oriented resources.



## 0 Introduction

Terminology is one of the main elements of study in translators' work. The rapid advance of new technologies in recent years has transformed translation needs. The number of specialized texts to be translated is steadily increasing, as is their thematic variety. These market requirements justify the fact that translators are demanding tools capable of solving their terminological doubts as quickly as possible (Alcina 2008: 81).

However as dictionaries are lacking or are of poor quality, translators are required to manage terminological tasks on their own, which implies loss of time and efficiency in translation (Alcina 2009a: 33). This lack of quality dictionaries is caused somewhat by the high cost of certain tasks involved in creating dictionaries, especially in those phases where work is not automatic, such as analyzing concepts and preparing definitions.

The TecnoLeTTra<sup>4</sup> Team is working on the *ONTODIC Project: Methodology and technologies for the production of onomasiological dictionaries based on ontologies. Terminological resources for e-translation*<sup>5</sup>. This project aims to develop a systematic methodology for creating onomasiological terminological dictionaries, namely, dictionaries that contain specialized terminology in a field of knowledge that allow users to query from a meaning to obtain a word, and not just from a word to obtain a meaning.

We are also developing the *ONTODIC II Project: Methodology and techniques for the production of collocations dictionaries based on ontologies. Terminological resources for e-translation*<sup>6</sup>.

The resulting methodologies and techniques may be applied to the design of an onomasiological dictionary of industrial ceramics. However, we will attempt to automate most tasks so that they can be used to develop dictionaries in any subject area.

Among the ONTODIC Project objectives we find, on the one hand, those related to the analysis of terminological concepts and their implementation, which relies mainly on paradigmatic relations between terms. Moreover, the ONTODIC II Project studies how to represent the syntagmatic relations in the dictionary; i.e. collocations. On the other hand, if our intention is for this new type of dictionary to be useful for linguistic mediators, it is

---

<sup>4</sup> TecnoLeTTra Team: Language, Terminology, and Translation Technologies (<http://tecnolettra.uji.es>).

<sup>5</sup> Directed by Prof. Alcina and funded by the Spanish Government (TSI2006-01911).

<sup>6</sup> Directed by Prof. Alcina and funded by the Spanish Government (TIN2009-07690, subprogramme TSI).

necessary to analyze what their needs are and how a new onomasiological dictionary must provide access to information in order to meet current demands, or any others that may be required.

Within this second objective, we have written the doctoral thesis *Onomasiological Search Strategies in Translation Activity. An Aid for the Design of Terminological Dictionaries*, which intends to find out which search mechanisms would be useful to translators for querying an onomasiological dictionary. The idea is to improve the onomasiological dictionary results by designing a query system that allows the combination in one tool of all the search strategies that may be useful for translators, which would facilitate the solution of certain terminological problems.

## **0.1 Hypothesis**

We assume that, on the one hand, translators need to perform onomasiological queries and, on the other, that resources are scattered and are not accessible to perform a complete onomasiological search in the same tool. We also believe that, to date, linguistic mediators have worked mainly with semasiological resources to solve their terminological doubts, and that they use onomasiological search strategies in these resources. Translators have to devise strategies to acquire certain information that resources do not offer directly. For example, some of the techniques used by translators are querying terms, contexts and collocations in specialized corpora or online texts.

The creation of onomasiological dictionaries would allow translators to query from the meaning so that the dictionary would automatically provide the word representing a particular concept.

## **0.2 Objectives**

In this thesis, we intend to meet the following objectives.

1) To check if translators have onomasiological needs that require conducting searches in different resources to access a word that expresses a specific meaning.

2) To identify how translators have performed onomasiological searches to date. On the one hand, which resources and which functions are implemented in the resources (*search techniques*). On the other hand, which *search strategies* are used by translators to retrieve unknown terminology in the current resources.

3) To collect or classify the resource functions and translators' strategies which are useful to conduct onomasiological searches (onomasiological search strategies). We

believe that this type of classification can help develop translation-oriented onomasiological resources.

### **0.3 Structure and methodology**

To achieve the aforementioned objectives, this thesis consists of three parts and five chapters. In addition, we include seven annexes in an electronic format on the disk enclosed with the thesis.

In the **first part**, we present the **background** of our work, which comprises two theoretical chapters.

In the **chapter one**, *The onomasiological search (from the meaning to the word). Onomasiological dictionaries*, we briefly review the history of onomasiological dictionaries since they emerged to the present-day. We observe that little attention has been paid to onomasiological searches (from the meaning to the word) in dictionaries, whereas the development of alphabetical dictionaries that mainly allow semasiological searches (from a word to access its meaning) has been encouraged. However, in this chapter we also note that future prospects suggest that the evolution of dictionaries to electronic formats should facilitate the integration of more advanced searches, which could also include onomasiological searches, as well as guaranteeing that these onomasiological dictionaries are useful for users.

In **chapter two**, *The translation process, translators' terminological competence, and the relationship between translation and terminology*, we attempt to examine to what extent there is an onomasiological need in translation activity, and to what extent translators are required to perform searches in resources in order to solve this and other problems. We observe that the translation process consists mainly in two phases: comprehension and expression. In the *comprehension phase*, translators' activity is *semasiological* because translators need to find the meaning of certain elements of the source text, while this activity is *onomasiological* in the *expression phase* because translators need to find words or expressions in the target language (TL) that express the meaning of the source language (SL).

In line with this, we also note that there is a growing need for terminology in translation, particularly in specialized translation. We examine the ongoing emergence of terminological competence in translation competence. Finally, we study translators' need to know the resources and how to master their searches for terminology in these resources in order to conduct an effective search that facilitates translation work.

From this background theory, we consider the following: if translators have different kinds of needs, we can assume that the resources required to meet these needs or problems should also be different. Therefore, we present a two-way investigation. We must firstly know what functions (or *search techniques*) are currently offered in the resources (which we address in the second part of this paper), and secondly, how translators use these functions (which *search strategies* they use) to successfully meet all their needs (in the third part of this work).

The **second part** of the thesis is on **search techniques** and consists in a single chapter that combines theory and empirical analyses.

In the **chapter three**, *Search techniques in resources and our proposed classification*, we enroll the analysis of the search features, or options, that are allowed in the resources, which we call *search techniques* in this thesis. We divide the chapter into three types of resources: dictionaries, corpora and the Internet. For all three resources, we first examine the studies that have contributed to not only describe the features and search possibilities, but hence to also develop electronic dictionaries, corpora and the Internet. Nevertheless, it is true that different classification criteria have been used in the literature. Sometimes different criteria have also been mixed, and these resources have been described from different points of view. Thus, we perceive the need for an in-depth study into what search options are like in the resources, and how they can be organized. We analyze electronic resources because, as we will see later, they integrate the newest searches as compared to traditional paper resources in which the search is more limited.

Therefore, we conduct an empirical analysis of a set of electronic dictionaries (Pastor and Alcina 2010, forthcoming), electronic corpora (Pastor and Alcina 2009) and Internet search tools (Pastor and Alcina 2011), and we classify all the search techniques that they incorporate by using the same classification criteria. This allows us to present a classification that combines search techniques in dictionaries, corpora and the Internet at the end of Chapter three.

Our categorization of search techniques is based on three elements: *query*, *resource* and *result*, which cover and systematize all the search possibilities we have observed in dictionaries, corpora and the Internet. This classification is valid for being applied to the assessment of any resources in general, and can be adapted to any specific resource. In addition, its structure is flexible in that new elements can be incorporated if needed.

The **third part** of the thesis, on **search strategies**, consists of two chapters.



In **chapter four**, *Search strategies in translation and our proposed classification*, we also combine both theory and empirical analyses. Whereas in Chapter three we focus on analyzing the resources from a technical perspective (search techniques), Chapter four incorporates new factors involved in a search, which do not depend on the resource, but on the user. The combination of all these elements in a search is what we call a *search strategy*.

To begin this chapter, we discuss some current research theories by advocating a user-oriented approach. Subsequently, we examine a set of studies on translators' search strategies in different resources; this analysis is divided into dictionaries, corpora and the Internet (in parallel with the previous chapter). However, and as we will see, the same search strategy may involve different types of resources. In each section, we first provide an introduction to translators' search strategies in each resource. Then we separately summarize different studies on translators' search strategies by focusing on objectives, subjects, methodologies and main results. At the end of each section on translators' search strategies in dictionaries, corpora and the Internet, we provide shorter summaries of the most important ideas of the analysis.

In parallel to the analysis of the literature on translators' search strategies in dictionaries, corpora and the Internet, we synthesize the results of this analysis in Annexe 2. Thus we detect some trends in translators' search strategies that are repeated in the literature. Finally, from the synthesis and arrangement of these search strategies identified in the studies, we put forward a preliminary proposal for the classification of search strategies in translation, which we present at the end of Chapter four.

Accordingly, we firstly observe in Chapter four the state of research on translation search strategies, and secondly we classify a number of search strategies that have been identified in the literature. This allows us to draw some conclusions which we take as a starting point to design our empirical research on onomasiological search strategies in translation.

At this point, we wish to explain to readers that Chapters three and four of this paper are not necessarily intended to be read from the beginning to the end. The presented analyses and descriptions of our classifications of *search techniques* and *search strategies* are quite exhaustive and detailed because we intended to include the whole scenario about the observed search techniques and strategies in the same paper. Besides at the end of each section on the analysis of search techniques and strategies in dictionaries, corpora and the Internet, we include summaries of all our analyses, while at the end of each chapter, our

proposed classifications summarize the findings from the whole analysis. Therefore, initially it would suffice to access the summaries of our analysis and our two proposed classifications of search techniques and search strategies. However, if readers are interested in further exploring any of the mentioned search techniques or strategies, they can also find all the descriptions in the sections of these chapters where we analyze and explain all the search techniques and strategies.

Finally, in **chapter five**, we present our *Empirical study on onomasiological search strategies in translation*. In the previous analyses throughout this thesis, we find that the onomasiological search function in dictionaries and translators' onomasiological search strategies have not received enough attention if compared to other search types. Furthermore, most of the analyzed research subjects are translation students, and not professional translators, so we cannot directly extrapolate the results of these studies to professional translation. However, some authors state that translators have onomasiological requirements to access unknown words, and they also suggest the need to develop resources which provide onomasiological access to terminology for their users.

These assumptions justify our empirical study in which we aim to contribute to research on translators' search strategies by focusing on a particular type of strategies, *onomasiological search strategies*. In this study, we assume that translators often find themselves in the situation where they understand an idea or concept of the SL, but do not know the word or its equivalent in the TL. In this case, the word cannot be searched for alphabetically in a dictionary, and translators may have to resort to a series of search strategies to find the correct word. In our study, we wish to know the experience of translators who may have had to perform such searches, which we call *onomasiological search strategies* (because they go from the meaning to obtain the word).

Our research on translators' onomasiological search strategies comprises two phases. Firstly in a pilot study (Pastor 2008a, 2008b, 2010), we handed out a questionnaire to 15 professional translators in order to assess which questions and which measuring instruments may prove useful to acquire the information we needed, and to also establish a first contact with translators' onomasiological search strategies. Secondly, we present our main study consisting of an oral interview which we conducted with 48 professional translators. Here we further examine translators' search strategies in a variety of resources, as well as the demands or suggestions made by these users to adapt the resources to their needs. From the information provided by the interviewed translators, we developed a new classification of search strategies in translation in which we identify and place

onomasiological search strategies. Finally, we make some proposals for the creation of (onomasiological) translation-oriented resources.



## **PARTE I. ANTECEDENTES**



# 1 La búsqueda onomasiológica (del concepto a la denominación). Los diccionarios onomasiológicos

An extreme but by no means exceptional case was the university student who, when asked in an examination to describe a thesaurus, replied: "Isn't it a kind of prehistoric animal?"

(Snell-Hornby 1987: 168)

La elaboración de diccionarios onomasiológicos no es una novedad. La tradición lexicográfica onomasiológica se remonta a varios siglos de antigüedad, tal y como han reflejado algunos autores que hacen un repaso de la lexicografía onomasiológica (Casares 1941; Hartmann 2005; Hüllen 2005; Martín Mingorance 1998).

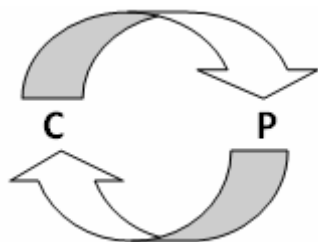
Sin embargo, estos estudios también reflejan que los diccionarios onomasiológicos que se han publicado tanto en el pasado como en el presente, se encuentran a la sombra de los diccionarios semasiológicos, cuyo orden alfabético parece que ha resultado más atractivo y que se ha impuesto en el mercado (véase Hüllen 2005: 4).

En este capítulo veremos brevemente qué se ha entendido tradicionalmente por diccionario onomasiológico, haremos una revisión de los tipos de recursos lexicográficos que se han considerado de naturaleza onomasiológica, como tesauros, diccionarios de sinónimos, diccionarios analógicos, diccionarios de ideas, etc. y citaremos algunos títulos tradicionales, disponibles sobre todo en formato papel, que se han destacado en la bibliografía. Asimismo, mencionaremos algunos proyectos de diccionarios onomasiológicos electrónicos más actuales y de recursos lingüísticos basados en ontologías.

Finalizaremos con unas reflexiones sobre aquellos planteamientos que no han recibido suficiente atención en la *onomasiografía* y cuál es el camino que debería seguirse en el futuro (Hartmann 2005).

## 1.1 Antecedentes de diccionarios onomasiológicos

Casares (1941: 51-86), en el discurso del acto de su recepción en la Real Academia Española, trató de convencer a los académicos de esta institución normalizadora sobre la necesidad de crear diccionarios que permitan buscar del concepto a la palabra, en lugar de la palabra al concepto. La relación entre palabra y concepto la representó como se muestra en la Ilustración 1.



**Ilustración 1. Representación de la conexión que une el concepto con la palabra en sus dos direcciones C→P y P→C (Casares 1941: 51).**

Para justificar la necesidad de diccionarios onomasiológicos, Hartmann (2005: 6) propone un ejemplo muy ilustrativo.

We were in a department store, and the directory boards on each floor helped us find the furniture department. Between the sofas and armchairs, there were some strange-looking seats (for one? for two?) –‘they’re called *love-chairs*’. This recent personal experience may well illustrate some of the points under discussion here, e.g. how do we find our way in unfamiliar territory? How do we name new objects? How is such information made available in reference works that we call onomasiological?

(Hartmann 2005: 6)

*Love-chair* podría definirse como un «canapé de dos asientos, especialmente aquel cuya forma permite a una persona sentarse enfrente de otra», pero ¿qué nombre recibe esta pieza de mobiliario en español? La respuesta es *confidente*. Sin embargo, en el *Diccionario de la Real Academia* podemos realizar la consulta de la palabra *confidente* para acceder a la definición, pero no podremos partir del conocimiento de su significado y realizar una búsqueda en el diccionario para acceder a la denominación.

Autores como Casares (1941), Martín Mingorance (1998) o Hartmann (2005) realizan una revisión de cuál ha sido la tradición de la lexicografía onomasiológica desde sus inicios, desde las primeras obras que se basaron en un esquema aristotélico (Martín Mingorance 1998: 264); explican que estas obras han evolucionado a lo largo de la historia y han influido unas en otras en la lexicografía de distintos países, como el *Roget’s International Thesaurus of English Words and Phrases* (publicado en 1852) que fue el referente de la lexicografía onomasiológica, por ejemplo en Alemania en el diccionario de Dornseiff (de 1933) *Der deutsche Wortschatz nach Sachgruppen*, en Francia en el *Dictionnaire idéologique* de Robertson (1859), o en España en el *Diccionario ideológico de la lengua española* de Casares (que se publicó por primera vez en 1942).

Según apunta Hartmann (2005) los tipos y subtipos de diccionarios onomasiológicos han recibido distintas denominaciones dependiendo de los países. Así, la tradición anglosajona ha publicado *thesauruses*, *dictionaries of synonyms*, *idioms*, *collocations*, *paronyms*, *pictorial dictionaries*, *rhyming dictionaries*; en inglés americano también se ha



hablado de *word-finder*, *word menu*, *descriptive* y *reverse dictionary*; *WordNet* (que trataremos más adelante) también está siendo la base de diccionarios de naturaleza onomasiológica (como señala Lew 2011: 240). Los diccionarios franceses descartan el nombre de *tesauro* y parece que se decantan más por *diccionarios analógicos*. Los *diccionarios de ideas* españoles, como los de Casares (1942), Corripio (1983) o Alvar Ezquerro (1998) también parece que siguen el modelo analógico francés.

A continuación, a modo ilustrativo, describiremos algunos de los diccionarios mencionados, como el *Roget's Thesaurus*, diccionarios inversos publicados en papel<sup>7</sup> como el *Illustrated Reverse Dictionary* y el *Oxford Reverse Dictionary*, el *Diccionario de ideas afines* de Casares y los diccionarios combinatorios *Redes* de Bosque (2004) y *Práctico* (Bosque 2006).

### 1.1.1 Roget's International Thesaurus of English Words and Phrases

Un ejemplo de tesauro en inglés es el *Roget's International Thesaurus of English Words and Phrases*, publicado por primera vez en 1852 en papel. Actualmente, el tesauro se encuentra disponible en línea<sup>8</sup>, lo que facilita su consulta.

Se establecen 6 clases de categorías principales (*relaciones abstractas, espacio, mundo material, intelecto, voluntad y sentimiento y moralidad*). Estas categorías generales incluyen, a su vez, subcategorías.

- Abstract relations: Existence, Resemblance, Quantity, Order, Number, Time, Power.
- Space: Motion, Change of place.
- Material World: Properties of Matter (Solidity, Fluidity, Heat, Sound, Light), Phenomena, Perceptions.
- Intellect: Acquisition, Retention, Communication of Ideas.
- Volition: Voluntary and Active Powers (Choice, Intention, Utility, Action, Antagonism, Authority, Compact, Property).
- Sentient and Moral Powers: Feelings, Emotions, Passions, Moral and Religious Sentiments.

---

<sup>7</sup> Asimismo, en el apartado 3.2.3 hacemos un análisis de las funciones de búsqueda en diccionarios electrónicos, donde incluimos el diccionario inverso en línea *One Look Reverse Dictionary*.

<sup>8</sup> Véase <http://www.bartleby.com/110/>

La consulta puede realizarse a partir del índice de categorías o el índice alfabético. Dentro de cada entrada del diccionario se remite a otras palabras relacionadas. Por ejemplo, con la consulta de la palabra *food* (que engloba cualquier tipo de comida, y se incluye a su vez en *eating*), si descendemos por el artículo de *food*, llegamos al apartado de *fruit*, en el que aparece una relación de frutas, como se puede observar en la Ilustración 2.

[FRUIT] figs, *figues* [F.], raisins, nuts, *noisettes* [F.], almonds: *les quatre mendicants* [F.]; apple, *pomme* [F.], pear, *poire* [F.], alligator pear, avocado, apricot, *abricot* [F.], peach, *pêche* [F.], plantain, banana, breadfruit, grapefruit, mango, mangosteen, grapes, pineapple, *ananas* [F.], walnuts, *noix* [F.], orange, lemon, lime, cherries, *cerises* [F.], watermelon, currants, cranberry, loganberry, blueberry, blackberry, gooseberry, whortleberry, huckleberry, raspberry, strawberry.

**Ilustración 2.** Porción de la entrada de *food* relacionada con *fruit* en el *Roget's Thesaurus*.

### 1.1.2 Diccionarios inversos

El *Illustrated Reverse Dictionary* es un diccionario en lengua inglesa, compuesto por 608 páginas. Contiene más de 400.000 palabras en más de 70.000 entradas. En este diccionario el acceso a las entradas es alfabético. Cada una de las entradas se considera una palabra clave que remite a otras palabras relacionadas semánticamente.

Por ejemplo, la palabra *acid* remite a tipos de ácidos y a palabras relacionadas con el concepto. En la Ilustración 3 observamos que *una botella en la que se introducen ácidos* se denomina *carboy*. Además, se remite a una tabla en la que se ofrecen sinónimos o características de distintos tipos de ácidos.

**acid** See chart, page 14  
 - acid-alkali indicator PHENOLPHTHALEIN; LITMUS PAPER; PH TEST  
 - acid as used for etching on a printing plate MORDANT  
 - acidlike substance, the universal solvent that dissolves anything, believed by alchemists to be possible and discoverable ALKAHEST  
 - acid or other destructive chemical CORROSIVE  
 - able to react as both an acid and a base, as some chemicals are AMPHOTERIC; AMPHIPROTIC  
 - bottle for acids and other corrosive liquids CARBOY  
 - chemical compound that reacts with acids to form salts BASE  
 - soluble base that can neutralize an acid ALKALI

ACIDS	
acetic acid	found in vinegar
amino acid	basis of proteins
ascorbic acid	vitamin C
bile acid	active in absorbing fat from intestine
carbolic acid	phenol; used in disinfectant soap
citric acid	found in lemons and other citrus fruits
conjugate acid	from addition of a proton to a base
deoxyribonucleic acid	DNA; basis of chromosomes, and hence of genetic transmission
essential fatty acid	vitamin F
fatty acid	associated with fat in the body
folic acid	vitamin B <sub>9</sub> or vitamin M
formic acid	naturally occurring in ants
free fatty acid	found in the blood
hydrochloric acid	spirits of salt; found in dilute form in the digestive juices; wide industrial application

Ilustración 3. Entrada de la palabra *acid* y tabla relacionada en la página 14 del *Illustrated Reverse Dictionary*.

Otra de las formas de búsqueda en este diccionario es a través de tablas e imágenes. Como se puede ver en la Ilustración 4, en las primeras páginas del diccionario aparece un índice de imágenes y tablas ordenadas alfabéticamente por palabras clave. Por ejemplo, si buscamos la imagen que se corresponde con la palabra clave *flowers*, encontraremos en la página 209 del diccionario una imagen de una flor en la que se señalan los nombres de todas sus partes.

Charts and Illustrations		
Quick-Reference Features That Pinpoint the Word You Need		
Page numbers in regular type refer to charts. Page numbers in <b>bold</b> type refer to illustrations.	<b>C</b> Cakes, Cookies, and Biscuits 83 Camels <b>84</b> Canonical Hours 85 Car <b>86</b> Card Games 87 Castle <b>88</b> Cats <b>90</b> Chameleon <b>94</b> Cheese <b>96</b> Chemistry Terms 98 Church <b>101</b> Classification 104 Clergymen and Clergywomen 105 Clerical Clothing <b>106</b> Clothes 108 Clouds <b>110</b> Coins 112	<b>F</b> Fabrics 192 Feather <b>195</b> Fencing Terms 197 Feudal and Medieval Terms 199 Figures of Speech and Rhetorical Devices 200 Films and Filming Terms 202 Fish <b>205</b> Fish Dishes 204 Fishing Terms 206 Flags <b>207</b> Flowers <b>209</b> Football <b>212</b> Footprints <b>213</b> Fractures <b>216</b> French Terms 218 Frogs <b>219</b>

Ilustración 4. Índice de tablas e ilustraciones en el *Illustrated Reverse Dictionary*.

Otro ejemplo de diccionario inverso en inglés es el *Oxford Reverse Dictionary*. Contiene 416 páginas y 31.000 entradas. La presentación de este diccionario también es alfabética. Sin embargo, está compuesto por palabras clave que permiten acceder a otras palabras relacionadas por medio del significado.

Por ejemplo, como se puede ver en la Ilustración 5, la entrada de la palabra *agreement* contiene una serie de definiciones, entre las que se incluye la del término *settlement* (*agreement reached by parties to a dispute or lawsuit*).

**agreement** • apprenticeship agreement: **indentures** • formal agreement: **compact, pact** • general agreement: **consensus** • interim agreement: **modus vivendi** • legally enforceable agreement: **contract** • agreement between the pope and a secular state: **concordat** • agreement between states: **treaty** • imagined agreement between members of a society to forego certain personal freedoms for the general good: **social contract** • agreement to suspend hostilities for a stated period: **armistice** • agreement reached through concessions: **compromise** • agreement reached by parties to a dispute or lawsuit: **settlement** • informal treaty: **convention** • informal understanding reached between states: **entente** • draft of an agreement: **protocol** • bargaining to reach an agreement: **negotiation** • bargains made to mutual advantage: **horse-trading** • demand that must be satisfied before agreement is reached: **stipulation** • condition attached to an agreement: **proviso** • exceptions specified in an agreement: **saving clause** • proviso added to an agreement:  **rider** • in full agreement: **unanimous** • one side to an agreement: **party** • denoting an agreement between two parties: **bilateral** • denoting an agreement between three or more parties: **multilateral** • person forwarding an agreement though not party to it: **intermediary, mediator** • breach of an agreement's terms: **infringement** • formal announcement of intention to end an agreement: **notice** • reject an agreement already made: **renounce, repudiate, rescind**

Ilustración 5. Entrada de la palabra *agreement* en el *Oxford Reverse Dictionary*.

### 1.1.3 Diccionario ideológico de la lengua española

La primera publicación del *Diccionario ideológico de la lengua española* fue en 1942. Tal y como afirma Soler (1999: 29), «la novedad que supone el diccionario ideológico de Casares es que sigue un punto de vista onomasiológico frente a los diccionarios tradicionales, que siguen un proceso semiológico. Es decir, va del significado al significante, puesto que su finalidad es facilitar instrumentos para construir mensajes, al contrario del resto de diccionarios, que son instrumentos para descifrarlos».

El diccionario está dividido en tres partes:

La parte sinóptica en la que se estructura la obra en 38 categorías temáticas a las que remiten cada una de las entradas de la parte analógica.

La parte analógica (acceso onomasiológico a la información), en la que se recopilan las palabras por grupos. El diccionario presenta 2.000 grupos que contienen unas 80.000 palabras.

La parte alfabética (acceso semasiológico), con un diccionario que define más de 100.000 voces, pero también marca con un asterisco (\*) las palabras que pueden consultarse en la parte analógica.

Las 38 categorías que se incluyen en la parte sinóptica del diccionario son las siguientes:

1. Religión	20. Colocación
2. Física y química	21. Tiempo
3. Geografía, astronomía y meteorología	22. Cantidad
4. Geología, mineralogía y minería	23. Inteligencia
5. Botánica	24. Juicio, valoración
6. Zoología	25. Voluntad
7. Anatomía	26. Conducta
8. Fisiología	27. Acción
9. Alimentación	28. Comunicación de ideas y sentimientos
10. Vestido	29. Arte – Lenguaje
11. Vivienda	30. Nación
12. Medicina	31. Costumbre
13. Sensibilidad	32. Derecho y justicia
14. Sentimiento	33. Propiedad
15. Existencia y cambio	34. Milicia
16. Relación, orden y causalidad	35. Comercio, banca, bolsa
17. Espacio y geometría	36. Agricultura
18. Forma	37. Zootecnia
19. Movimiento	38. Transportes

**Tabla 1. Categorías de la parte sinóptica del *Diccionario ideológico de la lengua española*.**

En cada una de las categorías, se establecen dos columnas en las que se contraponen conceptos antónimos o asociados. Por ejemplo, la palabra *belleza* pertenece a la categoría 24 (*juicio, valoración*) y aparece en contraposición a *fealdad* o *ridiculez* y relacionada con *elegancia, estimación* o *excelencia*.

En la parte analógica, las palabras se distribuyen por grupos. Los epígrafes de los grupos están ordenados alfabéticamente y unos grupos remiten a otros grupos relacionados. Dentro de cada grupo, las palabras están clasificadas por categorías gramaticales, primero aparecen los sustantivos y sus derivados, después los verbos, adjetivos, adverbios y modos adverbiales, preposiciones e interjecciones.

El vocablo que encabeza al grupo aparece destacado en mayúsculas. Debajo, en negrita, aparecen términos asociados. Estos términos asociados contienen, a su vez, otros términos asociados. Por ejemplo, como se puede observar en la Ilustración 6, la palabra *arrepentimiento* se incluye en las categorías 14 y 26 (*sentimiento* y *conducta*). Debajo de esta palabra aparecen en negrita *aflicción*, *penitencia* o *desesperanza*. Debajo de *aflicción* encontramos *compunción*. Debajo de *penitencia* aparece *remordimiento*. En relación con *desesperanza* se encuentran expresiones como *acto de contrición*, pero también verbos como *arrepentirse*, *lamentar*, y adjetivos como *compungido*.

ARREPENTIMIENTO (14, 26)	arrepentirse	arrepentido
arrepentimiento	dolerse	repiso
pesar	compungirse	arrepiso
dolor	pesarle a uno	pesaroso
sentimiento	concomerse	pesante
aflicción	recomerse	contrito
compunción	deplorar	compungido
penitencia	lamentar	atrito
contrición	llorar	penitente
atrición	sentir	magdalena
remordimiento	corroer	remordedor
escarabajeo	acusar la conciencia	remordiente
desesperanza	escarabajar la conciencia	lamentable
voz de la conciencia	morderse las manos	deplorable
gusano de la conciencia	llorar con lágrimas de sangre	
acto de contrición	compungir	ARREPENTIRSE
golpe de pecho		(V. Arrepentimiento)

**Ilustración 6.** Entrada de la palabra *arrepentimiento* en la parte analógica del *Diccionario ideológico de la lengua española*.

### 1.1.4 Diccionarios combinatorios

El diccionario *Redes. Diccionario combinatorio del español contemporáneo* (Bosque 2004) se define en su prólogo de la siguiente manera:

Este diccionario es combinatorio, pero se fundamenta en las relaciones semánticas que existen entre las palabras. Especifica en cada entrada las restricciones semánticas que las palabras se imponen unas a otras. El aspecto que resulta quizás más sorprendente de esta información es que no aparece en los diccionarios, pero a la vez es absolutamente esencial para el conocimiento del idioma, tanto el que corresponde a un hablante nativo como el que intenta adquirir el que lo estudia como segunda lengua. Así pues, este diccionario se diferencia de los de ideas afines y de los ideológicos en que no analiza el mundo a través del léxico, sino el léxico mismo a través de la relación que existe entre el significado de las palabras y la forma en que las combinamos.

(Bosque 2004: 6)

Cada una de las entradas de este diccionario da acceso a sus colocaciones típicas y añade o explica el significado que refleja dicha combinación. El *Diccionario combinatorio práctico del español contemporáneo* (Bosque 2006) se desarrolló a partir del diccionario *Redes* y recoge las combinaciones más habituales de las palabras. Estas combinaciones se agrupan por significados y, a su vez, dentro de cada significado por categorías gramaticales. También se incluyen expresiones o locuciones. Como ejemplo presentamos la entrada de *anillo*<sup>9</sup> en la Ilustración 7.

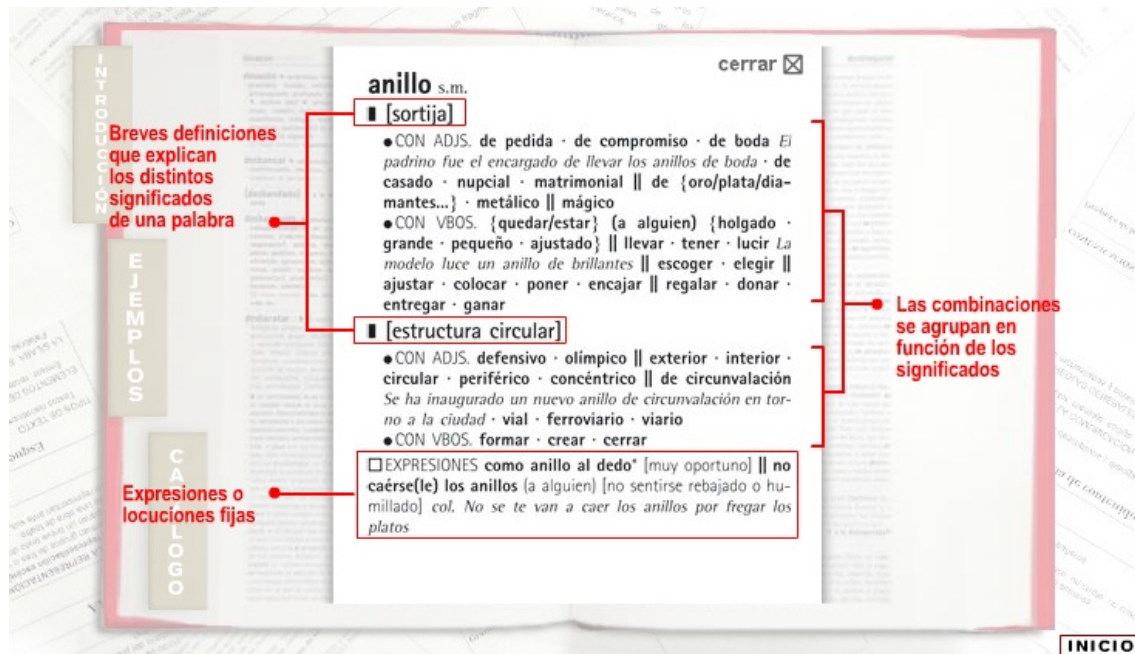


Ilustración 7. Entrada de la palabra *anillo* en el *Diccionario combinatorio práctico del español contemporáneo*.

<sup>9</sup> Ilustración obtenida de la página <http://diccionariopractico.grupo-sm.com>



## **1.2 Proyectos de investigación sobre diccionarios onomasiológicos electrónicos**

En los últimos años los proyectos de diccionarios onomasiológicos tratan de aprovechar las posibilidades que ofrece el medio electrónico (Hartmann 2005), ejemplo de ello son *Hideal* y *DEBO*.

*Hideal* es una herramienta que ha diseñado el Grupo de Estructuras de Datos y Lingüística Computacional del Departamento de Informática y Sistemas de la Universidad de Las Palmas de Gran Canaria<sup>10</sup>, y que consiste en un manejador informático de diccionarios ideológicos (también de sinónimos y tesauros) (Santana et al. 1997).

*El Diccionario Electrónico para la Búsqueda Onomasiológica (DEBO)*, desarrollado por el Instituto de Ingeniería de la Universidad Nacional Autónoma de México, dirigido por Sierra (1999; 2000; Sierra y McNaught 2000) es un diccionario onomasiológico electrónico, que describimos a continuación.

### **1.2.1 Diccionario Electrónico para la Búsqueda Onomasiológica**

Según las propias palabras de Sierra (1996) «el DEBO es una interfaz que, de manera general: lee la información introducida por el usuario; identifica las “palabras clave” en una primera base de datos; determina, en otra base de datos, una lista de términos relacionados con cada palabra clave; y obtiene, con un análisis de frecuencias, el o los términos más probables».

El *DEBO 1* permite buscar 33 términos del área de desastres por medio de la introducción de palabras clave. El sistema consta de un archivo inverso o base de datos que relaciona el índice de posibles palabras clave de la búsqueda introducida por el usuario con un índice de paradigmas predeterminado que agrupa las palabras clave en conjuntos semánticos. A su vez, el índice de paradigmas se asocia a la base de datos de términos (Sierra 1996).

En cuanto a la interfaz de búsqueda, consta de varias opciones que permiten seleccionar las necesidades de consulta del usuario. La primera opción facilita la selección del idioma en que se realiza la búsqueda. Esta opción también prevé el caso de que la base de datos se elabore para terminologías multilingües.

Otra de las opciones se refiere al área de conocimiento en que se realiza la búsqueda.

---

<sup>10</sup> Véase la página <http://www.gedlc.ulpgc.es/hideal.htm>



También existe la posibilidad de seleccionar cualquier tema, si el concepto puede englobarse en varias áreas.

Una vez seleccionadas las opciones anteriores, el usuario describe el concepto que está buscando con el uso de términos (palabras clave) que se relacionan con el concepto que quiere expresar. Cuando se hace clic en el botón de buscar, el sistema advierte si el concepto que se está buscando se incluye en varios temas, por lo que el usuario podrá elegir uno para acotar la búsqueda.

Por último, la herramienta muestra una lista de posibles términos que se corresponden con las palabras clave introducidas. Estos términos se dividen, según Sierra (1996), en «términos cabezas de familia o *header term*» y «términos relativos o *related terms*». Una vez encontrado el término que estábamos buscando, otra de las opciones del programa es la de acceder a la definición del mismo para asegurarnos de que el resultado obtenido ha sido satisfactorio.

En la Ilustración 8, se busca en el diccionario *DEBO*<sup>11</sup>, en lengua inglesa y en el ámbito del clima, la denominación utilizada para designar a un *fuerte movimiento de viento*. En la columna de la izquierda se pueden observar los *header terms* y en la derecha se encuentran los *related terms* de la palabra *wind*. Por último, el diccionario incluye debajo una definición de la palabra seleccionada, en este caso *wind*.

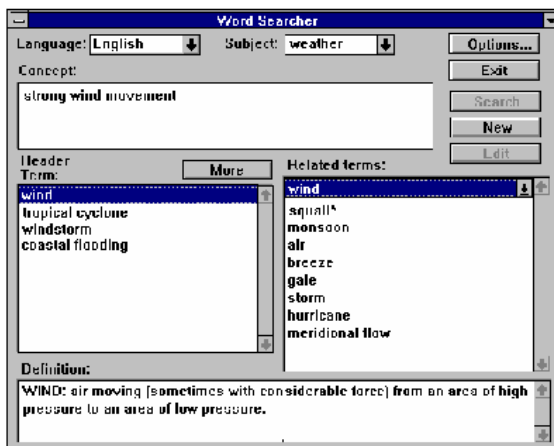


Ilustración 8. Ejemplo de búsqueda en el diccionario *DEBO* (Sierra 1996).

<sup>11</sup> El ejemplo se ha tomado del artículo de Sierra (1996). No hemos tenido acceso a la consulta en el diccionario *DEBO*.

### **1.3 Recursos lingüísticos basados en ontologías**

La Ingeniería Ontológica se está aplicando actualmente en el desarrollo de sistemas basados en el conocimiento, en ámbitos de la informática en general y también la terminografía, que permiten el acceso a la información tanto por humanos como el procesamiento del lenguaje natural por máquinas.

Se han observado las aplicaciones de las ontologías en la elaboración de bases de datos basadas en conocimiento, como en los proyectos *OncoTerm* (Faber 2002b, 2009; López Rodríguez et al. 2006a; López Rodríguez et al. 2006b), *Genoma-KB* (Cabré et al. 2004a; Cabré et al. 2004b; Feliu et al. 2002) y *ONTODIC* (Alcina 2009a), también en la traducción automática, como en el proyecto *Mikrokosmos* (Mahesh 1996; Mahesh y Nirenburg 1995a, 1995b), o en el desarrollo de la Web Semántica (véase Alcina 2009a). Los bancos de conocimiento *OncoTerm* y *Genoma* los describimos en la parte II de este trabajo.

### **1.4 Perspectivas para el futuro de los diccionarios onomasiológicos**

Por último, incluimos una serie de sugerencias que según Hartmann (2005: 15-16) deberían contemplarse para el desarrollo de la lexicografía onomasiológica. En relación con la práctica lexicográfica:

- mayor concienciación sobre los diccionarios por parte de editores, profesores y usuarios;
- mayor conocimiento sobre los procesos de compilación lexicográficos;
- mayor información en los catálogos y en las bibliografías sobre productos, funciones y contenidos;
- mayor reconocimiento en el *marketing* a los compiladores y sus ideas;
- mayor investigación y evaluación de formatos y tecnologías alternativos;
- en la práctica comercial, los desarrolladores de diccionarios deberían ponerse de acuerdo en las denominaciones empleadas para etiquetar y comercializar sus productos.

También explica que en la lexicografía teórica sería deseable mejorar la investigación en algunas áreas que se encuentran desatendidas.

- se precisa mayor investigación sobre las relaciones genealógicas en y entre las tradiciones;
- existen pocos estudios estructurales sobre diccionarios monolingües e interlingua;
- no existen estudios actualizados sobre tipologías de obras onomasiológicas más allá de los tesauros;
- existen enfoques y metodologías sin explorar, como la perspectiva del usuario;
- es necesario estudiar las relaciones entre la lexicografía y otras disciplinas;

- se han producido avances importantes en las tecnologías de la información, pero su aplicación está resultando más complicada de lo que predijeron los expertos;
- se precisan bibliografías de diccionarios onomasiológicos mejores, más exhaustivas y para todas las lenguas.

En relación con la perspectiva del usuario, en un estudio sobre la función onomasiológica de los diccionarios (Anaya Revuelta 2005: 25), el análisis de una serie de diccionarios ideológicos sugiere que la búsqueda en estos recursos debe partir del conocimiento que tiene el usuario sobre la organización del diccionario y de la intuición de la lengua, aunque parece que esto no ha propiciado la popularidad del diccionario ideológico.

Para cambiar esta situación, en el estudio de Anaya Revuelta (2005: 25) se plantea la necesidad de definir qué tipos de relaciones de significado establecen los usuarios entre las palabras, es decir, «por qué llegamos a averiguar la palabra que buscamos (formatear) por una vía (informática) y no por otras (disco, disquete), etc.».

### **1.5 Recapitulación**

En este apartado hemos hecho una pequeña revisión sobre el estado de la investigación en materia de diccionarios onomasiológicos desde sus inicios, pasando por el presente y las perspectivas de futuro. Hemos visto que aunque los diccionarios onomasiológicos han despertado el interés de los investigadores desde los comienzos de la lexicografía, estos han recibido un tratamiento insuficiente. Especialmente en la práctica editorial, abundan los diccionarios semasiológicos cuyo orden alfabético (aparentemente) ha resultado más práctico (Hüllen 2005: 4).

El escaso desarrollo de los diccionarios onomasiológicos puede deberse a que tradicionalmente en los diccionarios en papel difícilmente podía ofrecerse una búsqueda que no fuera la de localizar una palabra en un índice alfabético. Los intentos de diseñar modos de acceso al vocabulario de una lengua de forma más avanzada consistieron en la organización de índices analógicos, cuya consulta al parecer ha resultado compleja a los usuarios.

No obstante, el progreso tecnológico parece que aporta nuevas soluciones en la elaboración de diccionarios onomasiológicos de consulta más fácil por medios electrónicos, aunque estos avances no se están implantando con la rapidez que cabría esperar. En este sentido, parece que el futuro de la investigación en diccionarios onomasiológicos debería encaminarse, por un lado, a aprovechar las ventajas que ofrece el

formato electrónico sobre todo en materia de búsquedas avanzadas y, por otro, al diseño de estos diccionarios desde la perspectiva de los usuarios que los vayan a consultar para que resulten verdaderamente útiles.

## 2 El proceso de traducción, la competencia terminológica del traductor y la relación entre traducción y terminología

Sometimes when I translate, I am doping to release better words from my unconscious, my memory hold. I go on repeating the start of a sentence and stop, expecting something fresh to emerge, to pop out. Sometimes it does; it's a relief, and I start my smirk. More often, it doesn't, and I despair.

(Newmark 1993: 79-80)

La traducción es, según palabras de García Yebra (1984: 30): «enunciar en otra lengua [o lengua meta] lo que ha sido enunciado en una lengua fuente [o lengua original], conservando las equivalencias semánticas y estilísticas»<sup>12</sup>.

En este capítulo explicaremos, en primer lugar, aquellos aspectos del proceso de traducción y la competencia traductora que están relacionados o pueden ser útiles de cara a la competencia terminológica. Después, abordaremos la relación que existe entre la traducción y la terminología, que viene motivada por unas necesidades terminológicas que hacen que el traductor deba servirse de la terminología para llevar a cabo con éxito la actividad de traducción. Finalmente, veremos qué recursos terminológicos utilizan los traductores para realizar las labores de búsqueda y gestión terminológica que precisan.

Con ello, trataremos de averiguar en qué medida existe una necesidad onomasiológica en la actividad de traducción y hasta qué punto los traductores precisan la realización de labores de búsqueda en los recursos para resolver este y otros problemas.

### **2.1 El proceso de traducción y los problemas de traducción**

La traducción, como proceso mental, según autores como García Yebra (1984) y Hurtado Albir (2001), es una actividad que desarrolla un sujeto, el traductor, para la cual «ha de poseer ciertos conocimientos y habilidades (la competencia traductora) y ha de efectuar un proceso mental complejísimo, en el que intervienen múltiples operaciones cerebrales, y cuyos procesos básicos son la comprensión y la reexpresión» (Hurtado Albir 2001: 37).

---

<sup>12</sup> En lo sucesivo, emplearemos lengua meta (o LM) y lengua original (o LO).

Asimismo, García Yebra (1984: 30) explica que el proceso de traducción se compone al menos de dos fases: una fase de *comprensión* del texto original y una fase de *expresión* del contenido del texto original en la LM. Para este autor, en la fase de *comprensión* la actividad del traductor es *semasiológica* porque el traductor busca el contenido o el sentido del texto original; mientras que en la fase de *expresión* la actividad es *onomasiológica* porque el traductor busca en la LM palabras o expresiones que reproduzcan el contenido de la LO.

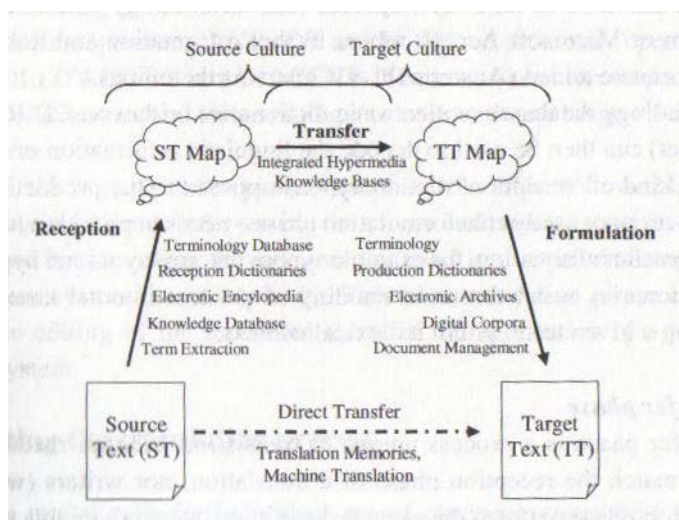
Por su parte, Delahaut y Demaecker (2002: 298) también señalan que en las etapas de traducción de un texto, la de lectura del texto original es *semasiológica*, ya que se parte de una palabra para extraer el significado, mientras que en la fase de reexpresión en la LM el tratamiento es *onomasiológico* porque se parte de la idea para encontrar una palabra o expresión: «[...] la lecture (qui est une démarche sémasiologique: partir du mot pour en extraire le sens), la compréhension du message et la ré-expression (qui est une démarche onomasiologique: partir de l'idée pour en trouver le mot ou l'expression)».

En este sentido, estos autores indican que el traductor deberá efectuar búsquedas documentales y terminológicas, que son las que marcan la diferencia entre una buena y una mala traducción, sobre todo en lo que respecta a la búsqueda del *término exacto* (*terme juste*), imprescindible para evitar el uso de paráfrasis innecesarias a falta de un término adecuado.

A su vez, según Hurtado Albir (2001: 367-368) algunos autores han identificado, además de estos procesos básicos, una *fase intermedia de carácter no verbal* o *fase no verbal*, que esta autora llama también *desverbalización* (denominación que toma de Seleskovitch y Lederer 1989) o *representación semántica* (por Bell 1991), y también en la traducción escrita una fase de *verificación* (por Delisle 1980).

En relación con las tecnologías, Austermühl (2001: 13) distingue tres fases que se encuentran presentes en el proceso de traducción, donde el traductor puede hacer uso de herramientas electrónicas: la fase de *recepción*, la fase de *transferencia* y la fase de *formulación*. En la recepción el traductor se forma un concepto mental del texto, este autor lo denomina *mapa del texto origen*, que está influenciado por factores contextuales, intertextuales, situacionales e individuales. El objetivo en esta fase es comprender el texto origen, lo que puede implicar dificultades como la obtención de conocimientos temáticos, así como la identificación de la terminología especial del texto. Las bases de datos y los diccionarios electrónicos pueden emplearse para decodificar la información lingüística del texto.

La fase de transferencia según este autor requiere conocimientos multiculturales contrastivos y gran habilidad de comunicación intercultural. En la fase de formulación el uso de diccionarios y bases de datos también pueden ayudar al traductor con las relaciones sintagmáticas provistas en guías de estilo o diccionarios de colocaciones. Según este autor, si los diccionarios proporcionan este tipo de información se convierten en recursos de asistencia a la producción de gran valor para el traductor. También los corpus e Internet pueden resultar de utilidad para verificar determinadas expresiones en LM. Las memorias de traducción permiten automatizar todo el proceso de traducción. En la Ilustración 9 mostramos la representación, según este autor, de las herramientas electrónicas del traductor en el proceso de traducción.



**Ilustración 9. Herramientas electrónicas del traductor en el proceso de traducción (Austermühl 2001: 13).**

### 2.1.1 Investigaciones empíricas sobre el proceso de traducción

Desde la segunda mitad de la década de los 80, a partir de la aparición del estudio de Krings (1986b) *Was in den Köpfen von Übersetzern vorgeht*, se llevan a cabo investigaciones empíricas que tratan de indagar sobre los procesos mentales del traductor, los contenidos de la *caja negra* (*black box*) de estos sujetos.

La mayoría de estos estudios utiliza la metodología de los Think-Aloud Protocols (TAP)<sup>13</sup>, por medio de la cual los sujetos del estudio traducen un texto y al mismo tiempo

<sup>13</sup> Para más información, véase Kussmaul (1995) y Kussmaul y Tirkonnen-Condit (1995). Orozco (2002) también realiza una compilación de estudios empíricos sobre el proceso de traducción.

(método introspectivo) o una vez finalizada la traducción (método retrospectivo) verbalizan oralmente cuáles son sus procesos cognitivos. Los protocolos de verbalización se han obtenido de verbalizaciones individuales (Jääskeläinen 1996), o en grupo (por ejemplo en Pavlović 2009; Séguinot 1996). Además, dado que algunas actividades cognitivas del traductor son inconscientes, y las actividades que se exteriorizan son aquellas que permanecen en la memoria a corto plazo, es decir, aquellas que se efectúan de manera consciente (según han indicado Ericsson y Simon 1984), los TAP en ocasiones se complementan con otros métodos como cuestionarios y entrevistas, observación, monitorización (con el programa *CamtasiaStudio*), seguimiento de los ojos (*eye tracking*), registro de pulsaciones de teclas (*key-logging*, con el programa *Translog*), etc. De esta manera, con la triangulación de los datos pueden obtenerse resultados más objetivos (estas metodologías se describen por ejemplo en Alves 2003; Hansen 1999; Hansen 2002; Tirkkonen-Condit y Jääskeläinen 2000).

Además, respecto a los sujetos, se ha estudiado y comparado el proceso de traducción en estudiantes de lenguas extranjeras (Gerloff 1986; Krings 1986a; Krings 1986b; Lörcher 1986, 1991; Mondahl y Jensen 1996), estudiantes de traducción (Gile 2004; Jääskeläinen 1989; Livbjerg y Mees 1999, 2003; Pavlović 2009) y traductores profesionales (Alves y Liparini Campos 2009; Asadi y Séguinot 2005; Fraser 1999; Massey y Ehrensberger-Dow 2010a, 2010b; Séguinot 1996), con pocos años de experiencia, o traductores noveles, y con 8 a 10 años de experiencia, es decir, traductores expertos (Jääskeläinen 1989; Jensen 1999).

Asimismo, muchos de estos estudios observan y comparan diferencias en el proceso de traducción de sujetos variados, por ejemplo entre no traductores y traductores, y también entre estudiantes de traducción, traductores profesionales noveles y traductores expertos (Göpferich y Jääskeläinen 2009; Jensen y Jakobsen 2000; Kiraly 1995; Künzli 2001; Massey y Ehrensberger-Dow 2011; Ronowicz et al. 2005), por citar algunos ejemplos.

### **2.1.2 Resultados de las investigaciones empíricas sobre problemas y estrategias de traducción**

Los estudios sobre el proceso de traducción identifican una serie de problemas que el traductor trata de resolver por medio de la puesta en marcha de estrategias. Nord (1991: 151) distingue entre dificultad y problema de traducción. Esta autora indica que la dificultad de traducción es subjetiva, ya que cada traductor puede tener diferentes



dificultades, mientras que el problema de traducción es un problema objetivo que cualquier traductor debe resolver en el desarrollo de una tarea de traducción, independientemente de su nivel de competencia o de las condiciones técnicas de su trabajo. En este sentido, identifica cuatro tipos de problemas: textuales, pragmáticos, culturales y lingüísticos. Por su parte, Lörscher (1991: 202) señala problemas léxicos, sintácticos o léxico-sintácticos.

Roberts y Bossé-Andrieu (2006: 203-204) clasifican los problemas de traducción, en primer lugar, en función de si se trata de problemas relacionados con el texto origen o con el texto meta, en este sentido apuntan problemas de comprensión del texto origen o transferencia a la LM; en segundo lugar, en función de la naturaleza del problema, señalan problemas enciclopédicos, lingüísticos o textuales. Se entienden por enciclopédicos problemas relacionados con la falta de conocimientos del tema del texto. Los problemas lingüísticos se refieren a las dificultades con palabras o secuencias específicas, tanto en lo que respecta a la comprensión o la traducción. Los problemas textuales corresponden a la identificación o reproducción de un determinado tipo de texto. Según estas autoras, mientras que los problemas enciclopédicos tienen lugar principalmente en el texto origen y los problemas textuales en el texto meta, los problemas lingüísticos pueden aparecer tanto en el texto origen como en el texto meta.

Hurtado Albir (2001: 288), que se basa en estudios empíricos con estudiantes de traducción y traductores profesionales, clasifica los problemas de traducción en cuatro grupos: lingüísticos, extralingüísticos, instrumentales y pragmáticos.

- Problemas lingüísticos. Son problemas de carácter normativo, que recogen sobre todo discrepancias entre las dos lenguas en sus diferentes planos: léxico, morfosintáctico, estilístico y textual (cohesión, coherencia, progresión temática, tipologías textuales e intertextualidad).
- Extralingüísticos. Son problemas que remiten a cuestiones de tipo temático, cultural o enciclopédico.
- Instrumentales. Son problemas que derivan de la dificultad en la documentación (por requerir muchas búsquedas o búsquedas no usuales) o en el uso de herramientas informáticas.
- Pragmáticos. Son problemas relacionados con los actos de habla presentes en el texto original, la intencionalidad del autor, las presuposiciones y las implicaturas, así como los derivados del encargo de traducción, de las características del destinatario y del contexto en que se efectúa la traducción.

Las estrategias de traducción se definen como «potentially conscious plans for solving a translation problem» (Krings 1986a: 268) o «[...] the means which the translator, within the confines of his or her existing knowledge, considers to be best in order to reach the goals set by the task. Translators need to be able to determine the most effective way to act in each translating situation» (Jääskeläinen 1993: 111). Bergen (2006: 111) indica que

una forma de categorizar las estrategias en traducción es por medio de la fase del proceso de traducción en la que tienen lugar, así se refiere a *estrategias de comprensión*, *estrategias de transferencia* y *estrategias de producción*. Dentro del conjunto de estrategias de traducción que proponen los estudios, se ha constatado el uso de recursos por parte de los traductores.

Krings (1986a: 269) distingue las siguientes estrategias para resolver problemas de traducción: «strategies of comprehension, strategies of equivalent retrieval, strategies of equivalent monitoring, strategies of decision making and strategies of reduction». En la Ilustración 10 podemos ver cómo unas estrategias se relacionan con otras, donde según este autor: «A wide range of strategies could be observed in the way reference books were used. A particularly frequent strategy consisted in looking up the unknown item in a bilingual dictionary and subsequently checking the appropriateness of the given equivalents in a monolingual dictionary» (Krings 1986a: 270).

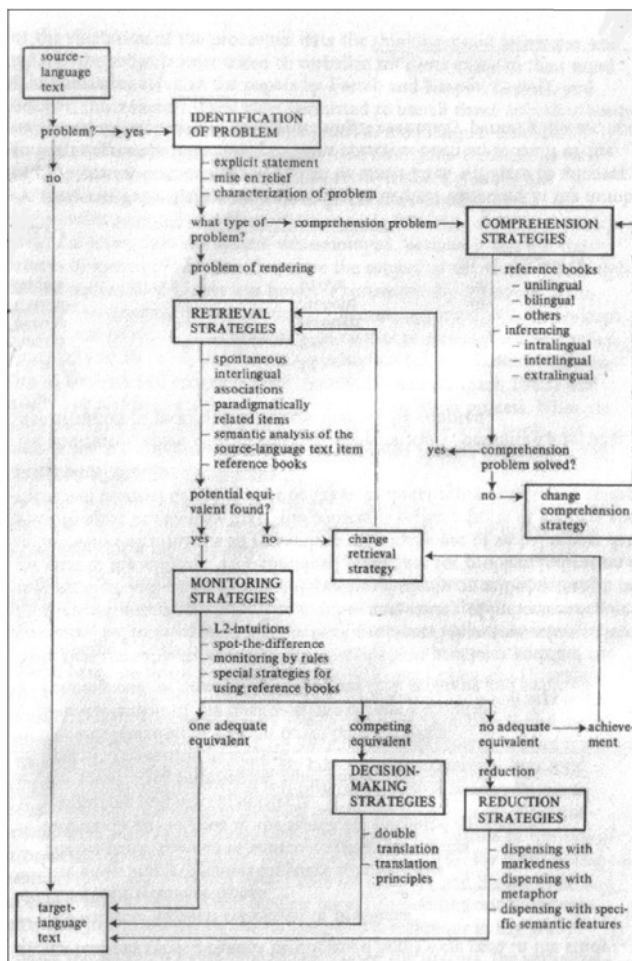
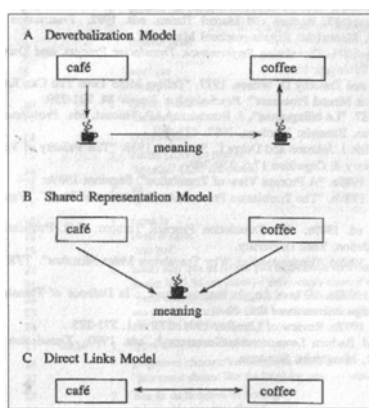


Ilustración 10. Modelo de proceso de traducción, según Krings (1986a: 269).

Gerloff (1986: 253), dentro de las estrategias de procesamiento textual, en la categoría tres, *storage and retrieval*, distingue dos subcategorías *memory search* «e.g. searches memory for English or French equivalent, identifies a word as remembered or never seen before, or waits for word “to emerge into consciousness”» y *dictionary search* «either says would use a dictionary at this point (when no dictionary is available) or uses one –searches dictionary for English equivalent, for English or French synonyms».

Séguinot (1996) distingue entre «interpersonal strategies, search strategies, inferencing strategies, and monitoring strategies». Dentro de las estrategias de búsqueda, menciona *externas* (acceso a recursos) e *internas* (los propios conocimientos del traductor).

Además, presenta tres modelos actuales de acceso al léxico en una traducción (modelo de desverbalización, representación compartida y acceso directo), como vemos en la Ilustración 11.



**Ilustración 11. Modelos actuales de acceso al léxico en una traducción, según Séguinot (1996: 86).**

Los estudios han detectado diferencias en el uso de los recursos en el proceso de traducción por parte de distintos sujetos. Por ejemplo, Krings (1986b) observa en los estudiantes de segundas lenguas un uso mayoritario del diccionario bilingüe, frente al monolingüe, que se consulta únicamente cuando el diccionario bilingüe resulta insatisfactorio. Jääskeläinen (1989) identifica que los estudiantes de traducción principiantes hacen un uso mayor de los recursos y sobre todo del diccionario bilingüe que los estudiantes más avanzados, que utilizan el diccionario monolingüe. Esta autora identifica estrategias de búsqueda en el diccionario bilingüe con comprobación posterior en el diccionario monolingüe (Jääskeläinen 1989: 190-196). En Ronowicz et al. (2005: 590) señalan que ninguno de los estudiantes de su estudio realiza búsquedas en el diccionario monolingüe ni suele consultar más de un diccionario, mientras que los profesionales y paraprofesionales sí que hacen uso del diccionario monolingüe y especializado.

En Jääskeläinen (1996) y Göpferich y Jääskeläinen (2009, que revisan los estudios de Gerloff (1988) y Jääskeläinen (1990)) se comprueba que los estudiantes de traducción y los traductores profesionales utilizan recursos más diversos que los no traductores. De manera similar, Künzli (2001) indica que los traductores profesionales utilizan recursos más variados que los estudiantes de traducción. Este autor señala que los traductores profesionales, aparte de diccionarios bilingües y monolingües (el primer y el segundo recurso más habitual respectivamente), utilizan otras fuentes (enciclopedias, textos paralelos, gramáticas, corrector ortográfico de Word; también una de las traductoras profesionales indica que consulta en diccionarios electrónicos y a expertos humanos, aunque no se dan más datos al respecto).

Igualmente, Massey y Ehrensberger-Dow (2011: 197-199) muestran que los estudiantes de traducción para resolver problemas lingüísticos hacen un mayor uso de diccionarios: del diccionario multilingüe en línea, luego impreso y en CD o DVD y a continuación el diccionario monolingüe impreso y en CD o DVD, mientras que los profesores de traducción optan primero por el diccionario monolingüe y después el multilingüe impreso o en CD o DVD y textos paralelos. En la resolución de problemas extralingüísticos los estudiantes consultan otros bancos de datos y también el diccionario monolingüe o multilingüe en línea, mientras que los profesores señalan los textos paralelos y los buscadores de Internet; los estudiantes apenas mencionan los buscadores especiales. Finalmente, los traductores profesionales hacen un mayor uso de buscadores, diccionarios multilingües en línea y bases de datos terminológicas.

Asimismo, según Künzli (2001), los profesionales muestran un coeficiente de consulta más elevado que los estudiantes, realizan más consultas para resolver un mismo problema, seguramente porque verifican el resultado obtenido en distintos recursos. Los resultados de su estudio indican, al igual que Nord (1999), que no puede afirmarse que el diccionario monolingüe garantice necesariamente la calidad de una traducción, por oposición al diccionario bilingüe, que algunos estudios pedagógicos en traducción consideran menos fiable, sino que el diccionario bilingüe es un recurso valioso, también para los traductores profesionales, junto con otros recursos.

Por su parte, Jensen (1999) constata diferencias en la cantidad de consultas en el diccionario, los no traductores son los que más los utilizan, después los profesionales con pocos años de experiencia (1-3) y finalmente los traductores expertos (más de 8 años de experiencia) son los que menos emplean el diccionario.

Asimismo, Massey y Ehrensberger-Dow (2010a: 136) muestran un nivel de acierto mayor en aquellos estudiantes que realizan la búsqueda en un buscador de Internet, que aquellos que recurren al diccionario bilingüe. Finalmente, Alves y Liparini Campos (2009: 204) señalan que el uso de memorias de traducción tiene un impacto descendente en la búsqueda de textos paralelos en Internet y diccionarios y que la búsqueda en estos dos últimos recursos también se utiliza para verificar las opciones que ofrece la memoria.

De todo lo anterior, hemos visto que el uso de recursos de referencia es una parte importante de las estrategias de los traductores, aunque la mayoría de estos estudios no profundizan propiamente en lo que son las estrategias de búsqueda, el uso real de los recursos por parte de los traductores (Künzli 2001: 507). Algunos anuncian la observación de «systematic research procedures» (Massey y Ehrensberger-Dow 2011: 206). Por eso, estudios como el de Ronowicz et al (2005: 592) reclaman lo siguiente: «[...] we probably need to re-examine anecdotal (written from experience), theoretical and empirical studies of strategies of use of dictionaries and other reference sources by professionals and design empirical studies to obtain supporting evidence». En este sentido, en el capítulo 4 recopilamos otros estudios que recogen estrategias de búsqueda en diccionarios, corpus e Internet por estudiantes de traducción y traductores profesionales.

## **2.2 La competencia traductora y la competencia documental y terminológica**

El traductor debe abarcar una serie de destrezas, habilidades, aptitudes, capacidades o competencias para poder desarrollar una buena labor de traducción. Los autores han utilizado distintas denominaciones, que se han recopilado por ejemplo en Hurtado Albir (2001: 383): competencia de transferencia, competencia traslatoria, competencia del traductor, actuación traductora, habilidad traductora, destreza traductora, etc. Orozco y Hurtado Albir (2002: 375) se decantan por la utilización de *competencia traductora*.

El grupo PACTE de la Universidad Autónoma de Barcelona ha centrado sus investigaciones en la competencia traductora y su adquisición. Los resultados obtenidos en sus numerosos estudios han permitido que se identificaran una serie de subcompetencias que forman parte de lo que denominan *competencia traductora*. Dichas subcompetencias se especifican en la siguiente cita.

**Competencia lingüística en las dos lenguas:** competencia gramatical, textual, ilocutiva, sociolingüística.

**Competencia extralingüística:** Conocimientos acerca de la organización del mundo en general y de ámbitos particulares: conocimientos biculturales, enciclopédicos, temáticos y sobre la traducción.

**Competencia instrumental/profesional:** Conocimientos y habilidades relacionados con el ejercicio de la traducción profesional: conocimiento del mercado laboral y comportamiento del traductor profesional, conocimiento y uso de las fuentes de documentación de todo tipo y de las nuevas tecnologías necesarias para el traductor.

**Competencia psicofisiológica:** Habilidad de aplicación de recursos psicomotores, cognitivos y actitudinales.

**Competencia de transferencia:** Capacidad de recorrer el proceso de transferencia desde el texto original hasta la elaboración del texto final.

**Competencia estratégica:** Procedimientos conscientes e inconscientes, verbales y no verbales utilizados para resolver los problemas encontrados en el desarrollo del proceso traductor.

(PACTE 2002: 131)

Hurtado Albir (2001: 397) señala que la competencia más importante es la *competencia de transferencia* y que en ella se integran el resto (véase también el modelo de PACTE 2001: 40, en la Ilustración 12). Asimismo, esta autora indica que la importancia de cada una de las subcompetencias dependerá de distintos factores, como la dirección de la traducción (directa o inversa), la combinación de lenguas, la especialidad (técnica, jurídica, literaria, etc.), la experiencia del traductor o el contexto de la traducción (tipo de encargo, plazos). Por ejemplo, en la traducción inversa influye en mayor medida la competencia instrumental, cada especialidad requiere distintas habilidades psicológicas, etc.

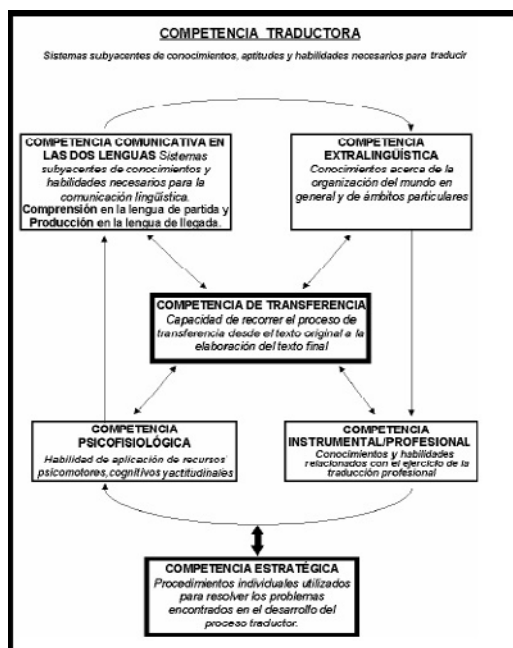
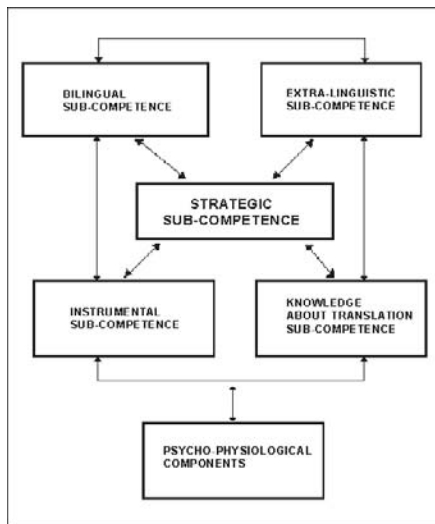


Ilustración 12. Modelo de competencia traductora, según PACTE (2001: 41).

En PACTE (2003) se remodela la competencia traductora, como vemos en la Ilustración 13, en la que se indica que es un conocimiento experto que no poseen todas las personas bilingües y es un conocimiento principalmente procedimental. El modelo se compone de 5 subcompetencias, donde la subcompetencia estratégica se encuentra en el centro ya que está presente en todo conocimiento procedimental.



**Ilustración 13. Modelo de competencia traductora, según PACTE (2003: 60).**

La investigación empírica en la competencia traductora sugiere que la subcompetencia estratégica, la subcompetencia instrumental y la subcompetencia de conocimientos sobre traducción son específicas de los traductores, a diferencia del resto de subcompetencias que comparten con los sujetos bilingües (PACTE 2005a: 611).

En el estudio sobre la competencia traductora y su adquisición<sup>14</sup>, el grupo PACTE identifica los siguientes problemas de traducción: *linguistic problems*, *textual problems*, *extralinguistic problems*, *problems of intentionality*, *problems relating to the translation brief and/or the target-text reader* (PACTE 2011). Para la resolución de estos problemas intervienen secuencias de acciones que agrupan en cinco categorías en función del grado de implicación cognitiva del usuario (*apoyo interno*), y del uso de recursos de documentación (*apoyo externo*). Dentro de este último, se refieren a la *consulta simple*, en un diccionario bilingüe o monolingüe, o *consulta compleja*, realización de búsquedas en buscadores y también búsquedas contextualizadas.

<sup>14</sup> Los sujetos de este estudio son 35 traductores expertos y 24 profesores de lenguas extranjeras.

Las acciones llevadas a cabo por los traductores pueden ser, por tanto, (1) *apoyo interno*, cuando el traductor propone una solución de traducción por sí mismo sin realizar ninguna consulta y la mantiene en el texto meta definitivo; (2) *predominio de apoyo interno con recurso a apoyo externo*, lo que significa que el traductor realiza consultas elaboradas, pero la solución definitiva no proviene de las consultas sino del apoyo interno; (3) *interacción equilibrada entre apoyo interno y apoyo externo*, que puede ser de tres tipos: uso de una consulta de cualquier tipo para confirmar una decisión, uso de una consulta simple que no se acepta, o uso de apoyo interno y consulta simple y se acepta; (4) *predominio de apoyo externo con recurso a apoyo interno* cuando las consultas elaboradas son la base de la solución definitiva; y (5) *apoyo externo simple*, la búsqueda se efectúa en un diccionario bilingüe y se adopta la solución propuesta (PACTE 2005a, 2005b). En PACTE (2009: 223) se reduce a cuatro categorías: *apoyo interno*, *predominio de apoyo interno*, *predominio de apoyo externo* y *apoyo externo*.

Una de las conclusiones que se derivan de este estudio es que cada grupo experimental utiliza secuencias de acciones diferentes para resolver un mismo problema (PACTE 2009: 225). Además, los autores señalan que las soluciones de traducción más aceptables provienen de los traductores expertos que combinan el uso de apoyo interno y apoyo externo, es decir, la consulta de distintos tipos de recursos (de ahí la importancia que se confiere a la subcompetencia instrumental). Además, señalan que el predominio de apoyo externo es más frecuente en la traducción inversa que en la directa, ya que las deficiencias de conocimiento en una lengua se compensan con la competencia instrumental.

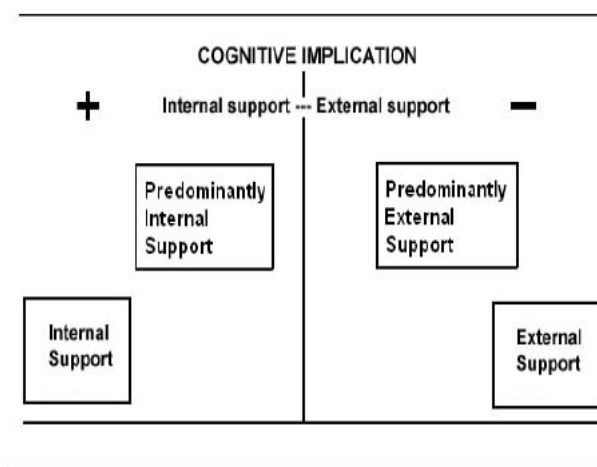
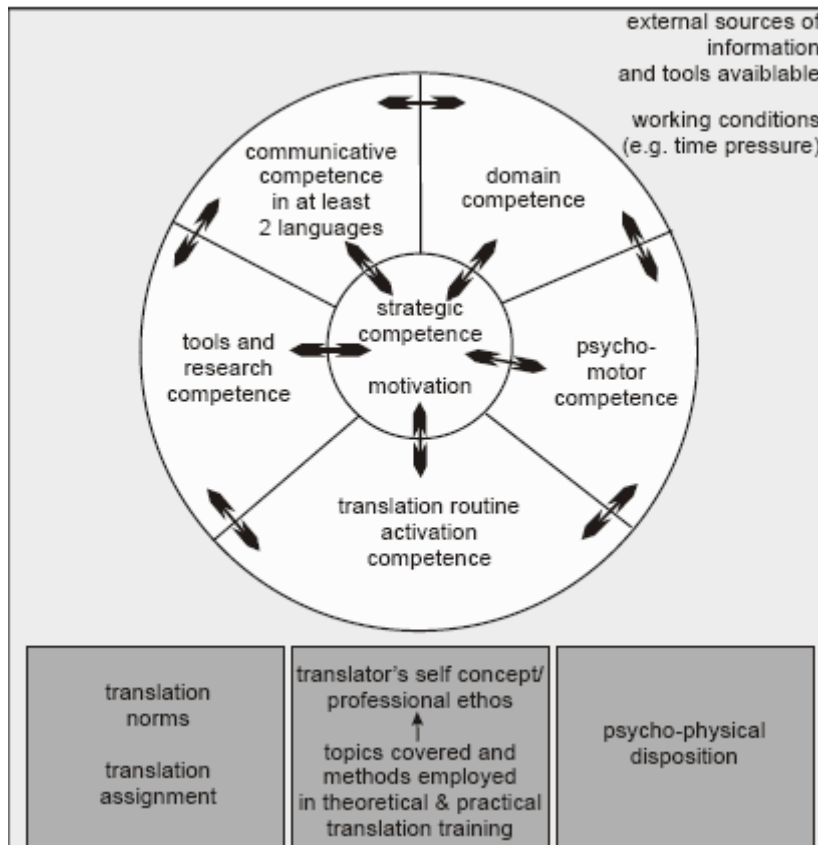


Ilustración 14. Implicaciones cognitivas en secuencias de acciones, según PACTE (2009: 223).



Seguendo el modelo de PACTE, Göpferich (2009: 20) se centra en los siguientes componentes de la competencia traductora: *strategic competence*, *translation routine activation competence* y *tools and research competences*. Según se indica, estas competencias son las competencias específicas que poseen los traductores, que los diferencia de las personas bilingües que no han recibido formación en traducción (Göpferich y Jääskeläinen 2009: 185).



**Ilustración 15. Modelo de competencia traductora en el estudio TransComp (Göpferich 2009: 20).**

Algunos autores han llamado a lo que se concibe por competencia instrumental, según PACTE, «*competencia documental* entendiendo por ella las habilidades y destrezas específicas del traductor para saber manejar con precisión y rapidez las fuentes de información de cualquier tipo y en cualquier formato necesarias para su quehacer traductor» (Palomares Perraut y Pinto 2000: 100). Otros autores destacan también el papel de la documentación en traducción (Gonzalo García 2004: 74; Palomares Perraut y Pinto 2000; Pinto 2000; Recoder y Cid 2004).

Como se puede observar en la Tabla 2, Gonzalo García (2004: 278) insiste también en la necesidad de que el traductor de textos especializados sea capaz de documentarse para obtener conocimientos sobre la temática, las reglas de confección del texto especializado, el género textual y la terminología y fraseología del campo de especialidad.

Competencias requeridas	Mecanismos para identificar y resolver problemas documentales
Conocimiento de la temática especializada	Consulta de fuentes de información especializada Consulta a especialistas Consulta a servicios lingüísticos Consulta a organismos de normalización oficiales
Conocimiento de las reglas de confección del texto especializado	
Conocimiento de la diversidad de tipos o géneros textuales	
Conocimiento de la terminología y la fraseología propias de la especialidad	

**Tabla 2. Capacidad para documentarse del traductor de textos especializados (Gonzalo García 2004: 278).**

Gamero (2001: 48) clasifica las competencias para la traducción técnica, en las que también incluye los conocimientos del campo temático, de los géneros técnicos y de la terminología técnica específica.

Características de funcionamiento textual	Competencias requeridas	
Importancia del campo temático técnico	→ Conocer los ámbitos técnicos	Ser capaz de documentarse en relación con los textos técnicos
Utilización de terminología técnica específica	→ Saber aplicar la terminología técnica adecuada en lengua de llegada	
Presencia de géneros técnicos característicos	→ Dominar los rasgos convencionales de los géneros técnicos en lengua de llegada	

**Tabla 3. Características textuales de los textos técnicos y competencias requeridas para la traducción técnica (Gamero 2001: 48).**

Sobre el campo temático, se precisa una competencia de comprensión de la materia técnica que se traduce, lo que supone adquirir conocimientos sobre los conceptos y las relaciones lógicas de los conceptos entre sí (Gamero 2001: 42). Esta autora añade que la comprensión que adquiere el traductor es pasiva, ya que no es necesario que se convierta en un experto en la materia, por ejemplo sobre cómo diseñar una determinada máquina, pero sí que tenga unas nociones básicas sobre cómo funciona dicha máquina.

Sobre la utilización correcta de la terminología técnica, el traductor debe ser capaz de encontrar en la LM los equivalentes de la LO. Para ello, deberá realizar una búsqueda puntual de terminología en obras de referencia de tipo terminológico. En esta labor, el traductor también recurre a la documentación por medio de textos especializados y a la consulta de especialistas, mediante preguntas concretas (Gamero 2001: 43).

Para esta autora, las fuentes más útiles para el traductor son las fuentes terminológicas (glosarios, diccionarios especializados, etc.), la documentación especializada (manuales, monografías, artículos de investigación, etc.) y las fuentes bibliográficas (repertorios bibliográficos, catálogos de bibliotecas, etc.) (Gamero 2001: 45).

Gamero (2001: 46) afirma que la recuperación de la información suele realizarse por medios informáticos porque de este modo es más rápida. En este sentido, destaca la creciente utilidad de Internet en el acceso a fuentes documentales, bibliográficas y terminológicas. Aquí aparece la red como un medio de búsqueda terminológica. Según esta autora, los servicios telemáticos más utilizados por los traductores son el correo electrónico, la propia Web, las listas de correo, FTP, los grupos de noticias y las discusiones interactivas.

Así, vemos que dentro de la competencia documental o instrumental hay una necesidad de acceso a la terminología. Por ejemplo, Sager (1993: 261) indica que la recuperación terminológica puede considerarse una aplicación especial de la recuperación de información. Este autor señala que la recuperación terminológica depende, en gran medida, de las tendencias y los avances actuales de los sistemas de recuperación de documentos. Por su parte, Mayoral (1994: 117) habla de un profesional «traductor-terminólogo o traductor-documentalista».

Gietz (1990) alude a la búsqueda terminológica entre las destrezas que implican dominar la competencia documental: «Según su mayor o menor nivel de competencia, el traductor deberá realizar un esfuerzo inversamente proporcional. En menor o mayor grado, deberá efectuar una investigación o búsqueda documental y terminológica».

Por ejemplo, Cabré (2004: 101) advierte que el traductor de textos especializados debe gozar de «una competencia paralela a la del especialista que de manera natural se comunica sobre la especialidad». Para conseguirlo, según indica, deberá simular que es un especialista, que domina la materia y su especificidad cognitiva y que maneja los mismos elementos léxicos que los expertos de dicha especialidad, debe conocer la terminología.

A esta competencia similar a la del especialista se le ha denominado, *competencia terminológica* (véase por ejemplo Faber 2002a: 4, 2003; Montero Martínez y Faber 2008: 100, 2009; Vázquez Béjar 2008: 713). Faber (2002a: 4) vincula cuatro habilidades principales: el almacenamiento de conocimientos especializados, los mecanismos para la adquisición de terminología, la creatividad en la formación de términos y la capacidad del traductor para resolver problemas en la adquisición de conocimientos durante el proceso traductor.

Montero Martínez y Faber (2008: 100) señalan que para poder traducir eficazmente, los mediadores lingüísticos deben aplicar conocimientos teóricos y habilidades que forman parte de una *macrocompetencia traductora* dinámica, que puede variar dependiendo de las necesidades del mercado. Estas autoras explican que uno de los factores que contribuyen a conseguir una buena traducción es la *subcompetencia terminológica*, que abarca aspectos como las búsquedas terminográficas, las estrategias documentales, así como la reconstrucción parcial de los ámbitos de conocimiento especializados. Montero Martínez y Faber (2008: 88) señalan que las características específicas del proceso de traducción son las que determinan el tipo de *competencia terminológica* que se precisa.

Estas autoras especifican una serie de habilidades o estrategias que forman parte de esta *subcompetencia terminológica*:

- la identificación y adquisición de los conceptos especializados del discurso;
- la evaluación, consulta y elaboración de recursos de información;
- el reconocimiento de correspondencias interlingüísticas basadas en los conceptos del ámbito de conocimiento especializado;
- la gestión de la información y el conocimiento adquirido y su reutilización en futuras traducciones.

(Montero Martínez y Faber 2008: 92)

Vázquez Béjar (2008: 713) ofrece una definición de *competencia terminológica*: «Conjunto de conocimientos, hábitos y habilidades para el trabajo con los términos, que abarca las esferas lingüística, cognitiva y pragmático-funcional de los mismos». Esta autora considera que dicha *competencia terminológica* desempeña un papel fundamental en el ejercicio profesional del traductor. Por tanto, explica, en la formación del futuro traductor profesional destacan los aspectos relacionados con la terminología y el acceso a recursos y fuentes terminológicas.

Los futuros profesionales de la traducción tienen como objetivo de formación general los siguientes aspectos:

- Traducir correctamente un texto.
- Redactar correctamente dicho texto.

- Resolver los problemas terminológicos que pueda presentar dicho texto.
- Acceder correctamente a los recursos y fuentes terminológicas.
- Elaborar un glosario temático.

Para ello, el futuro traductor deberá poseer buenos conocimientos en:

- La lengua de partida (LP).
- La lengua de llegada (LLI).
- Los recursos terminológicos disponibles.
- Las fuentes documentales para la terminología.
- La formación de términos (creación neológica).
- La materia especializada del texto a traducir.

(Vázquez Béjar 2008: 714)

### **2.3 La relación entre traducción y terminología**

La relación existente entre la traducción y la terminología es un tema recurrente en la investigación de ambas disciplinas. Montero Martínez y Faber (2008: 91) señalan que esta relación es asimétrica porque la terminología no tiene una necesidad intrínseca de recurrir a la traducción y, sin embargo, la traducción necesita de la terminología como un medio para efectuar con éxito la transferencia interlingüística de las unidades de conocimiento especializadas.

Ningún hablante es capaz de dominar todo el vocabulario existente en una lengua, y mucho menos de todos los ámbitos y disciplinas que existen. Sin embargo, el traductor tiene que enfrentarse diariamente a textos de temáticas muy distintas, que en la mayoría de los casos desconoce y que debe traducir en el menor tiempo posible y con la precisión de un especialista en la materia. Por tanto, para desarrollar efectivamente la actividad de traducción, se precisa realizar labores de búsqueda y gestión terminológica.

La traducción precisa del léxico y de la terminología en todo tipo de traducciones, pero especialmente en el caso de la traducción especializada, donde el léxico especializado aumenta.

Montero Martínez y Faber (2008: 91) señalan que la terminología dentro de la traducción se centra en los términos utilizados en contextos reales y en procesos de comunicación dinámicos, y cuando se analiza el uso de las unidades terminológicas tal y como aparecen en los textos, en el discurso oral o en las situaciones comunicativas, se hace evidente que el lenguaje general y el especializado no son tan diferentes. Por ejemplo, en el discurso especializado se encuentran habitualmente casos de polisemia y sinonimia, que provocan problemas de traducción, mientras que cuestiones como la estandarización o los

neologismos (que normalmente son esenciales para la Terminología) pasan a un segundo plano.

Todo mensaje escrito se compone de una combinación de palabras y términos, por lo que la terminología está presente en todo proceso de traducción, pero sobre todo en la traducción especializada (Sager 1992). Arntz (1993: 5) explica que el traductor tiene que estar familiarizado con la terminología de los textos antes de empezar a traducir. Para Bowker y Pearson (2002: 193) la traducción de textos especializados supone una serie de retos, por ejemplo el traductor puede tener dificultades para encontrar un equivalente, e incluso habiendo encontrado un equivalente adecuado, puede haber problemas para emplear un estilo adecuado en dicho campo de especialidad.

Como señala Cabré (2000: 2), para que un texto traducido sea igual que el original es necesario que se aproxime al máximo a los usos léxicos de un productor nativo en la LM. Para los traductores «la terminología es el conjunto de unidades que usan los especialistas en un ámbito y que aparecen en un texto de traducción. Su objetivo consiste en disponer en la lengua de traducción de unidades equivalentes que posean el mismo valor de uso» (Cabré 2004: 115).

Cabré (1999: 28) indica que la traducción, concebida como actividad práctica, se enfrenta a problemas de terminología que debe solventar y especifica que dicha terminología es imprescindible en la traducción especializada por dos motivos: porque en los textos especializados los términos son los que concentran el conocimiento y porque la calidad de una traducción especializada depende del uso que se hace de la terminología. La terminología que se emplee debe corresponderse con las unidades terminológicas que utilizaría un especialista en su campo de trabajo. Autores como Le Poder (2004: 27) y Montero Martínez y Faber (2008: 90), entre otros, afirman que los traductores de textos especializados deben ser capaces de identificar y solventar las dificultades terminológicas para no obstaculizar el proceso de traducción. Por su parte, Velásquez (2002: 144) habla de «una relación *sine qua non* entre traducción y terminología en la comunicación bilingüe mediada de lenguajes especializados».

La constante evolución en los ámbitos científicos y técnicos ha generado una demanda cada vez mayor de traductores especializados en las temáticas más diversas. Faber (2005: 13) indica que en las últimas décadas la traducción de textos científico-técnicos se ha intensificado enormemente. Según esta autora, en Europa se estima que el 50% de los 376 millones de páginas traducidas en el año 2000 eran textos de esta especialidad, en comparación a un 1% de textos literarios.

Quiroz et al. (2002) señalan que gran parte del tiempo y el esfuerzo empleado en la traducción de un texto especializado se dedica a la resolución de dificultades terminológicas. En concreto, Walker (1993: 222) ha estimado que este tiempo es de cerca del 40%. Varantola (2002a: 173) indica que el 50% del tiempo de una traducción está destinado a tratar de encontrar información léxica relevante en distintos recursos de referencia (Varantola 1998: 179). Tomaszczyk (1989) en un estudio de casos en el que se traduce un libro sobre el uso industrial de los diamantes del polaco al inglés, el 80% de las consultas las abarca la terminología especializada, el resto colocaciones, propiedades sintácticas de las palabras, preposiciones, etc.

En cuanto al nivel de relación que debe existir entre la traducción y la terminología, la mayoría de autores aclaran que no resulta necesario que el traductor se convierta en un experto en la material para poder llevar a cabo la traducción de un texto especializado (Neubert 2000: 9). Cabré (2000: 2) explica que un traductor puede mostrar diferentes grados de compromiso con la terminología.

Cabré (2004: 117) denomina a la labor terminológica del traductor *terminología puntual* o *terminología ad hoc* para diferenciarla de la tarea que llevan a cabo los terminólogos. Cabré et al. (2002: 165) tratan de generalizar las necesidades terminológicas del traductor científico, para lo cual toman como punto de partida un estudio efectuado por Estopà (1999) en el que se pidió a profesionales distintos (especialistas, documentalistas, traductores y lexicógrafos) que realizaran, con el mismo texto especializado, un vaciado terminológico útil para su trabajo (transmitir conocimiento especializado, indexar el texto, preparar su traducción o elaborar un diccionario). Los vaciados de los distintos grupos no coincidieron, ni en el número de términos seleccionados, ni en la naturaleza de dichos términos, ni en la información sobre los términos que le resultaba útil a cada colectivo.

El experimento demostró que los distintos grupos de profesionales tienen necesidades terminológicas diferentes cuando trabajan con un texto especializado, ya que las tareas que deben desarrollar son distintas. Por lo que respecta al vaciado de los traductores, se observó que seleccionaban aquellas unidades terminológicas que podían suponer un problema en la traducción (bien porque desconocían su significado o podían plantear dificultades a la hora de encontrar un equivalente de traducción). Seleccionaron sustantivos, verbos, adjetivos, adverbios, siglas y fraseología y el tipo de información que les interesaba especialmente sobre los términos eran elementos de su contexto inmediato (colocaciones del término), palabras relacionadas y la frecuencia de los términos.

Resultaron de interés para los traductores vocablos que aparentemente no eran de carácter especializado, pero que formaban parte de una unidad terminológica, ya que este tipo de palabras no especializadas pueden adquirir un matiz específico en el discurso especializado y una traducción literal puede resultar en un falso amigo. Así, aunque cada traductor tiene necesidades terminológicas diferentes dependiendo de su conocimiento del tema, su experiencia profesional, etc., se extraen las siguientes conclusiones (Cabré et al. 2002: 166-167):

Las unidades terminológicas que no provocan problemas de traducción no interesan al traductor.

Las unidades terminológicas que más interesan a los traductores son las que pueden ocasionar problemas en la traducción, bien porque desconocen su significado o el sentido dentro del texto, bien porque intuyen que la búsqueda de un equivalente puede plantear dificultades.

Todos los elementos textuales que facilitan la búsqueda de un equivalente para una unidad terminológica se tienen en cuenta.

Cabré (1999: 30) clasifica las necesidades propiamente terminológicas de los traductores en dos tipos: las de la terminología de la lengua B y las de la terminología de la lengua A. En cuanto a la terminología de la LO (o lengua B) cabría mencionar el conocimiento del significado de un término origen y sus distintas denominaciones dependiendo del contexto en el que aparece. De la terminología de la LM (o lengua A) resulta importante averiguar si existe un equivalente en LM para un término original, conocer la denominación más apropiada en el contexto de la traducción y asegurarse de que la denominación utilizada en la lengua término se corresponde con el concepto que se quería expresar en el texto de partida.

Sobre la terminología de la lengua B, el traductor puede necesitar:

- Conocer o precisar el significado de una unidad terminológica del texto de partida.
- Confirmar su carácter especializado.
- Conocer sus alternativas denominativas y sus condiciones de utilización en los textos, etc.

Sobre la terminología de la lengua A o la lengua de traducción, necesita posiblemente:

- Saber si existe una unidad terminológica equivalente en la lengua de traducción.
- En el caso de que no exista, saber qué recursos puede o debe utilizar para adaptar, formar o crear una denominación razonable y cómo advertir del carácter neológico de la unidad.
- Conocer la unidad denominativa equivalente más adecuada teniendo en cuenta el tema, el enfoque o la perspectiva del trabajo.
- Conocer si un término debe ser usado gramaticalmente de manera especial o restrictiva.
- Conocer la combinatoria prototípica de los términos.



- Conocer la fraseología común del ámbito especializado.
- Conocer las alternativas denominativas de un mismo concepto y sus condiciones pragmáticas.
- Asegurar que la unidad denominativa que selecciona corresponde exactamente al concepto al que se atribuye, etc.

(Cabré 1999: 30)

En otra clasificación de dificultades terminológicas a las que pueden enfrentarse los traductores, Cabré (2004: 105) destaca la carencia de los diccionarios en cuanto a la búsqueda de equivalencias; la falta de contextos en los diccionarios bilingües, lo que imposibilita seleccionar la equivalencia más adecuada para cada texto, especialmente cuando el diccionario ofrece varias alternativas; y la dificultad de encontrar una equivalencia exacta entre lenguas, por razones culturales o de otro tipo.

- No saben si la lengua A dispone de una unidad terminológica lexicalizada para expresar una idea que en la lengua B se expresaba mediante un término porque no figura en los diccionarios de la materia;
- no saben si las unidades que les recomiendan los diccionarios bilingües especializados son las más adecuadas para el texto de traducción;
- no saben qué unidad deben seleccionar en caso de tener diferentes alternativas que les ofrecen los glosarios;
- no saben si la equivalencia semántica entre terminología es posible en muchos casos fuertemente marcados por el sesgo cultural o institucional, etc.

(Cabré 2004: 105)

Por su parte, Sánchez-Gijón (2003a: 355) propone tres dificultades en torno a la búsqueda de equivalencias terminológicas en la traducción especializada, dependiendo de si el traductor desconoce información lingüística de un término o información conceptual.

	<b>Informació lingüística</b>	<b>Informació factual</b>
Possibilitat 1	Coneguda	Desconeguda
Possibilitat 2	Desconeguda	Desconeguda
Possibilitat 3	Desconeguda	Coneguda

**Tabla 4. Tipología de problemas suscitados por unidades de significación especializada en la traducción especializada (Sánchez-Gijón 2003a: 355).**

Tal y como indica Sánchez-Gijón (2003a: 355) una unidad determinada de la que el traductor conoce tanto la información lingüística como el significado, o la información factual (conceptual) que representa dentro de un área temática determinada, no supondrá ningún problema de traducción. Según esta autora, los problemas a los que el traductor se enfrenta en un texto especializado se corresponden con los que se reflejan en la Tabla 4.

En cuanto a la primera posibilidad, cuando el traductor conoce una unidad lingüística, pero desconoce su información factual, puede tener dificultades en identificar

que existe un problema terminológico. Esta situación puede llevar al traductor a cometer un error de traducción al utilizar un término que cree que conoce pero que en el contexto de la traducción no es correcto. Un ejemplo sería el uso de un falso amigo.

La segunda posibilidad, y la más frecuente en la traducción especializada, es que se desconozca tanto la información lingüística como la factual de un término porque el traductor no domine la temática del texto. Normalmente, en la búsqueda para resolver este problema se encuentra la información factual necesaria para obtener la información lingüística.

La tercera posibilidad que apunta esta autora es similar a la anterior. La única diferencia es que, en este caso, el traductor ya conoce la información factual que rodea al término y a partir de ella busca la información lingüística.

Wright y Wright (1997: 148) exponen una serie de desventajas concretas a las que se enfrenta el traductor de textos especializados cuando desarrolla labores terminológicas, en comparación con un experto en la materia. Entre ellas, destacan el desconocimiento de la temática del texto y la terminología asociada a la misma, las carencias que presentan los recursos tanto en LO como en LM y la falta de tiempo para desarrollar la tarea de traducción, lo que dificulta la búsqueda terminológica y la creación de recursos terminológicos propios que se puedan aprovechar en futuras traducciones.

[...] Translators, in contrast to the experts for whom the original text was written, are frequently dropped down on these isolated information atolls, and left to fend for themselves, unfamiliar with the sea of knowledge that surrounds them.

Translator-terminologists experience several disadvantages:

- As noted, they are not subject-field experts.
- They may even have difficulty determining the field the text actually belongs to (seemingly logical, but totally incorrect assumptions can lead to disastrous translations).
- Available research materials in both the target and the source language are inadequate.
- They lack access to subject-field specialists.
- Due to short delivery deadlines, they lack time to pursue extensive (or even cursory) research activity.
- Even when information is available, they lack time to create extensive, thoroughly documented terminological entries.

(Wright y Wright 1997: 148)

Kiraly (2000: 136) señala que los traductores deberán gestionar la terminología, tanto si trabajan como autónomos, o como trabajadores en plantilla para alguna empresa o para la administración. De hecho, autores como Bowker y Marshman (2009: 60) indican que, aunque la tarea terminológica pueden llevarla a cabo otros profesionales, es un componente indiscutible de la carga de trabajo del traductor. Por ello, distintos autores se

refieren a los traductores como traductores-terminólogos (Arntz 1993: 17; Fulford 2001: 260; Mayoral 1994: 117). Esto también se refleja en el interés creciente de la formación de los traductores en terminología (Maia 2002b), como se refleja en distintos estudios y publicaciones, como por ejemplo en el reciente volumen especial de la revista *Terminology* (Alcina 2009b; Bowker y Marshman 2009; Montero Martínez y Faber 2009; Sánchez-Gijón et al. 2009, entre otros).

## **2.4 Herramientas terminológicas en traducción**

Ahora veremos algunas clasificaciones de recursos en traducción. Cabré (1999: 19) identifica cuatro tipos de fuentes de información para el traductor: fuentes gramaticales y discursivas, principalmente gramáticas y manuales de estilo; fuentes lexicográficas, diccionarios generales monolingües y bilingües; fuentes terminológicas, glosarios, vocabularios y diccionarios especializados; y fuentes especializadas, manuales, normas, monografías y artículos especializados.

Por su interés terminológico para el traductor, esta autora (Cabré 1999: 32) muestra las siguientes fuentes de consulta dependiendo del grado de especialización del texto para traducir: diccionarios de lengua general, enciclopedias, diccionarios generales de la ciencia y de la técnica, diccionarios especializados o glosarios terminológicos, diccionarios visuales especializados, léxicos o vocabularios, bases de datos terminológicos.

Montero Martínez y Faber (2008: 147-187) realizan una clasificación de recursos que pueden ser utilizados por el traductor en la documentación terminográfica. En primer lugar, aluden a obras terminológicas o lexicográficas especializadas, entre las que incluyen las enciclopedias especializadas, que pueden ser sistemáticas (también temáticas o metódicas porque su sistema de ordenación no es alfabético) o alfabéticas; los diccionarios monolingües especializados que pueden ser generales o específicos de una materia, estas obras según las autoras suelen contener información sobre el significado y no sobre el uso; los diccionarios bilingües especializados, también generales o específicos, que suelen estar ordenados alfabéticamente y contienen equivalentes pero no definiciones, ni información sobre colocaciones y compuestos, a excepción de algunos diccionarios; los diccionarios multilingües especializados que contienen todavía menos información que los bilingües; los diccionarios visuales que contienen imágenes en las que se marcan los nombres de

objetos, máquinas o procesos; los vocabularios y glosarios de ordenación alfabética<sup>15</sup>; los tesauros, las clasificaciones, las nomenclaturas, los sistemas de codificación, que son vocabularios controlados con fines específicos; las bases de datos terminológicas normalmente de organismos internacionales y empresas multinacionales, que suelen organizar los datos lingüísticamente y no conceptualmente, aunque está surgiendo una nueva generación de bases de datos basadas en el conocimiento, como veremos.

Aunque para la búsqueda de información terminológica son en principio más útiles las obras especializadas, según Montero Martínez y Faber (2008), igualmente podrían interesar las obras de lexicografía general, por ejemplo diccionarios monolingües y enciclopedias de carácter general porque también pueden contener información especializada. Además, incluyen obras de contenido especializado, donde puede accederse a textos paralelos. Finalmente, los traductores pueden necesitar documentarse en Internet o elaborar y consultar un corpus ad hoc.

Alcina (2008: 92) efectúa una clasificación de recursos para el traductor en seis fases de la traducción (véase la Tabla 5). En la fase tres, de documentación y búsqueda de terminología en la LM, intervienen buscadores de Internet, bases de datos en línea, diccionarios electrónicos, corpus electrónicos en línea o CD-ROM, software de análisis textual, foros, listas de distribución, grupos de noticias y portales, entre otros.

---

<sup>15</sup> Tal y como señalan distintos autores (Bowker y Pearson 2002: 137; Quiroz et al. 2002; Wright y Budin 1997: 325) un glosario es un listado de palabras de una misma área temática, ordenadas alfabéticamente, que pueden presentar equivalencias en otros idiomas, pero que no suelen incluir ninguna otra información adicional.

Comparing the phases of the translation process with and without the help of computer tools

	Cycle of translation before using the computer	Cycle of translation with the aid of a computer	Programs and resources
Phase 1	- Look for clients and translation work through adverts, colleagues, former teachers, acquaintances, and so forth	- Getting a website up and running on the Internet with enough information to orient clients about working languages, specialised area(s), approximate price per word, formats used for working, ways of submitting source texts, infrastructure available, and so forth - Contacting other professionals by means of e-mail, mailing lists, forums, chats.	HTML editor E-mail Mailing lists
Phase 2	- Receive translation work (by fax, courier, post, express mail)	- Use of e-mail, sending the file as an attachment, use of virtual collaborative work environments or downloading files by FTP	E-mail Virtual collaborative work environments FTP
Phase 3	Understanding the text: - Documentary research on the speciality by means of specialised books, obtaining documents from libraries and research centres, consulting experts and specialists, terminology lookup in specialised monolingual and multilingual dictionaries, and so forth - Search for target language terminology in bilingual or multilingual dictionaries, consulting specialised texts in target language	- Documentary research by means of web pages on the Internet - Consulting bibliographical resources in libraries and institutions through their websites on the Internet - Use of general and specialised, monolingual, bilingual and multilingual electronic dictionaries on the Internet or on CD-ROM - Consulting electronic text corpora on the Internet or on CD-ROM - Looking up information in a personal text corpus by means of text analysis or concordance software - Consulting specialists in a particular subject by means of expert forums - Consulting other translators and terminologists by means of mailing lists, news groups and virtual communities	Internet browser Online databases Electronic dictionaries Text corpora on line or on CD ROM Text analysis software Forums, mailing lists, news Portals
Phase 4 Translation	Writing the target text - Correction of spelling, grammar and style, and revision of the target text by the same translator - Correction and revision of the target text by a reviewer or a second translator	- Use of word processors - Use of a spelling, grammar and style checker and/or corrector, or the tools doing a similar job included in the word processor - Use of the revision tools in the word processor (protect document and track changes) - Use of interactive assisted translation (translation memories) - Use of machine translation	Word processor Text editor Spelling, grammar and style checkers Interactive assisted translation Translation memories Machine translation
Phase 5 Presentation	Sending translation back to client	- Preparation of the text in a suitable format for the client (the document may be in any of the different formats used by word processors like doc or rtf, those used by DTP programs such as mif, or in hypertext as html, and so forth) - Use of e-mail to return the translated text and, should it be the case, other materials (such as glossaries, translation memory, and so forth)	File compression and decompression software Word processor Desktop publishing Virtual collaborative work environments
Post-translation	- Writing up a project report, creation of terminology record cards with the data that was used during the translation (terminological equivalences)	- Alignment of original and translated texts so they can be added to a translation memory - Updating terminology databases - Automatic word-count - Incorporating the texts in a personal corpus of texts	Text alignment software Translation memory Terminology database management software

**Tabla 5. Comparación de las fases del proceso traductor con y sin ayuda de herramientas informáticas (Alcina 2008: 92).**

## 2.5 *Recapitulación*

En este apartado hemos visto que el traductor precisa de la terminología para realizar con éxito una traducción y, por ello, que deberá desarrollar una competencia terminológica como parte de la competencia traductora que se le exige. También hemos comprobado que las necesidades terminológicas pueden variar dependiendo del tipo de traducción, de la fase del proceso de traducción, de los conocimientos o experiencia del traductor en la temática del texto que traduce, etc.

Estas necesidades surgen principalmente en dos etapas del proceso traducción, en la de comprensión del texto origen y en la producción del texto meta. En los problemas de comprensión el traductor desconoce el significado de una denominación del texto origen, mientras que en la fase de producción, el traductor necesita encontrar un equivalente, es decir, una denominación en LM que exprese la misma idea o significado que la LO. Así, algunos autores se han referido a que las necesidades de comprensión son *semasiológicas*, mientras que las de expresión son *onomasiológicas*.

En este sentido, si los traductores tienen necesidades de distinto tipo, es de suponer que los recursos para satisfacer estas necesidades o problemas también deberán ser diferentes. Por tanto, deberemos comprobar, por un lado, qué funciones (o técnicas de búsqueda) ofrecen actualmente los recursos que puede utilizar un traductor (tema que abordaremos en la parte II de este trabajo) y, por otro lado, de qué forma utilizan los traductores estas funciones (qué estrategias de búsqueda emplean) para satisfacer con éxito cada una de sus necesidades terminológicas (como veremos en la parte III).

## **PARTE II. TÉCNICAS DE BÚSQUEDA**





### 3 Técnicas de búsqueda en los recursos y nuestra propuesta de clasificación

There are two types of knowledge. One is knowing a thing. The other is knowing where to find it.

(Samuel Johnson, escritor y lexicógrafo inglés, 1709-84)

En este capítulo vamos a analizar las técnicas de búsqueda que ofrecen una serie de recursos. La evolución de las tecnologías de la información ha propiciado el desarrollo de opciones de búsqueda innovadoras que se están incorporando en los recursos y ha permitido la rápida transición de los diccionarios en papel a diccionarios electrónicos, así como la aparición de nuevas herramientas que están transformando los modos de búsqueda tradicionales.

Sin embargo, la diversidad de opciones de búsqueda dificulta su clasificación. La sistematización de las técnicas de búsqueda en los recursos favorecería tanto el proceso de enseñanza como de aprendizaje, y también facilitaría la labor de lexicógrafos y terminólogos de creación de nuevos diccionarios electrónicos más estandarizados. A pesar de estas reflexiones que aparecen en la bibliografía, no hemos encontrado trabajos que establezcan una clasificación u ordenación de las técnicas de consulta en los recursos con una vocación *universal*, es decir, válida para ser aplicada a las herramientas electrónicas en general, y adaptable a cualquier recurso particular. Una clasificación que sirva para especificar y destacar las funciones de búsqueda que presenta un recurso y las que no presenta. Esta clasificación podría servir también como hoja de ruta para diseñar las formas de búsqueda en la creación de nuevos recursos electrónicos.

En el trabajo que estamos desarrollando, dentro del proyecto ONTODIC y ONTODIC II, abordamos la reflexión sobre cómo incorporar nuevas técnicas de consulta en diccionarios, como por ejemplo las consultas onomasiológicas (Alcina 2009a). Esto nos ha llevado a darnos cuenta de que cada recurso establece sus propias denominaciones para referirse a las diversas formas de búsqueda que implementa. Los manuales también emplean distintas nomenclaturas para designar opciones de búsqueda similares, tanto en recursos del mismo tipo (por ejemplo entre diccionarios) y especialmente entre recursos de distinto tipo (por ejemplo entre diccionarios, corpus e Internet). Además, en ocasiones una denominación de una forma de búsqueda en un recurso y en otro corresponden a formas de búsqueda sensiblemente diferentes. Ello acarrea dificultades de cara a la comparación de

las herramientas de búsqueda y a la enseñanza de su uso. Esta situación tampoco favorece que en la colaboración entre lexicógrafos, terminólogos y técnicos informáticos que desarrollan recursos estos se entiendan fácilmente, o que los docentes que enseñan a los futuros lexicógrafos, terminólogos, filólogos o traductores puedan utilizar una clasificación estándar que sea útil para todo el conjunto de técnicas de búsqueda.

De ahí que resulte necesario establecer una clasificación y nomenclatura homogénea a la hora de evaluar los recursos, valorar sus funciones, enseñarlas a sus usuarios o diseñar nuevos recursos. Nuestro primer objetivo fue pues analizar las formas de búsqueda que ofrecen actualmente los diccionarios electrónicos (Pastor y Alcina 2010, en prensa), así como los corpus electrónicos (Pastor y Alcina 2009) y los recursos de Internet (Pastor y Alcina 2011).

En la selección de los recursos para el análisis hemos tenido en cuenta que fueran recursos que suele emplear un traductor en sus búsquedas. Por ello, hemos examinado los tres tipos de recursos más destacados en el apartado 2.4 donde abordamos clasificaciones de recursos para traductores.

Es necesario conocer en primer lugar qué es lo que ofrecen los recursos desde el punto de vista técnico (cuáles son las posibilidades de búsqueda que existen en los recursos), para poder determinar posteriormente qué uso hacen de los recursos los traductores en función de sus necesidades por medio de estrategias de búsqueda, que trataremos en el capítulo 4.

### **3.1 Elementos de la clasificación de técnicas de búsqueda**

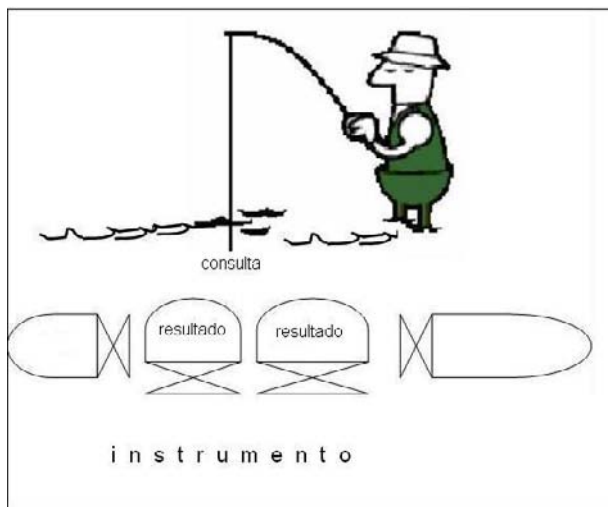
A través del estudio de las opiniones de otros investigadores y del análisis de uso de un conjunto de recursos, hemos sintetizado las técnicas de búsqueda en diccionarios electrónicos, corpus e Internet en torno a tres grandes ejes: la *consulta (query)*, el *instrumento (resource)* y el *resultado (result)*. Estos tres elementos reúnen todas las posibilidades de búsqueda que hemos observado en los recursos. Con esta clasificación, pretendemos contribuir a resolver el problema de la diversidad de nomenclaturas, tipos y subtipos de formas de búsqueda que encontramos en cada recurso en particular.

Las técnicas de búsqueda son opciones que ofrecen los recursos para obtener un determinado resultado. Un recurso puede interrogarse introduciendo una determinada búsqueda, a la que denominamos *consulta*. El recurso electrónico al que se interroga y el elemento concreto del recurso electrónico al que el usuario interroga es lo que llamamos *instrumento*. Por último, al realizar una consulta en un recurso, el usuario obtiene un

*resultado*, que es el tercer elemento de nuestra clasificación. Por tanto, distinguimos tres elementos en una técnica de búsqueda: una consulta, un instrumento y un resultado. La consulta es una palabra o expresión que introduce el usuario en la interfaz de un recurso. El instrumento es el recurso o parte de un recurso en el que se busca una palabra o expresión de búsqueda. El resultado de la búsqueda es el elemento al que se accede cuando se consulta un recurso.

En este punto cabe señalar que nuestro estudio de técnicas de búsqueda se centra en las búsquedas a partir de la realización de una consulta (*query*). En ocasiones nos referiremos también, aunque brevemente, a las búsquedas por exploración (*browsing*) y navegación (*navigating*)<sup>16</sup>.

En la Ilustración 16 representamos una técnica de búsqueda en la imagen de un pescador, que introduce una consulta que utiliza como cebo, en un instrumento que en este caso es el agua para pescar un pez que sería el resultado. La metáfora del pescador pretende ilustrar una técnica de búsqueda, en la que se ven involucrados los tres componentes de la clasificación<sup>17</sup>.



**Ilustración 16. Representación de técnica de búsqueda en los recursos.**

<sup>16</sup> Ténganse en cuenta las siguientes definiciones: «A query is a formulation of a user information need. In its simplest form, a query is composed of keywords and the documents containing such keywords are searched for. [...] a query can be (and in many cases is) simply a word, although it can in general be a more complex combination of operations involving several words» (Baeza-Yates y Ribeiro-Neto 1999: 100). «Navigation refers to following a chain of links, switching from one view to another, toward some goal, in a sequence of scan and select operations. Browsing refers to the casual, mainly undirected exploration of information structures, and is usually done in tandem with selection, although queries can also be used to create subcollections to browse through» (Baeza-Yates y Ribeiro-Neto 1999: 265).

<sup>17</sup> También otros autores, como por ejemplo Maia (2000: 56), se refieren a la pesca en el estudio de técnicas de búsqueda en los recursos. Esta autora señala que más que *surfear* en Internet debería hablarse de *pescar* en Internet, donde existen determinadas *técnicas* que sirven para obtener la información.

A continuación, explicamos con más detalle las técnicas de búsqueda, primero en diccionarios electrónicos (apartado 3.2 y subapartados), luego en corpus (apartado 3.3 y subapartados) y finalmente en Internet (apartado 3.4 y subapartados). Para cada uno de los recursos (diccionarios, corpus e Internet) empezamos con un análisis de la bibliografía para observar en qué punto se encuentra el estudio de las formas de consulta. A continuación examinamos una serie de diccionarios, corpus y recursos de búsqueda en Internet que se encuentran disponibles actualmente. A partir de estos análisis generamos clasificaciones homogéneas de técnicas de búsqueda y aportamos ejemplos concretos de cómo se implementan estas técnicas de búsqueda en los distintos recursos que hemos analizado.

### **3.2 Diccionarios electrónicos**

En este apartado entendemos por diccionarios electrónicos aquellos recursos léxicos en formato electrónico (algunos disponibles también en línea) que se alejan del modelo clásico de diccionario, lo que significa que no son meras versiones de diccionarios en papel disponibles en formato electrónico. Estos diccionarios reciben también otros nombres en la bibliografía, por la forma de almacenamiento no lineal de la información en bases de datos, o también por tratarse en su mayoría de recursos en línea que se actualizan constantemente, razón por la cual los autores Jousse et al. (2008) hablan por ejemplo de *sites lexicaux*.

En las siguientes subsecciones reflexionamos, en primer lugar, en el apartado 3.2.1, sobre la utilidad del diccionario como recurso de búsqueda, lo que justifica el análisis de este recurso. En segundo lugar, recopilamos ideas que se han manifestado en la bibliografía en relación con las características de los diccionarios electrónicos y las formas de consulta, especialmente en contraposición al diccionario en papel (apartado 3.2.2).

A continuación (apartado 3.2.3), nos hemos ocupado de examinar las posibilidades de búsqueda que permiten algunos diccionarios electrónicos disponibles. Además de las búsquedas clásicas más conocidas, nos interesamos especialmente por las más innovadoras, a partir del análisis en particular de una serie de diccionarios electrónicos.

Seguidamente, en el apartado 3.2.4, presentamos la clasificación de técnicas de búsqueda en diccionarios electrónicos que hemos elaborado, basándonos en las reflexiones de los autores y el análisis empírico de diccionarios. En esta clasificación, como se verá, utilizamos los tres elementos básicos de los que partimos para describir las diferentes técnicas de acceso, que hemos denominado: consulta, instrumento y resultado. A partir de la combinación de estos tres elementos y sus subtipos se estructuran todas las técnicas de acceso que hemos encontrado en los diccionarios.

Finalmente, en el apartado 3.2.5, incluimos un resumen y una valoración de nuestra clasificación de técnicas de búsqueda en diccionarios electrónicos.

### **3.2.1 Utilidad del diccionario como recurso de búsqueda, según la bibliografía**

Muchos autores coinciden en que el diccionario es una obra de referencia obligada y de indudable valor para los traductores (Newmark 1998: 29; Roberts 1996: 30) y suele ser la primera que utiliza el traductor en una secuencia de búsqueda (Gonzalo García 2004: 297). De hecho, Williams (1996: 276) señala que la consulta en el diccionario se emplea en dos tercios de los problemas de traducción.

Sin embargo, múltiples trabajos disponibles en la bibliografía que han abordado el uso de los diccionarios coinciden en que los usuarios tienen dificultades cuando consultan el diccionario, fundamentalmente por dos razones: el diccionario no facilita el acceso a la información (los diccionarios presentan deficiencias y no son intuitivos), y los usuarios desconocen cómo utilizar los diccionarios para acceder a la información (Béjoint 1989b: 280; Cowie 1999: 188; Fernández-Pampillón Cesteros y Matesanz del Barrio 2006: 150; Hartmann 1999: 40; Nesi y Hail 2002: 285; Sánchez Ramos 2004a: 482). No en vano, afirmaciones de este tipo son habituales: «Swearing at the dictionary must be any translator's favourite pastime» (Newmark 1993: 82); «Translators often enjoy a love-hate relationship with the main tools of their trade: dictionaries» (Rogers y Ahmad 1998: 193).

Estas dificultades implican que los diccionarios no se exploten de manera adecuada y se saque completo provecho de las posibilidades de búsqueda que ofrecen y los usuarios se planteen utilizar otros recursos, como Internet y los corpus, para acceder a la información que necesitan. Por ejemplo, Varantola (2002a: 173) señala que los traductores muestran una insatisfacción respecto de diccionarios y glosarios ya que en ocasiones no les proporcionan el acceso a información que puede resultar vital en el proceso de toma de decisiones, incluso en los casos en los que el traductor ha sido capaz de identificar el problema.

Frente a esta situación, algunas autoras sugieren incorporar en los estudios de traducción programas destinados a la enseñanza del uso del diccionario, como ha señalado Corpas Pastor et al. (2001a: 255): «[...] el diseño curricular de la licenciatura debería incluir una parte dedicada exclusivamente a los diccionarios, ubicándola dentro de las distintas fuentes de documentación a disposición del traductor e intérprete»; y también Roberts (1997): «[...] 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». Es evidente que el éxito de una búsqueda depende en gran medida de una utilización inteligente de las herramientas de consulta (Varantola 1998: 180), y que los traductores tienen que conocer las limitaciones y las virtudes de sus diccionarios y tienen que saber utilizarlos de una forma eficaz (Fraser 1999: 32).

Por otra parte, las mismas voces que abogan por un perfeccionamiento de los usuarios, igualmente demandan una mejora de los diccionarios. Por ejemplo, los diccionarios no siempre explotan todas las ventajas que ofrece el formato electrónico. En ocasiones, los diccionarios electrónicos presentan una sencilla traslación del diccionario en papel al formato electrónico, como un simple texto en el que las distintas informaciones aparecen con estilos y fuentes diferentes, pero sin que haya un aprovechamiento de las posibilidades que ofrece un sistema de base de datos avanzado. En este sentido, tampoco parece que se esté evolucionando tan rápido como se desearía en el desarrollo de diccionarios que incorporen nuevas técnicas de consulta, como por ejemplo las consultas onomasiológicas (Alcina 2009a: 37).

### **3.2.2 Las formas de búsqueda en diccionarios electrónicos, según la bibliografía**

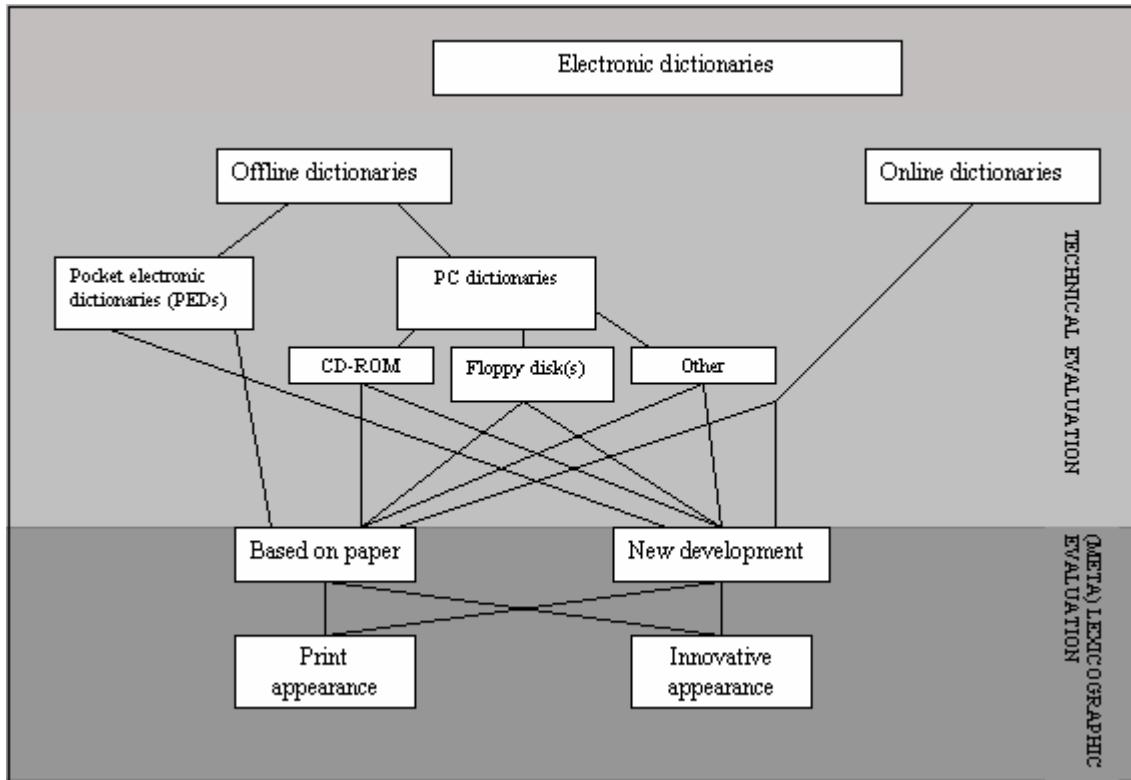
Con el desarrollo de las nuevas tecnologías y la aparición de Internet los diccionarios han cambiado progresivamente. Muchos diccionarios en papel han cambiado a formatos electrónicos, como el CD-ROM, y otros también están disponibles en línea<sup>18</sup>.

Los diccionarios electrónicos pueden agruparse en distintos tipos dependiendo de diversos criterios. Algunos autores han elaborado tipologías de diccionarios electrónicos (Ide 1993; Nesi 1998, 1999, 2000a; Sharpe 1995). Sirva de ejemplo la clasificación aportada por Lehr (1996: 315) en función de parámetros técnicos y metalexigráficos. Según los criterios técnicos, distingue entre diccionarios que están disponibles en línea o fuera de línea (diccionarios para PDA o diccionarios para ordenador en CD-ROM, entre otros). Según criterios metalexigráficos, distingue, por un lado, adaptaciones electrónicas de diccionarios en papel y diccionarios electrónicos genuinos (concebidos

---

<sup>18</sup> Por ejemplo, Kosem (2011: iv) señala lo siguiente: «It is now the time of online dictionaries and dictionary apps».

desde el principio para el medio electrónico) y, por otro, diccionarios con apariencia impresa o con aspecto innovador.



**Ilustración 17. Clasificación de diccionarios electrónicos según Lehr (1996: 315, traducida del alemán al inglés en De Schryver 2003:148).**

Dentro de esta clasificación, se distingue entre diccionarios electrónicos nuevos (*new development*) y versiones electrónicas de diccionarios en papel (*based on paper*) (Gross 1997: 55; Jacquet-Pfau 2002: 90).

La mayoría de autores (Abate 1985: 280; Nesi 2000b: 140; Somers 2003: 19) coinciden en que son más efectivos los diccionarios plenamente electrónicos que los diccionarios electrónicos creados a partir de un diccionario en papel: «Electronic dictionaries would be most effective if they were designed from scratch with computer capabilities and computer search mechanisms in mind» (Nesi 2000b: 140). Los diccionarios electrónicos se actualizan fácilmente (Carr 1997: 214; Harley 2000: 85; Kay 1984: 461) y además permiten una consulta más rápida, precisa y exhaustiva en la que pueden combinarse diversos criterios de búsqueda (De Schryver 2003: 157; Jacquet-Pfau 2002: 99). También se sugiere que las nuevas versiones de diccionarios electrónicos puedan ser útiles para su uso tanto por humanos como por máquinas (Gala y Lafourcade 2011: 89).

Como afirma Sharpe (1995: 48), las adaptaciones electrónicas de diccionarios en papel adoptan los mismos tipos de búsqueda que los diccionarios impresos y ofrecen, por tanto, menor flexibilidad en las formas de búsqueda. Lew (2011: 238 y 242) señala incluso la existencia de algunos diccionarios en línea más incómodos de consultar que los diccionarios en papel porque no utilizan adecuadamente el potencial que ofrece el formato electrónico.

Algunos autores señalan la evolución que han experimentado los diccionarios a lo largo de los años y las diferencias entre los distintos formatos (Poirier 1989). De Schryver (2003: 153-159) determina ventajas tanto de diccionarios en papel y electrónicos, tanto fuera de línea como en línea. Los diccionarios en papel resultan más familiares y fiables, tienen un valor simbólico como un objeto propio valioso, pueden utilizarse por el simple placer de su lectura, son mejor para la vista, la anotación es sencilla, se puede escribir en ellos, perduran en el tiempo y se pueden transportar, sin miedo a que sufran daños o pierdan funcionalidad, su existencia es independiente, no requieren la conexión de cables de ningún tipo, ni del acceso a un ordenador, pueden utilizarse desde el mismo momento de la adquisición sin necesidad de un ordenador, de la compra de actualizaciones o de tener que estar conectado en línea.

El diccionario electrónico fuera de línea incluye información sobre pronunciación y archivos de sonido, permite comparar la pronunciación del usuario con la almacenada en el diccionario, ocupa muy poco espacio y no pesa, el acceso es muy rápido, ofrece formas alternativas de mostrar la información, por medios gráficos, dispone de distintos métodos para la búsqueda, una búsqueda aproximada, multiacceso, hiperenlaces, facilita la interactividad y la personalización del diccionario, utiliza interfaces intuitivas y permite emplear la función de copiar y pegar.

El diccionario en línea tiene el atractivo de ser una novedad, es más ecológico, se libera a los usuarios del orden alfabético, se eliminan las restricciones del texto linear, se optimizan las referencias cruzadas, no hay restricciones de espacio y cantidad de información, se ofrece acceso a gran cantidad de información en corpus, así como a secuencias de video y animación, se vincula a otros recursos, la información está actualizada, permite la consulta simultánea en distintos diccionarios y los costes de uso son muy reducidos o gratuitos.

Forget (1999: 76-89), por su lado, señala que los diccionarios electrónicos se diferencian de los diccionarios en papel según criterios como uso, presentación, capacidades de búsqueda, aspectos técnicos y naturaleza de los contenidos (elementos



multimedia). En cuanto al uso, por ejemplo, esta autora menciona la utilización del hipertexto en los diccionarios electrónicos y una mayor flexibilidad en la consulta con respecto a los diccionarios en papel. También explica, sin embargo, que en los diccionarios electrónicos cada interfaz es diferente, lo que requiere emplear más tiempo en aprender a utilizar cada uno de los distintos diccionarios electrónicos disponibles.

En cuanto a la presentación, se observa el uso de distintos colores para destacar la información en los diccionarios electrónicos y el uso de interfaces más intuitivas. Respecto a las capacidades de búsqueda, se afirma que los diccionarios electrónicos implementan distintos tipos de búsqueda, que la búsqueda electrónica siempre es más rápida que la manual, pero se advierte de las posibles pérdidas de tiempo por probar distintas técnicas en el diccionario electrónico que no ofrecen resultados.

Finalmente, en lo que respecta a los aspectos técnicos, se señala la interactividad de los diccionarios electrónicos, que permiten al traductor añadir a la entrada del diccionario algún comentario que le resulte útil para futuras consultas como una nota de uso, un contexto, una traducción utilizada para un cliente en concreto, etc. Además, los traductores pueden copiar los términos directamente del diccionario sin necesidad de teclearlos otra vez en la traducción, los diccionarios electrónicos ocupan menos espacio que los diccionarios en papel y no se deterioran con el paso del tiempo.

El uso de diccionarios electrónicos y en línea conlleva numerosas ventajas frente a los diccionarios en papel. Sin embargo, el acceso a la información del léxico y la terminología contenido en los diccionarios plantea dificultades en el acceso eficaz a los datos que contiene, en parte debido al desconocimiento de los usuarios (incluso expertos en lenguas como los traductores) sobre las formas de consulta de los diccionarios para acceder a esa información, y en parte debido a la propia diversidad de formas de consulta (en diferentes campos, con diferentes comodines, guiados por interfaces muy variadas) que cambia de unos diccionarios a otros.

De todas las diferencias que hemos sugerido entre diccionarios impresos y electrónicos, la principal diferencia que señalan algunos autores y que queremos destacar en este apartado es la manera en la que se accede a la información, que en los recursos electrónicos es más eficaz (De Schryver 2003: 147; L'Homme y Leroyer 2009: 259; Nesi 1998; Sager 1993: 237). La búsqueda en los diccionarios en papel es limitada, ya que está restringida a la disposición de los contenidos por parte del lexicógrafo, que es siempre alfabética y se limita a la *búsqueda exacta* de un lema, sin flexiones ni derivaciones (Santana et al. 1996: 70). En cambio, la consulta en los diccionarios electrónicos es más

rápida y flexible, ya que se utilizan técnicas de búsqueda más avanzadas (Kaalep y Mikk 2008: 388; Sánchez Ramos 2004a: 181). Muchos diccionarios electrónicos actualmente permiten buscar información en todo el contenido del diccionario o buena parte de él y no solo en las entradas (Roberts y Langlois 2001: 712), por ejemplo en definiciones o contextos del diccionario.

Encontramos en la bibliografía algunos intentos de recopilar o clasificar las formas de búsqueda de los diccionarios electrónicos. Por ejemplo, Fernández-Pampillón Cesteros y Matesanz del Barrio (2006) distinguen los siguientes tipos de búsqueda. La *búsqueda en entrada* se efectúa mediante la introducción de la palabra exacta en la casilla de búsqueda. La *búsqueda en lista de entradas* es alfabética en orden creciente A-Z, o decreciente Z-A. En la *búsqueda asistida* el diccionario sugiere palabras al usuario cuando la palabra exacta introducida no se encuentra en el diccionario, se basa en sistemas de lematización y corrección ortográfica. La *búsqueda múltiple* combina distintos elementos de búsqueda y permite buscar en todo el contenido del diccionario para que se extraigan todos los artículos del diccionario que satisfacen una determinada búsqueda, por ejemplo localizar palabras en el diccionario a partir del contenido de los artículos (definición, ejemplos, etc.). Dentro de esta forma de búsqueda se incluyen *búsquedas con comodines*, *patrones definidos por el usuario* y *filtros* y *búsqueda de formas complejas de un lema*. La *búsqueda mediante palabras relacionadas* permitiría el uso del diccionario como tesoro, accediendo a relaciones de sinonimia, antonimia o familia de palabras. La *búsqueda de anagramas* recupera palabras que resultan de cambiar el orden de las letras introducidas en la búsqueda. La *búsqueda utilizando abreviaturas y marcas* junto con la búsqueda de palabras relacionadas favorece el procesamiento semántico de la información; recupera las entradas del diccionario a partir de las abreviaturas o las marcas de uso que se utilizan en el diccionario, que se pueden agrupar en función de criterios *categoriales*, *lingüísticos en general*, *distribución geográfica*, *vigencia cronológica*, entre otros. Por último, las *búsquedas más refinadas* o más precisas avanzan en el procesamiento semántico de la información mediante la comprensión semántica de la consulta del usuario y de la información léxica del diccionario.

Sánchez Ramos (2004a: 193-194) basándose en la clasificación anterior destaca nueve tipos de búsqueda. La *búsqueda en entrada*, la *búsqueda en lista de entradas*, la *búsqueda asistida*, la *búsqueda múltiple o avanzada*, la *búsqueda con comodines*, *patrones definidos por el usuario* y *filtros* que ayudan a combinar palabras de búsqueda (con los operadores AND y OR) o buscar variantes ortográficas (con los comodines \*, ?, #, etc.), la

*búsqueda mediante palabras relacionadas, la búsqueda de anagramas, la búsqueda utilizando abreviaturas y marcas y las búsquedas más refinadas.*

De Schryver (2003: 173-181) presenta un artículo sobre *los sueños de los lexicógrafos en la era de los diccionarios electrónicos*. En esta publicación recoge entre otras informaciones los avances observados en la bibliografía hasta ese momento en temas de accesibilidad en los diccionarios electrónicos. Presenta unas tablas en las que, por un lado, recopila las reflexiones de los autores sobre los avances que serían deseables en los diccionarios en distintos aspectos (*columna dream*) y, por otro lado, valora si la mejora propuesta se había implementado en los diccionarios en el momento en el que se sugirió (*columna then*) o en el momento de la redacción del artículo de este autor (*columna now*). Para ello se emplean los siguientes códigos (– no está implementada, ± parcialmente incorporada, ✓ existe en los diccionarios).

De Schryver clasifica las citas de los autores en cuanto a la accesibilidad del diccionario en cinco bloques: comunicación con la base de datos, búsqueda en toda la base de datos, acceso a la macroestructura, acceso a la microestructura y acceso a la medioestructura. A continuación presentamos estos bloques.

En cuanto a la primera parte, sobre la comunicación con la base de datos, el autor incluye sistemas de reconocimiento en los diccionarios para idiomas con caracteres logográficos, acceso automático al diccionario mediante una ventana emergente con solo hacer clic en una palabra, consultas en lenguaje natural, posibilidad de cambio entre distintos tipos de diccionarios (monolingües, bilingües o multilingües en un mismo sistema), posibilidad de completar artículos, bases de datos para consultas múltiples.

#	<i>Dream</i>	<i>Then</i>	<i>Now</i>
51	for a language with logographic characters: <u>pen-shaped OCR scanner as input to the keyboard, or light-pen input</u> that can be 'matched' by software (Sharpe 1993, 1995)	—	±
52	<u>reading pen which translates and pronounces words scanned in directly from the printed page</u> ( <i>Seiko Instruments Quicktionary</i> 1998)	✓	✓
53	<u>pop-up access</u> (i.e. the user only needs to point the mouse at (and sometimes single-click) the to-be-looked-up word(s)) (Kay 1983; <i>Babylon</i> 1997; <i>MoBiMouse</i> 1998; <i>Macmillan English Dictionary for Advanced Learners</i> 2002; <i>Oxford Genie</i> 2002; ...)	—	✓
54	information-handling devices able to deal with natural language (queries); <u>enter into real-language dialogue</u> with the database (Abate 1985; Zgusta 1991; Rogers 1996)	—	±
55	users should be in a position to switch, within one system, <u>D types</u> (mono-, bi- and multilingual) at will (Dodd 1988; Atkins 1996; Gamper & Knapp 2001)	—	—
56	transform the lexical database development environment into a lexical database <u>consultation environment</u> (Amsler 1984)	—	±
57	the consultation-reading of a database should give way to <u>exploration-research</u> of that database and thus the possibility to <u>complete D articles</u> (Gorcy 1990)	—	±
58	many a D of the future may well be a <u>multiple queriable database</u> (Docherty & Heid 1998)	—	±
59	<u>printed D enriched by the corresponding e-database</u> (Zgusta 2000)	—	±

**Tabla 6. Sueños sobre accesibilidad parte 1: comunicarse con la base de datos (De Schryver 2003: 174).**

Sobre la búsqueda en toda la base de datos, se refiere al acceso a elementos relacionados a partir de una lista de lemas (significados, ortografía, etimología, etc.), búsquedas paradigmáticas, búsquedas con filtros, búsquedas en todo el contenido con ayuda de operadores booleanos; nuevas formas de acceso a la información (tiene cierta etimología, es de determinado año o siglo, se emplea en un estilo concreto, etc.), combinación con búsqueda aproximada, diferentes operadores y analizadores morfológicos y sintácticos; acceso a los significados de forma directa por búsquedas complejas o aproximadas de grupos de palabras que contengan las mismas características fonológicas, sintácticas o semánticas; coexistencia en un mismo sistema de acercamientos semasiológicos y onomasiológicos, búsqueda temática estructurada.

#	<i>Dream</i>	<i>Then</i>	<i>Now</i>
60	create the ultimate D that can be used by the entire language community, through a lemma-sign list of limitless extent and the potential to <u>bring together related items</u> (e.g. senses, spellings, etymologies, etc.) <u>drawn from anywhere in the D</u> (Abate 1985)	—	—
61	paradigmatic search: <u>access directly and purposively right-hand fields</u> (in addition to 'in-and-out' consultation of <u>left-hand side</u> ) (Knowles 1990; ...)	—	✓
62	' <u>filtered searches</u> ' (i.e. selection of data types to be included in a search); full text searches <u>with the aid of Boolean operators</u> (Poirier 1989; ...)	—	✓
63	a multitude of <u>new routes to the data</u> (...), has an etymology of D, dates from year/century E, is used in style of F, ... in combination with <u>fuzzy matching, various operators, and morphological and grammatical analysers</u> (Dodd 1989)	—	±
64	various enumerations like: <u>heterographic homophones, (L1-sensitive) paronyms and confusibles, words within or outside a particular frequency, familiarity or polysemy band, words whose associated visual representations</u> (in a multimedia D) <u>meet certain conditions, etc.</u> ; all such criteria with <u>multiple Boolean searches</u> (Sobkowiak 1999)	—	±
65	<u>access meanings directly, fuzzy or complex searches for groups of words containing the same phonological, syntactic or semantic features</u> (Abate 1985; Webster & Ning 1996; Nesi 2000a)	—	—
66	within the framework of a large lexical database system: successful coexistence of <u>semasiological and onomasiological approaches, thematically structured search path</u> (Kay 1983; <i>Microsoft Encarta 99 Encyclopedia – World English Edition</i> 1998; Geeraerts 2000)	—	±

**Tabla 7. Sueños sobre accesibilidad parte 2: búsqueda en toda la base de datos (De Schryver 2003: 176).**

En cuanto a la parte 3 de acceso a la macroestructura, se libera a los usuarios del orden alfabético, se guía por medio de búsquedas visuales en aquellos idiomas con caracteres logográficos, se ayuda con las incorrecciones ortográficas introducidas<sup>19</sup>, se permite un acceso fonético, búsquedas con sugerencias (correcciones ortográficas, homófonos, formas flexivas), funciones de uso de comodines, índices inversos y anagramas, búsquedas fonéticas, recuperación rápida haciendo clic en una palabra o tecleando solo las primeras letras de una palabra, también con analizadores morfológicos,

<sup>19</sup> En un estudio de los principales diccionarios electrónicos monolingües de inglés para estudiantes se concluye que los diccionarios todavía no han avanzado suficiente en la corrección de las consultas ortográficamente incorrectas (Lew y Mitton 2011: 173).

inclusión de todas las formas flexivas de un lema en el diccionario, guiar al usuario en un determinado significado por medio de ventanas con homónimos y colocaciones.

#	<i>Dream</i>	<i>Then</i>	<i>Now</i>
67	liberate users from alphabetical order (Dodd 1988, 1989; Atkins 1996; Leech & Nesi 1999)	+	✓
68	for a language with logographic characters: provide (incorrect/dummy) lemma signs for those characters whose stroke numbers are commonly miscounted and <u>re-route the user</u> to the correct one (visual look-up) (Sharpe 1995)	—	±
69	help with poor spelling: through <u>fuzzy spelling</u> options (spelling corrections) and browsing links from <u>sight-words</u> (i.e. words users recognise the shape of) (Corris et al. 2000)	±	±
70	<u>phonetic-access D (PAD)</u> (i.e. access 'through sound', with wildcards for certain segments), <u>phonetically-aware fuzzy search algorithms</u> (both strictly linguistic and algebraic fuzzy searches) and <u>phonetic similarity searches</u> (Kay 1983; Sobkowiak 1994, 1996, 1998, 1999; <i>Macmillan English Dictionary for Advanced Learners</i> 2002)	—	±
71	macrostructural retrieval through <u>suggestive searches</u> (spelling corrections, near-homophones (by means of <u>diphone</u> data), <u>inflectional forms</u> ) (Gecraerts 2000)	✓	✓
72	<u>edutainment</u> : search functions featuring wildcards, reverse indexes and anagrams; <u>rhyiming D</u> (by means of <u>diphone</u> data) (Gecraerts 2000)	✓	✓
73	new means to access data: <u>start keying</u> an item and select from a list, a <u>phonetic tool</u> to find the correct spelling, searches with <u>wild cards</u> , an <u>automatic conjugation tool</u> (Duval 1992)	±	±
74	speedy retrieval through the possibility to click on a word and/or <u>focus-in typing</u> (i.e. to type the first letter(s) of a word) (Selva & Chanier 1998; Landau 2001)	±	±
75	speedy retrieval by means of a morphological analyser (Dodd 1988; ...)	—	±
76	inclusion of <u>all inflected forms</u> as lemma signs in the D (Corris et al. 2000)	—	—
77	guide users to a particular sense through the display of intermediate <u>homonym windows</u> and <u>collocation windows</u> (Selva & Chanier 1998)	±	±
78	<u>problems</u> with for example <u>homonymy vs. polysemy</u> status <u>cease</u> to be relevant (Dodd 1989)	—	±

**Tabla 8. Sueños sobre accesibilidad parte 3: acceso a la macroestructura (De Schryver 2003: 177).**

En cuanto al acceso a la microestructura, menciona la función de hipertexto, acceso directo a sustantivos compuestos, fraseología y modismos, comprobación de colocaciones, información sobre agrupaciones de palabras, desambiguación de los significados, intentos de proporcionar un equivalente adecuado en contexto.

#	<i>Dream</i>	<i>Then</i>	<i>Now</i>
79	hypertext functionality <u>eliminating linear text restrictions</u> ; <u>not everything</u> needs to be written/visual and <u>constant</u> (Abate 1985; Atkins 1996; ...)	—	±
80	<u>plan of an article</u> + full treatment per clickable sense (Selva & Chanier 1998)	±	±
81	compound nouns, phrases and idioms should be made directly accessible, retrieval of idioms and proverbs through a <u>phraseological dialogue box</u> (Abate 1985; Lemmens & Wekker 1990; Gecraerts 2000)	—	±
82	check whether two particular words collocate; list words with particular collocates (Lemmens & Wekker 1990; <i>Collins COBUILD English Collocations on CD-ROM</i> 1995)	—	✓
83	information about the use of every single word (cluster) with each other word, of every single resulting <u>word cluster again with each other word</u> , etc. (Grefenstette 1998)	—	—
84	automatic <u>disambiguation</u> of meaning (Zgusta 1991; ...)	—	+
85	attempt to provide the user with the correct equivalent for a word as used in the passage at hand (contextivity) (Michiels 2000; Whitelock & Edmonds 2000)	±	±
86	what if the computer was able to <u>analyse any word in context</u> , extract all the patterns that the word was found in, cluster them in a coherent way, and display these further examples succinctly for the user? (Grefenstette 1998)	—	±

**Tabla 9. Sueños sobre accesibilidad parte 4: acceso a la microestructura (De Schryver 2003: 179).**

Finalmente, en el acceso a la medioestructura, se incluyen ventanas emergentes con información gramatical, hiperenlaces con notas gramaticales, vínculos externos, encadenamiento de significados por medio de enlaces y referencias cruzadas.

#	<i>Dream</i>	<i>Then</i>	<i>Now</i>
87	grammatical pop-up windows (with inflectional paradigms+ explanatory text) (Sobkowiak 1998; Geeraerts 2000)	±	✓
88	<u>hyperlink grammar notes and the unavoidable abbreviations to specific grammar cards explained in the learners' mother tongue</u> (Dodd 1988; Abel & Weber 2000)	—	±
89	meticulous <u>D-external links</u> ( <i>Collins COBUILD on CD-ROM</i> 1995; <i>Collins COBUILD III on CD-ROM</i> 2001)	±	✓
90	<u>definition chaining</u> (aka hyperlinked cross-referencing, searches by chaining or hyperlinking, internal cross-referencing) (Duval 1992; Nesi 1999, 2000a; Geeraerts 2000)	±	✓
91	<u>cross-references</u> only showing the <u>proper sense</u> in a parallel (non-overlapping) window (Selva & Chanier 1998)	±	±

**Tabla 10. Sueños sobre accesibilidad parte 5: acceso a la medioestructura (De Schryver 2003: 181).**

Otros autores se han referido igualmente a la necesidad de avances como Knowles (1989: 1656) que señala el uso de hipervínculos, quizá una de las principales características de los recursos electrónicos, que encadenan aquellas palabras que están relacionadas con otras (Gómez González-Jover 2005: 161; Nesi 1999: 61). Hamon y Nazarenko (2001: 187) identifican varios tipos de hipervínculos: en el índice de entradas, en las palabras clave y en las relaciones de sinonimia e hiperonimia entre entradas y subentradas.

Dodd (1989: 89) también se refería a las posibilidades de los diccionarios electrónicos y preveía que se pudiera buscar una palabra a partir de la fonética, aproximaciones de escritura, etimología, área temática, relaciones semánticas con otras palabras (sinonimia, antonimia, hiponimia), palabras que se incluyen en la definición, categoría gramatical, etc.

Austermühl (2001: 96-97) señala que la búsqueda en recursos electrónicos ofrece toda una serie de ventajas, como la búsqueda con operadores booleanos y de proximidad, filtros (de área temática, fecha), búsqueda desde distintas aplicaciones, como por ejemplo buscar una palabra en un diccionario haciendo doble clic en la palabra del procesador, funciones de copiar y pegar, hipervínculos entre palabras relacionadas, añadir notas a una entrada, marcar entradas como favoritas, función de historial, información adicional en forma de vínculos a textos en CD-ROM o en línea.

Abate (1985: 278) añade el uso del lenguaje natural en las consultas. Nesi (2000b: 142) menciona la búsqueda simultánea en varios diccionarios. Geeraerts (2000: 80) nombra funciones de búsqueda por índices inversos, anagramas o similitud fonética.

Corris et al. (2000: 175-176) reclama funciones del diccionario electrónico que sugiera palabras similares ortográficamente o fonéticamente a la palabra de búsqueda<sup>20</sup>.

<sup>20</sup> En este estudio se proponen estas características para los diccionarios electrónicos porque se observan necesidades de hablantes de lenguas minoritarias, en concreto hablantes y estudiantes de lenguas

Respecto de la inserción de elementos multimedia en los diccionarios, otro progreso que se repite es el de la utilización de imágenes y gráficos (Abate 1985: 272; Lew 2011: 245). Los estudios señalan que resulta de gran utilidad incluir imágenes en los diccionarios, que complementen la información lingüística (Faber et al. 2007: 41; Montero Martínez y Faber 2008: 151; Prieto Velasco y López Rodríguez 2009: 179), tal y como se han implementado en las bases de datos *OncoTerm* y *EcoLexicon*, que analizaremos más adelante.

Por otro lado, los archivos de sonido pueden ser de utilidad para proporcionar información sobre la pronunciación de las palabras (Sobkowiak 1999). Lew (2011: 245) señala que los diccionarios que incluyen información sobre sonido, solo reproducen el sonido de la palabra que encabeza la entrada del diccionario, sin embargo, en escasas ocasiones se reproducen los contenidos de otros campos como las definiciones o los ejemplos, una excepción es el *Longman Dictionary of Contemporary English* en el que se pueden escuchar grabaciones de los ejemplos. Finalmente, este autor constata que los diccionarios no incluyen videos o animaciones, pese a que este tipo de archivos multimedia son frecuentes en la red (Lew 2011: 246).

Otra opción de los diccionarios electrónicos es la posibilidad de acceder a las últimas entradas consultadas por el usuario, similar a la función de *historial* en los exploradores de Internet (Forget 1999: 10; Rizo Rodríguez y Valera Hernández 2000: 369). También se ha sugerido que los usuarios puedan ejercer su control sobre los diccionarios y editar, añadir o eliminar contenidos en las entradas o incluso crear nuevas entradas (Somers 2003: 19).

Por otro lado, hay que tener en cuenta que el significado de un término viene determinado por el concepto al que se refiere pero también por el contexto en el que se enmarca (Corpas Pastor et al. 2001: 253; Kussmaul 1995: 89; Robinson 2003: 113). De ahí la importancia de la búsqueda de los términos dentro del contexto en el que se utilizan. Actualmente, los diccionarios, incluso los electrónicos, no disponen de un campo contextual tan amplio como para resolver todas las necesidades de los traductores.

Debido a la importancia de los contextos<sup>21</sup> para la traducción, algunos estudios abogan por integrar a los corpus dentro de los sistemas de búsqueda de los diccionarios

---

aborígenes de Australia, que tienen entre otros problemas la dificultad de buscar en un diccionario por orden alfabético (cuando el orden alfabético de las lenguas aborígenes es diferente). Aquí es donde resulta especialmente útil la función de sugerencia de palabras similares ortográficamente.

<sup>21</sup> Kugler et al. (1995: 13) indican que los traductores necesitan más información que solo los equivalentes en lengua meta de un término de la lengua origen. Los ejemplos contextuales son muy útiles.

electrónicos. De esta forma, los corpus ya no solo se ven como una fuente para la creación de diccionarios, o como un complemento o sustituto de los diccionarios (Colominas 2004: 362), sino como una parte más de los mismos (Sager 1993: 192).

A este respecto, son de especial interés trabajos como el de Castagnoli (2008) en la creación del proyecto *EOHS Term*, una base de datos en la que se elimina el campo de contexto de las entradas y los términos se hipervinculan a un corpus de referencia, en el que pueden consultarse los contextos. Incluso, es posible consultar un corpus como si fuera un diccionario, por medio de interfaces de búsqueda avanzadas, como por ejemplo en los recursos de colocaciones en línea *For Better English*<sup>22</sup> y *Just The Word*, donde el usuario introduce una palabra en la casilla de consulta y el recurso recupera los resultados sobre las colocaciones (que se han extraído del corpus) con una presentación en forma de artículo de diccionario, en el que es difícil apreciar que no ha sido elaborado manualmente (Lew 2011: 246).

Muchos de los avances propuestos en la bibliografía respecto a los diccionarios electrónicos se benefician de la creación de grandes bancos de datos, que han experimentado un desarrollo espectacular gracias al avance de las nuevas tecnologías. Tal y como indica Alcina (1997: 223) se ha pasado de fichas terminológicas en papel y microfichas a bancos de datos terminológicos en soporte informático (principalmente CD-ROM) y bancos de datos en línea. Asimismo, desde el punto de vista de la información hemos visto que se ha producido una transformación entre las primeras bases de datos terminológicas orientadas al término y los bancos terminológicos orientados al concepto, o bases de datos terminológicas basadas en el conocimiento, que permiten la representación formal del conocimiento terminológico y conceptual (Montero Martínez y Faber 2008: 157).

Cabré y Tebé (1998: 1) afirman que los bancos de datos destacan entre los demás recursos terminológicos de traductores, redactores técnicos o terminólogos porque son capaces de recopilar de manera organizada cantidades ingentes de registros terminológicos en distintas lenguas. Respecto a la información que puede encontrarse en un banco de datos, Gómez González-Jover (2005: 164) incluye: una *entrada*, compuesta por una denominación; *informaciones relativas al concepto*, ámbito temático, subámbito, definición, conceptos relacionados, remisiones (hipónimo, hiperónimo, etc.);

---

<sup>22</sup> Actualmente en versión demo, disponible en [forbetterenglish.com](http://forbetterenglish.com). Las colocaciones se extraen de *Sketch Engine*, que trataremos en el apartado sobre corpus 3.3.3.



*informaciones relativas a la denominación*, categoría gramatical, términos relacionados (sinónimos, abreviaturas), fraseología; *informaciones pragmáticas*, fiabilidad de los términos, frecuencia de uso; las *fuentes* o referencias bibliográficas que han servido para documentar toda la información; y los *datos de gestión o mantenimiento de la ficha*, número de registro, fecha de creación, etc.

Además pueden incluirse equivalentes en una o más lenguas, información gramatical, contextos, notas de uso (poco frecuente, arcaico, británico), o cualquier otro comentario que el traductor considere relevante para utilizar el término correctamente, incluyendo múltiples contextos o clientes, las formas más comunes de un término, por ejemplo las conjugaciones más usuales, e incluso gráficos (Bowker 2003b: 51-59).

En cuanto a la búsqueda en estas bases de datos, es posible efectuar búsquedas de palabras truncadas con el uso de comodines como el asterisco (\*), o búsquedas aproximadas (*fuzzy matching techniques*) con las que se accede a entradas similares a la de la búsqueda, como variantes morfológicas (diferentes formas de un verbo, con sufijos o prefijos), variantes ortográficas (también errores ortográficos), o términos compuestos incluso si el usuario no ha introducido el orden exacto de los componentes de un término compuesto (Bowker 2003b: 55).

Finalmente, en algunas bases de datos, integradas en programas de traducción asistida es posible efectuar una búsqueda automática de terminología, donde el programa compara los términos del texto original con los de la base de datos y muestra las entradas de la base de datos en el caso de que encuentre alguna coincidencia; también es posible realizar una búsqueda e inserción automática de la terminología de la base de datos a la traducción mediante la función de pretraducción (Bowker 2003b: 56-57).

En los bancos de datos los términos están asignados a un campo temático y la información está estructurada y automatizada. Esto permite al usuario realizar una búsqueda por área temática y otros criterios, como el idioma, por ejemplo. Además, los bancos de datos pueden contener otra información de distinto tipo: definiciones, contextos, ejemplos de uso y relaciones entre las distintas entradas (traducciones, sinónimos, hiperónimos, etc.) (Kugler et al. 1995: 50).

Como hemos visto en este apartado, las versiones electrónicas de los diccionarios ofrecen nuevas formas de búsqueda. En este sentido, hay autores que sugieren que, en algunos casos, estas funciones avanzadas pueden desarrollarse en un diccionario de manera que se permita buscar en un diccionario tradicional monolingüe (el típico diccionario semasiológico) de una forma onomasiológica, del significado a la palabra (Geeraerts 2000:

81; Oppentocht y Schutz 2003: 222). En palabras de Oppentocht y Schutz: «The kind of features that form the basis of the onomasiological search should be systematically identified». Últimamente, los esfuerzos están dirigidos a que un mismo diccionario permita ambos tipos de búsqueda, tanto semasiológica como onomasiológica (Díez Orzas 2004: 426).

Para Oppentocht y Schutz (2003: 222) y Geeraerts (2000: 82) las características que constituyen la base de una búsqueda onomasiológica pueden abarcar distintos tipos de información (definiciones, etiquetas, sinónimos, antónimos, información etimológica, etc.), que caracterizan las unidades léxicas que son objeto de la búsqueda (tanto palabras como expresiones compuestas por más de una palabra, que pueden corresponderse tanto con la entrada de un artículo, una expresión idiomática o incluso una cita, o una colocación). En la Tabla 11, Oppentocht y Schutz (2003: 222) proponen un resumen de las principales características identificadas que pueden ser la base de una consulta onomasiológica (en el eje vertical) y en el eje horizontal incluyen ejemplos de los posibles valores de estas características, que se pueden introducir en un diccionario para acceder a una unidad léxica.

	Search for all words	Search for all 'examples'
Form	ends on <i>able</i>	contains the words <i>cat</i> or <i>dog</i>
Part of speech / Type	verbs, nouns	proverbs, similes, collocations
Label	informal, obsolete	medical, euphemistic
Etymology	< Italian	since 18 <sup>th</sup> century
Explanatory text	contains <i>horse</i>	contains friendship
Word field	antonymous with <i>good</i>	synonymous with <i>home sweet home</i>

**Tabla 11. Onomasiological search matrix, según Oppentocht y Schutz (2003: 222, adaptado de Geeraerts 2000: 82).**

Según Geeraerts (2000: 83) estas matrices de búsqueda permiten dar respuesta a determinadas preguntas siempre y cuando el usuario sea capaz de formular las consultas de forma flexible e ingeniosa.

The onomasiological search matrix [...] enables the dictionary user to actively look for linguistic means of expression. Given a certain flexibility and inventiveness in formulating queries on the part of the user, highly specific questions like the following may now be easily answered.

What is the specific name of those heavy towers that one finds in medieval castles and fortresses? Looking for definitions containing *toren* "tower" and *burcht* "fortress" or *kasteel* "castle" leads to the word *donjon*.

Does Belgian Dutch have a specific variant for the word *overgordijnen*? Looking for words with the label *Belg.* and with *overgordijnen* as part of the definition may lead to *draperie*.

Are there any euphemisms for talking about cancer? Looking for the label *euf.* and the string *kanker* on the definitional level yields the abbreviated, euphemistic form *K*.

(Geeraerts 2000: 83)

En esta línea, según afirma De Schryver (2003: 157), el acercamiento a los diccionarios electrónicos onomasiológicos está cambiando de un enfoque en el tamaño de las bases de datos al desarrollo de *mecanismos de búsqueda inteligentes* donde en los diccionarios alfabéticos tradicionales se consulta dentro de los artículos para llegar al lema adecuado. Para ello, entran en juego el uso de palabras de búsqueda específicas, etiquetas, operadores booleanos, diversos campos de los artículos del diccionario, etc.

[...] the approach to onomasiological EDs has started to shift from a mere focus on database size to clever search mechanisms by which traditional alphabetically organised dictionaries are searched from within an article to the lemma sign [...]. Here an inventive use of specific search words, labels, Boolean operators, article fields to be searched, etc., can for instance lead from a combination of keywords in a definition to the item(s) one is looking for. Some prototype dictionaries being developed do the reverse in that they are basically onomasiological EDs, with an alphabetical index. A unique implementation in this regard is WordNet which can be seen as a sophisticated variant of the old thesaurus concept [...].

(De Schryver 2003: 157)

Sager (1993: 236) aboga por que los diccionarios informatizados permitan la flexibilidad de búsquedas de forma que sea posible el tipo de consultas que se muestran en la Tabla 13, que vayan más allá de las respuestas directas que se obtienen de las distintas partes de una entrada de un diccionario convencional (véase Tabla 12). Este autor señala que «la mayoría de los datos necesarios para responder a las preguntas presentadas aquí aparecen en los bancos de datos de términos actuales pero la recuperación de esta información requiere una exploración complicada de toda la base de datos, es decir, los datos están escondidos».

PARTE	PREGUNTA	RESPUESTA
género	¿Qué género tiene “imprimante”?	- femenino
ortografía	¿Cómo se escribe la palabra francesa cuyo equivalente inglés es “Woodruff Key”?	- clavette Woodruff
equivalente	¿Cómo se dice en francés “laser printer”?	- imprimante laser
definición	¿Qué es una “impresora láser”?	- una impresora que...
sinónimo	¿Puedo utilizar “bit” como sinónimo de “dígito binario”?	- dígito binario
marca de campo	¿“Bit” está restringido a un campo temático?	- abrev.: bit
ejemplo	¿Hay una oración de ejemplo que contenga “bit”?	- Informática

**Tabla 12. Preguntas y respuestas de un diccionario convencional (Sager 1993: 235).**

CONSULTA	BÚSQUEDA DEL CAMPO
- Compila un glosario de todos los términos con una nota de uso “ICI”	restricción de uso o nota de alcance
- ¿Cómo llamas a una máquina que realiza X?	definición o enlaces conceptuales
- ¿Qué partes componen Y?	términos partitivos subordinados
- Busca todos los términos introducidos por “JCS” desde 1985	nombre del terminólogo durante un período de tiempo
- Compila un glosario de términos relacionados con “GPSG”	campo temático o subdivisión
- Imprime todos los términos con una fuente CCL	subconjunto del origen de los datos

**Tabla 13. Tipo de consulta y de preguntas en un banco de términos informatizado moderno (Sager 1993: 236).**

Varantola (2006: 223) también habla de la *inteligencia* de los diccionarios, que puede ser una *inteligencia superficial*, en aquellos diccionarios que incluyen sistemas para ayudar a determinar la ortografía de una palabra, o que añaden enlaces en las entradas para encontrar sinónimos, cuasisinónimos, antónimos u otra información en general; mientras que la *inteligencia profunda* serían los diccionarios completamente personalizables, que permitieran acceder a perfiles de búsqueda definidos por el usuario, especificación de filtros y modos de visualización, modos de navegación, modos de consulta, visualización de información al completo o reducida, alertas al usuario. Además, que incluyeran información detallada de colocaciones, prosodia semántica, valencia semántica,

diferenciación de significados, acceso a corpus compatibles y a fuentes de información alternativas (enciclopedias, información temática, visualización multimedia). Finalmente, sistemas de aprendizaje que permitan al diccionario aprender de los registros de búsqueda de los usuarios.

En relación con lo anterior, algunos autores abogan por diccionarios onomasiológicos de colocaciones (Jousse et al. 2011; L'Homme y Leroyer 2009; L'Homme et al. 2010). Así lo indica Siepmann (2006: 8) en la descripción de *Bilexicon*:

[...] aims at producing bilingual onomasiological dictionaries whose main entry type will be of a collocational nature. This represents a break with the word-based lexicography still current in both semasiological and onomasiological approaches. Semasiological dictionaries tend to consist of an alphabetical word list leading the user from the word to its meaning, while onomasiological dictionaries allow the user to proceed from a particular concept and find the most appropriate word for it. Both types of dictionary are therefore mainly based on individual words – although, perforce, including phraseology in sub-entries and examples.

(Siepmann 2006: 8)

Según este autor, la mayoría de diccionarios se basan en el paradigma saussureano<sup>23</sup> de que la unidad básica de significado en el lenguaje es la palabra, por lo que la mayoría de diccionarios están compuestos de listas alfabéticas de equivalentes de palabra a palabra y se incluye muy poca información del comportamiento sintagmático de las unidades de una

---

<sup>23</sup> Para más información sobre las relaciones paradigmáticas y sintagmáticas, véase Alcaraz Varó y Martínez Linares (1997: 422-423): «Como miembro solidario de la DICOTOMÍA PARADIGMA/SINTAGMA. Saussure estableció en el *Curso de lingüística general* que todas las unidades (fonológicas, morfológicas, semánticas, etc.) de una LENGUA están vinculadas entre sí formalmente por medio de dos tipos de relaciones: relaciones asociativas y RELACIONES SINTAGMÁTICAS. Las primeras (cf asociación), fundadas en el ASOCIACIONISMO psicológico, unen signos o elementos de los mismos que, por tener una función lingüística (fonológica, sintáctica, semántica, etc.), son equivalentes y se caracterizan por su sustituibilidad; estas relaciones son de tipo vertical o *in absentia* y se las ha llamado **relaciones paradigmáticas** desde que Hjelmslev (1938: 140) les dio este nombre a fin de evitar la connotación psicologista que conlleva el término *asociativo*. Saussure utilizó la palabra *enseñanza* para ilustrar este tipo de relaciones: la RAÍZ de *enseñanza* está relacionada con *enseño, enseñas, enseñamos*, etc.; y el SUFIJO *-anza* tiene relaciones paradigmáticas con *matanza, tardanza*, etc.; desde un punto de vista semántico, *enseñanza* está relacionada con *estudio, docencia, educación*, etc., y fonéticamente la palabra *enseñanza* está relacionada a cualquier palabra que acabe en *-anza*. El valor de cada unidad o elemento viene dado por la OPOSICIÓN que mantiene con con [sip] los demás elementos del paradigma.

Por el contrario, son **relaciones sintagmáticas** las que, basándose en el carácter LINEAL del lenguaje, mantienen las unidades lingüísticas de un mismo NIVEL en la sucesión de la CADENA HABLADA o escrita; estas relaciones, que son complementarias de las anteriores, son de tipo horizontal o *in praesentia*, ya que necesitan de la alineación de un elemento al lado del otro; el valor de cada unidad o elemento viene dado por el CONTRASTE con los elementos que le preceden y los que le siguen. Cf acento.

Después de todo lo dicho, se deduce que un 'paradigma' es un conjunto de unidades que pueden conmutar entre sí, es decir, que pueden manifestarse en el mismo CONTEXTO y, por tanto, las unidades de un 'paradigma' están en OPOSICIÓN entre sí, o sea, rivalizan por manifestarse en el **sintagma**. Éste se define como el conjunto de unidades del mismo nivel lingüístico en relación de CONTRASTE. Los 'paradigmas' y los 'sintagmas' están sometidos a dos operaciones: la elección y la combinación. En el HABLA o actuación comunicativa, todo hablante **elige** una unidad de entre las que constituyen el conjunto del paradigma y la **combina** en el sintagma la unidad siguiente de otro paradigma».

lengua. Sin embargo, un aspecto decisivo del significado de una unidad es su *prosodia semántica* que se refiere a que las unidades léxicas se ven influidas por las connotaciones que transmiten las coocurrencias habituales que acompañan a una palabra. En un diccionario onomasiológico las colocaciones no estarían clasificadas alfabéticamente, sino por áreas temáticas o tipos de situaciones, por ejemplo *Noise: Telling people to be quiet: don't say a word / don't make a sound / be quiet / hush / quiet, please / shut up / wrap up / belt up / put a sock in it* (Siepmann 2006: 21). Este autor concluye que «[...] the onomasiological dictionary of the future will constitute a new kind of dictionary of synonyms to the extent that it will contain collocational rather than one-word synonyms [...]» (Siepmann 2006: 34).

En la misma línea, Almela et al. (2011: 7) en la elaboración del *Dynamic Combinatorial Dictionary*, señalan: «A further argument for the incorporation of collocational data refers to the connection of form and meaning. The syntagmatic behaviour of words is closely associated with their semantic properties. Therefore, collocation is more than a surface co-occurrence pattern; it also provides a representation of word meaning». En este sentido, se refieren a las *constelaciones léxicas*<sup>24</sup> para mejorar la utilidad de los diccionarios de colocaciones, donde existe una interacción entre las diferentes colocaciones de un nodo, a diferencia de en los diccionarios de colocaciones tradicionales, en los que las colocaciones están relacionadas únicamente con la entrada y no entre ellas. Según estos autores: «What combinatorial lexicography has so far failed to reflect is the fact that the choice of a collocation can also affect the choice of other collocations in its vicinity» (Almela et al. 2011: 8).

En este apartado hemos recogido algunas de las reflexiones que los autores han manifestado respecto a los diccionarios electrónicos y, en especial, destacando sus diferencias con los diccionarios en papel. En particular, han destacado aspectos como la interactividad y la variedad y flexibilidad de las formas de consulta. En este sentido, hemos visto que los avances técnicos facilitan el desarrollo de búsquedas onomasiológicas, del significado a la denominación. Los estudios que hemos incluido en este apartado han contribuido a la descripción de las características y las posibilidades en materia de

---

<sup>24</sup> «A lexical constellation is a collocational network hierarchically organised in two or more centres of lexical attraction» (Almela et al. 2011: 2). Por tanto, estos autores indican que va más allá de la estructura simple de una *collocational network*. Alonso et al. (2011: 19-20) utilizan *collocational networks*, *collocational resonance* y *Corpus Pattern Analysis-CPA* para la elaboración de *DicSci*, que se concibe tanto para la búsqueda semasiológica como onomasiológica.

búsquedas y, por ende, al desarrollo de los diccionarios electrónicos. Si bien es cierto que en la bibliografía se han utilizado diferentes criterios de clasificación. En ocasiones también se han mezclado criterios y se han adoptado distintos puntos de vista para describir a los diccionarios. Por ello, hemos observado una necesidad de profundizar respecto a cómo son las formas de búsqueda en los diccionarios y cómo se pueden organizar.

### **3.2.3 Análisis empírico de técnicas de búsqueda en diccionarios electrónicos**

En este apartado hemos realizado un análisis de las opciones de búsqueda de 30 diccionarios electrónicos. Puesto que en este estudio nos interesa abordar todas las formas posibles de búsqueda, hemos decidido analizar directamente solo diccionarios electrónicos, por lo que no incluimos diccionarios en papel en el análisis. Según hemos visto en el estudio de la bibliografía, el medio electrónico ofrece más posibilidades de constatar técnicas, ya que son más versátiles en las búsquedas. En este punto, queremos señalar que nuestro objetivo no es elaborar un catálogo<sup>25</sup> de diccionarios electrónicos, además hay que tener en cuenta que el número actual de diccionarios electrónicos disponibles supondría que analizarlos todos fuera prácticamente inabarcable.

El objeto de este análisis son las funciones de búsqueda. Por tanto, el criterio utilizado para la selección de los diccionarios electrónicos ha sido que los diccionarios incorporaran técnicas de búsqueda innovadoras, además de la tradicional búsqueda alfabética. Asimismo, hemos analizado tanto diccionarios en línea como en CD-ROM. La mayoría de diccionarios en línea son gratuitos por lo que no es necesario comprarlos para poder utilizarlos. Además, los diccionarios en línea se pueden consultar desde cualquier ordenador (siempre que tenga conexión a Internet) por lo que no es necesario que el diccionario esté instalado en cada ordenador en el que vayamos a consultarlo. Por otro lado, los diccionarios en CD-ROM normalmente ofrecen técnicas de búsqueda más avanzadas que muchos diccionarios en línea, y son estables y duraderos, en comparación con los diccionarios en línea cuya virtualidad supone que pueden cambiar de dirección web de un día para otro o incluso dejar de estar disponibles.

---

<sup>25</sup> En otros trabajos se ofrecen catálogos de diccionarios (Möhrs y Töpel 2011).

Asimismo, aunque en nuestro proyecto nos interesa sobre todo la terminología, hemos revisado no solo diccionarios terminológicos, sino también lexicográficos ya que lo que buscamos es una amplia representación de técnicas de búsqueda que, en definitiva, podrían aplicarse a cualquier tipo de diccionario.

En esta sección describimos, en primer lugar, diccionarios en CD-ROM, en español, como el *Diccionario de uso del español* de María Moliner y el *Diccionario de la lengua española* de la Real Academia; diccionarios en CD-ROM, en inglés, como el *Oxford English Dictionary* y el *Collins English Dictionary*; y el diccionario en francés en CD-ROM *Le Grand Robert de la Langue Française*. Después, analizamos diccionarios y bases de datos en línea<sup>26</sup> en diferentes idiomas (tanto monolingües como bilingües), que presentamos por orden alfabético: *Base Lexicale du Français (BLF)*, *Cercaterm*, *Collins BETA*, *Diccionari de la llengua catalana*, *Dicouèbe* y *diccionarios similares (DiCoInfo, DiCoEnviro, DAD)*, *Diccionario de Colocaciones del Español (DiCE)*, *Dirae*, *EcoLexicon*, *EOHS Term*, *FrameNet*, *Le grand dictionnaire terminologique*, *Genoma*, *Goodrae*, *IATE*, *Just The Word*, *Lexical FreeNet*, *Macmillan English Dictionary*, el diccionario monolingüe en inglés, tesoro y diccionario visual de *Merriam Webster*, *OncoTerm*, *OneLook Reverse Dictionary*, *TERMIUM*, diccionarios *Trésor de la Langue Française informatisé*, *Ultralingua*, *UNTERM*, *WordNet*, los diccionarios de *WordReference* y el diccionario *Wordsmyth*.

A continuación describimos brevemente las formas de búsqueda que implementan estos diccionarios centrándonos en las peculiaridades que presentan cada uno de ellos. Con ello pretendemos mostrar un panorama amplio de las diversas y variadas formas de consulta que se pueden encontrar en los diccionarios electrónicos, antes de pasar, en el apartado siguiente, a su sistematización.

El *Diccionario de uso del español (DUE)* es un diccionario monolingüe en español del lenguaje general, en formato CD-ROM, desarrollado por la lexicógrafa María Moliner. Permite realizar cuatro tipos de búsqueda:

- en la lista de entradas: el diccionario ordena la lista alfabética de entradas del diccionario por la palabra o secuencia de caracteres que introduce el usuario;
- en las entradas: el diccionario genera una lista de palabras que se asemejan por escritura a la introducida por el usuario, en el caso de que la palabra introducida no se

---

<sup>26</sup> Última revisión del análisis con fecha 20 de abril de 2012. A partir de esta fecha las versiones en línea de los diccionarios pueden haber experimentado cambios.



encuentre en el diccionario. Además, esta opción tiene la particularidad de que recupera la forma canónica de la palabra que introduzcamos (un verbo conjugado, un sustantivo en plural, etc.);

- en las definiciones: el diccionario genera una lista de palabras que contienen en su definición todas las palabras introducidas por el usuario con el operador (&) o alguna de las palabras introducidas con el operador (), o que contienen en su definición una «expresión exacta» introducida entre comillas; también es posible buscar palabras que no contengan en su definición las palabras introducidas con el operador (!);

- anagramas: el diccionario genera una lista de palabras que resultan de combinar en distinto orden las letras introducidas por el usuario.

El *Diccionario de la lengua española (DRAE)* es un diccionario monolingüe de lenguaje general en español, desarrollado por la Real Academia Española. Está disponible tanto en CD-ROM, como en línea. La versión en línea solo permite la búsqueda de una palabra exacta, de una palabra aproximada, sin signos diacríticos o por semejanza fonético-ortográfica.

El diccionario *DRAE* en CD-ROM implementa más tipos de búsqueda. Presenta algunas formas peculiares de consulta como la búsqueda alfabética inversa, que ordena alfabéticamente las entradas desde la última letra hasta la primera, en lugar de hacerlo desde la primera a la última. La búsqueda de formas complejas recupera, a partir de la palabra introducida por el usuario, todas las entradas que incluyen esta palabra. También es posible buscar las entradas que contienen una determinada palabra en el campo de definición u otro campo de la entrada. Otra de las funciones es la búsqueda de anagramas con las letras introducidas por el usuario, o añadiendo tres letras más, o descartando tres letras de las introducidas por el usuario. Este diccionario también contiene un índice de abreviaturas y marcas a través del cual se pueden localizar entradas que contienen una determinada abreviatura. Asimismo, es posible el uso de los siguientes comodines: ? (sustituye a cualquier carácter), \* (sustituye a cero caracteres o una secuencia de caracteres), + (sustituye a uno o más caracteres), [...] (busca cualquiera de las letras incluidas entre corchetes), [!...] (busca cualquier letra que no esté incluida detrás de la exclamación entre corchetes), @ (sustituye a una vocal), # (sustituye a una consonante).

El *DRAE* en CD-ROM presenta también la opción de búsqueda en árboles. Los árboles son estructuras jerarquizadas de información de distinto tipo (etimológica, gramatical, de ámbito geográfico de uso, registro, construcción y régimen o cuestiones

semánticas). Al seleccionar alguna de estas ramas, el diccionario nos muestra un listado de palabras cuyas entradas cumplen esa descripción.

Por otro lado, también en línea se encuentra disponible *Dirae*, un diccionario inverso basado en el *DRAE* en línea, que recupera las palabras del campo de entrada de este diccionario a partir de la búsqueda en los campos de definición. Puede introducirse más de una palabra de búsqueda y todas las palabras introducidas se tendrán en cuenta. También es posible buscar una frase entrecomillándola. El diccionario busca tanto las palabras que coinciden exactamente con la consulta introducida como las que comparten lexema (lematización o stemming); para buscar solo palabras exactas, deberá introducirse cada término de búsqueda entre comillas. Asimismo, añadiendo una tilde (~) al final de una palabra de la consulta, el sistema recupera también palabras relacionadas. Se puede excluir un término de los resultados poniendo el signo (-) delante. El sistema de búsqueda no tiene en cuenta las mayúsculas.

*Goodrae* es un diccionario que convierte todo el contenido del diccionario *DRAE* en línea en un hipertexto, de manera que se puede hacer clic en cualquier palabra de los resultados (por ejemplo en cualquier palabra de una definición) para buscarla automáticamente en el diccionario. Además, sus autores lo consideran un diccionario inverso porque, como *Dirae*, busca dentro de las definiciones. Pueden buscarse dos o más palabras combinadas, o una expresión exacta (si se entrecomilla). También permite truncar palabras con el uso de un asterisco. Además, lematiza las palabras de búsqueda y recupera su definición y palabras de la misma familia.

El *Oxford English Dictionary (OED)* es un diccionario monolingüe en inglés, en CD-ROM, desarrollado por Oxford University Press. Permite la búsqueda en la lista de entradas por palabras, secuencias de palabras, formas y fonética (words, phrases, variant forms, phonetics) y permite el uso de filtros de fecha y categoría gramatical. Asimismo, es posible buscar en todo el texto del diccionario, en los campos de definición, etimología y quotation. Este diccionario también permite el uso de los comodines \* (que sustituye a cero o más caracteres) y ? (que sustituye a un carácter).

La versión en línea de *Oxford Dictionaries* ofrece la búsqueda en el diccionario monolingüe en inglés introduciendo una palabra en la casilla de búsqueda, que tiene función de sugerencias. En los resultados se presenta el artículo del diccionario, sin vínculos. Es posible acceder a palabras cercanas.

El *Collins English Dictionary (CED)* es un diccionario monolingüe en inglés, disponible en CD-ROM y desarrollado por HarperCollins Publishers. Este diccionario

tiene dos opciones de búsqueda generales: la opción de búsqueda en el diccionario (en la lista de entradas) y la opción de búsqueda en el tesoro (busca palabras relacionadas). En el diccionario es posible buscar anagramas, buscar en las definiciones del diccionario con operadores (AND, OR, NOT y paréntesis) y con palabras truncadas con comodines (\*, ?). Asimismo tiene una función de historial (*history*) para acceder a las últimas consultas realizadas.

El **Collins BETA** es una versión en línea que contiene el diccionario monolingüe y un tesoro en inglés, además de tres diccionarios bilingües inglés-francés, inglés-alemán e inglés-español. En el diccionario monolingüe inglés, con la búsqueda de una palabra exacta en la casilla de búsqueda, se recupera la definición, formas alternativas, etimología, sinónimos (con acceso al tesoro), citas, ejemplos de uso, traducciones a diferentes idiomas junto con la posibilidad de escuchar la pronunciación en cada uno, también términos relacionados, información sobre la tendencia de uso de la palabra a lo largo de los años y fotos extraídas de *Flickr*.

**Le Grand Robert de la Langue Française** permite buscar introduciendo una palabra en la casilla de búsqueda (con función de corrección ortográfica) o alfabéticamente en la lista de entradas, formas y locuciones (pestaña *nomenclature*). Con la pestaña *explorateur* se visualiza el esquema de los artículos del diccionario. La pestaña *recherche* permite realizar una búsqueda más avanzada dentro de distintos campos del diccionario. En el campo de entrada (*entrées*) puede introducirse una secuencia de letras utilizando comodines (? sustituye a un carácter, \* a más de uno, & a una vocal y # a una consonante), da la opción de incluir todas las palabras, verbos conjugados, plurales y femeninos y respetar los acentos), también permite la búsqueda de las letras en el orden introducido o desordenadas (función de anagrama), o utilizar la función de corrección, finalmente permite restringir por categoría gramatical y por morfología. En el campo *phonétique* es posible realizar la consulta introduciendo caracteres fonéticos que pueden buscarse en el orden introducido o desordenados, también pueden emplearse comodines. En el campo *citations* se realiza la búsqueda dentro del campo de citas, y en el campo *texte intégral* puede realizarse la búsqueda en todo el artículo, en el campo de etimología, de definiciones, ejemplos y expresiones, sinónimos y marcas de uso y dominio. Finalmente, la pestaña *historique* da acceso al historial de consultas.

La **Base Lexicale du Français (BLF)** está basada en el *Dictionnaire d'apprentissage du français langue étrangère ou seconde (DAFLES)* y en otros diccionarios. La opción *Find information on a word* permite introducir una forma (tiene lematizador) y comprobar

el género, la ortografía, la conjugación verbal, el significado, sinónimos, antónimos, traducciones, ejemplos de uso, frecuencia, palabras de la misma familia, colocaciones, pronunciación, artículos que contienen dicha palabra en la definición y artículos que contienen dicha palabra en el contenido completo, así como búsqueda de la palabra en recursos de la Web. La opción *Find information on a word combination/expression* permite introducir dos palabras combinadas y acceder a su significado, así como a ejemplos de aparición en un corpus de periódicos de *Le Monde* de 1998, y a textos del Parlamento europeo y *Wikipedia* en la herramienta *Corpuseye*; además traducciones del francés al neerlandés, inglés, alemán, español y sueco.

La opción *Find the translation of a word* permite buscar equivalentes en el diccionario en línea Interglot.com del neerlandés, inglés, alemán, español y sueco al francés, y consultar los ejemplos en el corpus. *Find the translation of a word combination/expression* permite introducir dos palabras combinadas o una frase exacta y acceder a las traducciones extraídas del corpus OPUS del neerlandés, inglés, alemán, español y sueco al francés. La opción *Check* verifica el uso de una palabra, combinación de dos palabras o secuencia en *Google News*, también verifica la validez de un equivalente de un idioma a otro en el corpus OPUS. La opción *Help me*, que está en proceso de desarrollo, tiene como objetivo ayudar al usuario en la comprensión, traducción o escritura de un texto. Las opciones *Learn* y *Do exercises* proporcionan materiales para aprender francés.

***Cercaterm*** es un diccionario multilingüe en línea, desarrollado por el Centre de Terminologia TERMCAT. Contiene terminología en lengua catalana, y otras como española, inglesa, francesa, etc. La diferencia con otros diccionarios es que los términos en este diccionario están clasificados por área temática. La interfaz de búsqueda permite restringir el término de la búsqueda a un área temática determinada por el usuario.

La búsqueda avanzada facilita recuperar términos por proximidad, también especificar opciones de truncamiento, determinar el campo en el que se realizará la búsqueda (denominación, definición o nota). Asimismo, es posible restringir el idioma de búsqueda, la categoría gramatical y la jerarquía.

El ***Diccionari de la llengua catalana***, segunda edición, es un diccionario monolingüe en línea de catalán. Permite una búsqueda simple en las entradas del diccionario de palabras que coincidan, empiecen, terminen, contengan, no empiecen, no terminen o no contengan la expresión de búsqueda. La búsqueda avanzada permite especificar más parámetros u opciones de búsqueda. Se puede indicar al diccionario que tenga en cuenta la puntuación gráfica de la expresión introducida, así se respetarán por ejemplo los acentos.

Se puede especificar que se busque una subentrada dentro de una entrada. Se puede restringir la búsqueda por categoría gramatical, por lengua de origen de la palabra, marca valorativa, información morfológica y por una o más áreas temáticas. También es posible buscar palabras dentro de las definiciones o de los ejemplos.

El *Dicouèbe*<sup>27</sup> (desarrollado por Mel'čuk y Polguère, grupo de investigación OLST de la Universidad de Montreal) es la versión en línea del diccionario de colocaciones<sup>28</sup> en francés *DiCo*<sup>29</sup>. Se basa en la Lexicología Explicativa y Combinatoria (LEC), el componente léxico de la Teoría Sentido-Texto. Según la LEC, una unidad léxica tiene tres componentes: un significado, una forma, y un conjunto de propiedades combinatorias (Mel'čuk et al. 1995: 16). En este sentido, cada entrada del diccionario<sup>30</sup> describe un único significado o acepción de una palabra, lo que se denomina unidad léxica. Las distintas acepciones se identifican con números romanos, cardinales y letras.

Por cada unidad léxica se describen las relaciones paradigmáticas (principalmente relaciones de tipo semántico, como por ejemplo las derivaciones semánticas) y las relaciones sintagmáticas (o también colocaciones). El etiquetado de este tipo de relaciones se efectúa por medio de funciones léxicas. Una función léxica (f) se aplica a una unidad léxica (L) para recuperar un conjunto de unidades léxicas que están asociadas al significado y a la configuración sintáctica de uso de f respecto de L. Por ejemplo, aplicando la fórmula f(L), dada la función léxica de sinonimia Syn y su aplicación a la unidad léxica *réfrigérateur*, *Syn(réfrigérateur)* designa los sinónimos de la unidad léxica *réfrigérateur*, así *Syn(réfrigérateur) = frigidaire, fam. frigo*.

Además, el significado de las unidades léxicas se describe por los participantes que intervienen en la situación que designa dicha unidad léxica, lo que se denominan actantes

---

<sup>27</sup> Otros diccionarios desarrollados por el OLST e inspirados en el *DiCo* son el *DAD (Dictionnaire analytique de la distribution)*, el *DiCoEnviro (Dictionnaire fondamental de l'environnement)* y el *DiCoInfo (Dictionnaire fondamental de l'informatique et de l'Internet)*. De estos, *DiCoEnviro* y *DiCoInfo* tienen la misma interfaz de búsqueda. En este apartado describiremos el *DiCoInfo* ya que de los dos diccionarios es el que más desarrollos ha experimentado. La interfaz *Olster* permite buscar una unidad léxica simultáneamente en el *Dicouèbe*, *DAD*, *DiCoEnviro* y *DiCoInfo*.

<sup>28</sup> También existen diccionarios impresos para la búsqueda de colocaciones. Según Jousse et al. (2011) en la mayoría las colocaciones se organizan de forma sintáctica, otros añaden información semántica, que consiste principalmente en listas de sinónimos. Algunos ejemplos de diccionarios de colocaciones en papel son *BBI Dictionary of English Word Combinations*, *LTP Dictionary of Selected Collocations*, *Oxford Collocations Dictionary for Students of English*, *Dictionnaire des combinaisons de mots*, *Dictionnaire d'apprentissage du français des affaires (DAFA)*, *Dictionnaire d'apprentissage du français langue étrangère ou seconde (DAFLES)*, *Diccionario combinatorio del español Redes*.

<sup>29</sup> Véase Polguère (2000) para una descripción del diccionario *DiCo*.

<sup>30</sup> Descripción y ejemplos extraídos de Jousse y Polguère (2005).

semánticos. Por ejemplo, la descripción del verbo *coller* (acepción I.1a) sería L'individu X ~ L'objet Y À L'objet Z AVEC LA substance W.

Respecto a la interfaz de búsqueda, permite introducir consultas en cada uno de los campos o informaciones que contiene el diccionario y recuperar también solo aquellos campos o informaciones que se especifiquen. A nivel léxico puede especificarse y recuperarse por unidad léxica, número de acepción, características gramaticales y etiquetas semánticas, también puede consultarse a partir de determinadas informaciones de una función léxica, locuciones construidas con una palabra clave y ejemplos de uso. Por ejemplo, puede especificarse la función léxica Magn<sup>31</sup> y el valor *ardent*, e indicar que se recuperen solo unidades léxicas, funciones léxicas, valores de la función léxica y ejemplos. El resultado de esta consulta es una lista de unidades léxicas (*curiosité, désir, gaieté, haine* y *zèle*) cuyo valor de la función léxica Magn es *ardent*. También se incluyen contextos de uso.

El *DiCoInfo* es un diccionario en línea en francés, inglés y español<sup>32</sup> del ámbito de la informática. En el *DiCoInfo* cada entrada describe un único significado de un término, o lo que es lo mismo, cada significado de un término se trata en entradas diferentes y se distingue con números y letras (L'Homme 2008). Por ejemplo, la palabra *adresse* tiene tres entradas, *adresse*<sub>1</sub>, *adresse*<sub>2</sub> y *adresse*<sub>3</sub>. Esta es una diferencia con la mayoría de diccionarios en los que todos los significados de una palabra polisémica estarían recogidos en una misma entrada. Otra diferencia es que cada una de las entradas del *DiCoInfo* permite consultar dos o tres contextos que reflejan el significado del término. Asimismo, en este diccionario pueden buscarse colocaciones de los términos; contiene un campo de relaciones léxicas (*liens lexicaux*) en el que se incluyen relaciones paradigmáticas y también sintagmáticas (es decir, colocaciones) (L'Homme 2008).

Existen diferentes opciones de búsqueda. El modo *Term* accede al conjunto de caracteres introducido (completo o parcial) que se encuentra en el campo de entrada. La opción *Mot* busca en el campo de entrada o en el de sinónimos.

Con la función de búsqueda *lien lexical* es posible acceder a entradas que contienen en el campo de relaciones léxicas las palabras que introduzcamos. Así, puede accederse, por ejemplo a colocaciones del verbo *modifier*, como *modifier un fichier, modifier un alias*,

---

<sup>31</sup> Según Mel'čuk et al. (1995: 136): «[...] modificateur adjectival ou adverbial de la lexie-clé L qui exprime le sens de 'très', 'intense', 'à un degré élevé'».

<sup>32</sup> Las versiones inglesa y española están en proceso de elaboración.

*modifier un attribut*, etc. La opción de búsqueda *expression* permite recuperar las entradas que contienen una expresión determinada, por ejemplo *modifier un fichier*.

La última aportación a la clasificación de las colocaciones del *DiCoInfo* es su agrupación por significados, lo que según se indica tiene como objetivo facilitar el acceso onomasiológico a las colocaciones<sup>33</sup>, por ejemplo, qué verbos expresan la idea de «utilizar» un cuadro de diálogo (*utiliser une boîte de dialogue*): *activer, afficher, ouvrir une boîte de dialogue* (Jousse et al. 2011). El *DiCoInfo* no es el único diccionario de colocaciones disponible actualmente, pero sí que ha implementado modos de acceso a las colocaciones más novedosos que en otras obras del mismo tipo (L'Homme y Leroyer 2009: 260-268). Además, ha incluido un nuevo módulo (que se encuentra en versión alfa), el *Visual DiCoInfo* que genera grafos y tablas en las que quedan representadas visualmente las relaciones léxicas de una palabra de búsqueda (taxonómicas, sinónimos, contrarios, derivados, actantes y vecinos).

El *Diccionario de Colocaciones del Español (DiCE)* es un diccionario que utiliza las funciones léxicas de la Teoría Sentido-Texto para clasificar las colocaciones del español. Permite una búsqueda en la lista alfabética de entradas.

También es posible efectuar consultas avanzadas. En las consultas *directas* puede accederse a los valores que tiene una determinada función léxica y también a los valores que tiene una función léxica combinada con un lema (unidad léxica). Por ejemplo, la consulta del lema *alegría* combinado con la función léxica Magn recupera los valores *a raudales, desbordante, enorme, extraordinaria, gran*, etc. En las consultas *inversas* puede accederse a los lemas que tienen un determinado valor (y opcionalmente una función léxica). Por ejemplo, la consulta del valor *desbordante* combinado con la función léxica Magn recupera los lemas *agrado, alegría, amargura, deseo*, etc.

La opción *¿qué significa?* recupera las funciones léxicas que codifican la relación entre un lema y un valor. Por ejemplo, la consulta del lema *alegría* combinado con el valor *desbordante* recupera la función léxica Magn. Finalmente, la opción *ayuda a la redacción* ofrece dos tipos de consulta. En primer lugar, comprobar si existe relación entre un lema o base y un valor, por ejemplo la comprobación de *alegría* como lema y *desbordante* como valor tiene un resultado positivo, mientras que la comprobación de *alegría* como lema e *inquietante* como valor tiene un resultado negativo. En la segunda forma de búsqueda

---

<sup>33</sup> Función implementada actualmente solo en la versión francesa.

puede indicarse una base o lema y un *tipo* (*atributo de los participantes*, ~ + *adjetivo*, *verbo* + ~, ~ + *verbo*, *nombre de* ~) para recuperar los valores. Por ejemplo, la consulta del lema *alegría* con el tipo *verbo* + ~ recupera los valores: Caus Degrad (*alegría* 1a) = empañar, Caus Func1 (*alegría* 1a) = despertar, Caus Func1 (*alegría* 2) = alegrar, etc.

**EcoLexicon** es una base de datos en línea (Prieto Velasco y López Rodríguez 2009) elaborada a partir de los proyectos PuertoTerm (Tercedor y López Rodríguez 2008), MarcoCosta y Ecosistema, de la Universidad de Granada. Este recurso está fundamentado en la Terminología basada en marcos (Frame-Based Terminology) (Faber 2010; Faber et al. 2009; Faber et al. 2005; Faber et al. 2006), que según estos autores es la aplicación al lenguaje especializado de una versión adaptada y modificada de los marcos semánticos de Fillmore (que se han utilizado para describir el lenguaje general en el proyecto *FrameNet*, que trataremos un poco más adelante). Tal y como indican Prieto Velasco y López Rodríguez (2009) esta base de datos contiene información de muy diversa índole: información lingüística (definiciones, sinónimos, equivalentes en distintas lenguas, información sintáctica, colocaciones), información conceptual (relaciones conceptuales y estructuras de dominio), información contextual (concordancias y contextos) e información gráfica (imágenes, videos, animaciones, etc.). Según se indica, este tipo de información gráfica es una novedad muy útil que ayuda al usuario a comprender los conceptos presentes en la ontología y que no se incluye en la mayoría de bases de datos disponibles, incluso en las institucionales (como *IATE*, *FAOTERM*, *ILOTERM*, *IMF Terminology*, *UNTERM*).

**EOHS Term** es una base de datos en línea (Castagnoli 2008) desarrollada en el marco del proyecto Linguaggi e Attività Produttive de la Advanced School of Modern Languages for Interpreters and Translators (SSLMIT), de la Universidad de Bolonia, en Forlì. La función principal observada de esta base de datos es que se suprime el campo de contexto de las entradas y las palabras del diccionario están vinculadas a un corpus de referencia en el que se pueden observar los contextos. Permite una búsqueda en el índice alfabético de entradas, también a través de la estructura de dominios, en el glosario multilingüe y la búsqueda avanzada facilita la consulta en los campos de término, texto y atributo.

**FrameNet** es otro diccionario electrónico en inglés, desarrollado por el *FrameNet Project* de Berkeley. Permite al usuario acceder a las posibilidades combinatorias para cada uno de los significados de una palabra (cada significado corresponde a una unidad léxica) (Ruppenhofer et al. 2010). Este recurso está basado en la Semántica de Marcos,



*Frame Semantics* (Fillmore 1985; Fillmore y Atkins 1998; Fillmore et al. 2003). La base de datos se divide en dos partes: la base de datos de marcos y la base de datos léxica. Cada entrada de *FrameNet* representa una unidad léxica, una combinación de un lema y un marco semántico (correspondiente a un significado de una palabra). En cada entrada el usuario puede acceder a los *Frame Elements* o los elementos del marco correspondientes a una unidad léxica determinada y a los patrones sintácticos en los que aparecen estos elementos. La entrada también contiene ejemplos anotados, extraídos de un corpus, para cada posibilidad combinatoria (Fillmore et al. 2002). Existen proyectos de *FrameNets* en otros idiomas<sup>34</sup>: chino, portugués de Brasil, alemán, español, japonés y sueco.

**Genoma** es un banco de datos basado en el conocimiento sobre la temática del genoma en inglés, español y catalán, disponible en línea (Cabré 2006), desarrollado por el grupo IULATERM, del Institut Universitari de Lingüística Aplicada (IULA), dentro de los proyectos TEXTERM y RICOTERM. Integra un banco de textos sobre el área temática donde pueden buscarse contextos; un banco de documentación con información bibliográfica; un banco terminológico consistente en fichas con información sobre el término, concepto, definición, categoría gramatical, género y contextos seleccionados, así como equivalentes en los distintos idiomas disponibles; y finalmente una ontología donde puede accederse a la estructura conceptual del campo. En el banco terminológico, puede realizarse una búsqueda tanto en español, francés o catalán de un lema exacto o de un término que empiece, contenga o termine con una determinada secuencia de caracteres. Además, puede consultarse en el corpus textual disponible tanto los contextos como la frecuencia del término que se desee, o combinaciones de términos (formas o lemas) y categorías gramaticales en el corpus textual, con una interfaz de búsqueda similar a la de *Bwananet* (que analizaremos posteriormente en el apartado de análisis de corpus).

El *Grand dictionnaire terminologique*, de la *Office québécois de la langue française*, es un diccionario que contiene terminología de especialidad clasificada por temas, en francés, inglés y latín. La búsqueda *par terme* permite buscar un término exacto (simple o compuesto) en la lista de entradas, aunque también admite el uso de comodines (\* para sustituir a más de un carácter y ? para sustituir a un carácter). Por ejemplo, la introducción de *\*installation* recupera *accessoire d'installation*, *accessoires complets d'installation à bord*, *accessoires d'installation de la piste*, etc. La búsqueda *dans la*

---

<sup>34</sup> En la siguiente dirección se puede acceder a las páginas de los distintos proyectos donde se muestran avances y publicaciones: [https://framenet.icsi.berkeley.edu/fndrupal/framenets\\_in\\_other\\_languages](https://framenet.icsi.berkeley.edu/fndrupal/framenets_in_other_languages)

*définition* recupera las fichas que contienen las palabras de búsqueda en el campo de definición.

*IATE* es la base de datos multilingüe de la Unión Europea disponible en línea. Permite buscar un término o una abreviatura, seleccionar una lengua de partida y múltiples lenguas de llegada para los resultados, así como restringir la búsqueda a una determinada área temática.

*Just The Word* es, según la definen sus autores, una base de datos de conocimiento que proporciona combinaciones de palabras en lengua inglesa. El objetivo es asistir al redactor a expresar una idea con la palabra idónea. Ha sido desarrollada por Sharp Laboratories of Europe. Las combinaciones de palabras propuestas se basan en un análisis estadístico de 80.000.000 de palabras del *British National Corpus (BNC)*. Tiene dos opciones de búsqueda. *Combinations* permite introducir una palabra o una frase exacta (si se introduce más de una palabra en la casilla de búsqueda). También permite la introducción de formas flexivas. Con esta función se accede a las colocaciones más frecuentes de dicha palabra o frase exacta en el corpus. Estas colocaciones aparecen clasificadas por categoría gramatical de la palabra de búsqueda y por la relación sintáctica que guarda con el resto de elementos de la combinación. Por ejemplo, si se consulta *house*, se recupera como sustantivo y como verbo. En las combinaciones como sustantivo podemos acceder a la estructura sintáctica V obj \*house\*, donde *house* es el objeto directo del verbo, por ejemplo *leave the house*.

La otra opción de búsqueda es *Alternatives*, con la que pueden introducirse varias palabras y, como resultado, el recurso muestra la posibilidad de combinación de dichas palabras en el corpus. La barra roja indica que es una combinación poco habitual. En este caso sugiere otras combinaciones más frecuentes (la frecuencia de combinación se refleja con la longitud de la barra verde, cuanto más larga es la barra, más ocurrencias de dicha combinación en el corpus). Con una barra azul se indica la similitud entre la palabra introducida originalmente por el usuario y la sugerida por la herramienta.

Asimismo, *Just The Word* permite que el usuario seleccione la forma de desambiguación en el caso de la introducción de palabras que puedan confundirse. En *confusibles*, la opción *Thesaurus* se basa en la proximidad semántica y la opción *Learner Errors* en un corpus de errores reales de estudiantes.

*Lexical Freenet* es una base de datos léxicos en lengua inglesa o una red semántica que combina las relaciones semánticas de *WordNet* con otras relaciones de datos y fonéticas<sup>35</sup>. Esta base de datos ha sido desarrollada por Datamuse Corporation en colaboración con una serie de grupos de investigación como el grupo WordNet 1.6 de Princeton University. Una de sus funciones es la de restricción de la búsqueda por categoría gramatical (sustantivo, adjetivo, verbo y adverbio), así como por tipo de relación (sinónimo, antónimo, pronunciación, parte de, generalización, etc.).

La opción de búsqueda *show related* sirve para buscar palabras relacionadas con un término introducido. La opción *connection* se utiliza para encontrar palabras que están relacionadas al mismo tiempo con dos términos que introduzcamos. La opción *rhyme coercion* exige la introducción de dos palabras de búsqueda; las palabras recuperadas están relacionadas al mismo tiempo con los dos términos de búsqueda introducidos y, además, las palabras recuperadas tienen una similitud fonética entre sí. La selección de *spell check* permite encontrar palabras que se asemejan por la escritura al término de la búsqueda, por ejemplo, la palabra *door* recupera *door*, *dolor*, *donor*, *doors*, etc. Por último, la opción *substring* busca palabras que contengan la cadena de caracteres de la búsqueda, por ejemplo, si introducimos la palabra *door* recupera *doorbell*, *doorcase*, *doorframe*, *indoors*, *outdoorsy*, etc.

La versión en línea del *Macmillan English Dictionary* contiene un diccionario monolingüe de inglés y un tesoro, desarrollados por Macmillan Publishers Limited. Este diccionario permite una búsqueda alfabética de palabras, en la que el usuario introduce una palabra exacta que el diccionario busca en la lista alfabética de las entradas. No obstante, existen algunas funciones que ayudan al usuario a encontrar la palabra que busca. El diccionario tiene una función de autocompletar, con la que mientras que el usuario está introduciendo una palabra en la casilla de búsqueda, aparece una lista desplegable con sugerencias de palabras que se incluyen en las entradas del diccionario. Por ejemplo, si el usuario empieza a escribir *uncon*, el diccionario sugiere que se busque *unconcealed*, *unconcern*, *unconcernedly*, *unconditional*, etc. Asimismo, el diccionario recupera un listado de palabras similares a la introducida por el usuario cuando la consulta no se encuentra en el diccionario. Como resultados de las búsquedas, este diccionario muestra las entradas que se corresponden con la palabra de búsqueda. Estas entradas incluyen

---

<sup>35</sup> En el artículo de Beeferman (1998) se describe este recurso y sus funciones.

información variada, como la definición, que además de aparecer en formato escrito, incluye también archivos de audio que representan el significado de algunas palabras. Por ejemplo, se incluyen archivos de audio en entradas como la de *whisper*, que designan conceptos difíciles de representar con una imagen, pero que pueden explicarse fácilmente a través de un sonido. Las entradas también incluyen información gramatical, pronunciación, ejemplos de uso (con enlaces en cada una de las palabras del ejemplo a las entradas del diccionario de dichas palabras), y frases en las que suele aparecer la palabra de búsqueda. Finalmente, las entradas dan acceso a un tesoro, donde el usuario puede encontrar sinónimos y palabras relacionadas.

Los diccionarios en línea de *Merriam Webster*, desarrollados por la editorial Merriam Webster, incluyen distintos tipos de diccionarios. El diccionario monolingüe en inglés del lenguaje general permite buscar palabras exactas en las entradas del diccionario y contiene hipervínculos a palabras relacionadas. El tesoro ofrece sinónimos, antónimos y palabras relacionadas. Una peculiaridad que introduce Merriam Webster es un diccionario visual que recupera imágenes que representan a una palabra introducida por el usuario, y que también están clasificadas y se puede acceder a ellas a través de la búsqueda en un índice temático. En las imágenes de este diccionario se muestran detalles como las denominaciones que reciben las distintas partes de un objeto. Por ejemplo, en la imagen de una flor se pueden observar los pétalos y se indica la denominación en inglés *petal*.

*OncoTerm* es un recurso de información de terminología del cáncer desarrollado por el grupo de investigación OncoTerm (véase también grupo de investigación LexiCon, de la Universidad de Granada). Permite la búsqueda a partir de la exploración en una red de conceptos que están conectados y que reflejan las relaciones conceptuales o de significado. A partir de estas relaciones, se accede a términos en inglés y español (formas, sinónimos, siglas, etc.) y, para cada uno de los términos, se ofrece información de definiciones, información gramatical, de género y número, etc. Asimismo, para algunos términos se incluyen imágenes, enlaces a páginas web y contextos (archivos con concordancias extraídas de un corpus).

El *OneLook Reverse Dictionary* es un diccionario monolingüe inglés, en línea, desarrollado por Beeferman. Este diccionario contiene sus propias entradas y además es un metadiccionario que busca las palabras consultadas en otros diccionarios. Permite buscar una entrada a partir de palabras que la definen, a partir de la descripción de un concepto, o preguntas en lenguaje natural. Por ejemplo, al introducir *tropical fruit*, el resultado es una lista de nombres de frutas tropicales (*pineapple, mango, guava, papaya, sapodilla, etc.*).

Asimismo, recupera palabras que empiezan, terminan o contienen la secuencia de caracteres que el usuario introduce combinada con comodines (\* o ?).

**Termium Plus** es el banco de datos terminológicos y lingüísticos del Gobierno de Canadá, disponible en línea. Contiene terminología en francés, inglés, español y portugués. Pueden realizarse cuatro tipos de búsqueda. La búsqueda simple, en la cual el usuario introduce un término y la herramienta recupera las entradas exactas que se corresponden con dicho término, y ordena la lista de entradas alfabéticamente por dicho término. Por ejemplo, al introducir en búsqueda simple de términos españoles la palabra *moneda*, **Termium** recupera las fichas cuya entrada corresponde al término exacto *moneda* y ordena la lista de entradas por la palabra *moneda*. En la búsqueda avanzada en las entradas, el usuario introduce palabras de búsqueda y **Termium** recupera aquellas fichas que contengan en el campo de entrada alguna, todas o no contengan alguna de las palabras de búsqueda (pueden utilizarse los operadores Y, O, Y NO). Por ejemplo, al introducir *moneda Y curso*, **Termium** recupera las entradas *dinero de curso legal* y *moneda de curso legal*. La búsqueda avanzada en las pruebas textuales recupera las fichas que contengan las palabras introducidas por el usuario en los campos de definiciones, contextos, ejemplos de uso, observaciones y fraseología. En la búsqueda avanzada en todas las fichas, **Termium** recupera aquellas fichas que contengan las palabras introducidas por el usuario tanto en el campo de entradas como en el campo de pruebas textuales.

El diccionario **Trésor de la Langue Française informatisé (TLFi)** es un diccionario monolingüe en lengua francesa, disponible en línea. Contiene distintas pestañas de búsqueda. La *recherche d'un mot* permite buscar una palabra alfabéticamente en la lista de entradas (igual que en un diccionario en papel) y realizar una búsqueda de una palabra exacta o por aproximación ortográfica. Con la opción *correcteur d'erreurs automatique*, si el diccionario no encuentra la palabra exacta introducida, busca otras por aproximación ortográfica. Con la opción *correcteur d'erreurs forcé* el diccionario busca términos por aproximación ortográfica, incluso si ha encontrado un término exacto en el diccionario. Por último, es posible buscar por similitud fonética, donde cada letra que se introduce en la casilla de búsqueda no corresponde a una letra, sino a un fonema.

La *recherche assisté* contiene distintas posibilidades de búsqueda. En la primera casilla puede introducirse una palabra para buscar en el diccionario las entradas correspondientes a la palabra exacta introducida, o entradas que contengan dicha palabra, por ejemplo en entradas de términos compuestos. La segunda pestaña permite especificar una categoría gramatical a la búsqueda, por ejemplo que se recuperen solo adjetivos. Con

la tercera pestaña es posible seleccionar un área temática de una lista. La cuarta pestaña permite escoger de una lista el tipo de uso de las palabras que se desean recuperar, por ejemplo intransitivo, peyorativo, etc. La quinta pestaña admite buscar palabras dentro del contenido de los artículos del diccionario, por ejemplo en las definiciones, en los ejemplos, en los sinónimos, etc. Si indicamos al *TLFi* que busque adjetivos, de uso peyorativo, que contengan en la definición la palabra *animal*, se recuperan las entradas de los adjetivos *bestial* y *vache*.

La *recherche complexe* permite combinar hasta seis criterios de búsqueda que deben encontrarse en el contenido de las fichas del diccionario. Por último, es posible crear listas de palabras para utilizarlas posteriormente en las búsquedas *recherche assisté* y *recherche complexe*. El usuario puede crear una lista de palabras propias, también puede crearse automáticamente una lista de las formas flexivas de una palabra, por ejemplo, el diccionario puede generar una lista de formas flexivas del verbo *faire*.

También es posible generar automáticamente una lista de palabras que se encuentren en el diccionario a partir de la introducción de una secuencia de letras. Por ejemplo, palabras de cuatro letras que terminan por *eul*.

La página web de *Ultralingua*, desarrollada por Ultralingua Inc., incluye diccionarios monolingües y de sinónimos en inglés y francés, así como diccionarios bilingües con distintas combinaciones de idiomas en inglés, francés, español, portugués, alemán, italiano, latín y esperanto. Como algunos de los diccionarios que ya hemos descrito, la opción *Word Hunt* genera listas de palabras completas que empiezan terminan o contienen la secuencia introducida con un comodín. Permite utilizar diversos comodines, la interrogación ? (que sustituye a un solo carácter), el asterisco \* (que sustituye a cero o más caracteres) y el símbolo más + (que sustituye a uno o más caracteres). Este diccionario también recupera la forma canónica de una palabra que se corresponde con la entrada del diccionario a partir de la introducción del usuario de una forma flexiva.

*UNTERM* es una base de datos que contiene terminología multilingüe en las seis lenguas oficiales de la ONU (inglés, francés, español, ruso, chino y árabe). Permite la consulta de una o más palabras (con los operadores and, &) exactas o truncadas (con un comodín \*). Puede seleccionarse que se busquen todas las variantes de una palabra (opción: *Use word variants*) y que se recuperen también palabras similares (opción: *Find similar strings, fuzzy search*). Puede buscarse en todos los campos de las fichas a la vez o restringir la búsqueda a los campos de entrada, acrónimos o notas.

**WordNet** es una base de datos léxicos (Fellbaum 1998) o un diccionario semántico<sup>36</sup> (Miller 1998a: 7) en lengua inglesa, desarrollado en la Universidad de Princeton, y disponible en línea. *WordNet* estructura las palabras por relaciones semánticas, principalmente por sinónimos. Los términos aparecen agrupados por conjuntos de sinónimos (*sets of synonyms* o *synsets*) (Miller 1998b: 23). Cada uno de los grupos está vinculado a los demás por medio de relaciones conceptuales, semánticas y léxicas. Con este diccionario pueden buscarse palabras relacionadas semánticamente con otra palabra (sinónimos, hipónimos, hiperónimos, etc.). Esta base de datos se ha utilizado como modelo en el desarrollo de otros diccionarios como *Eurowordnet*, una base de datos en diferentes lenguas. Las relaciones de palabras en este diccionario están basadas en el *WordNet* de la Universidad de Princeton.

La página **WordReference**, desarrollada por Kellogg, incluye diccionarios monolingües y bilingües que pueden consultarse en línea. Los diccionarios multilingües ofrecen una gran variedad de combinaciones de idiomas: español, inglés, francés, portugués e italiano. Además *WordReference* contiene diccionarios monolingües y de sinónimos en inglés y español.

Una novedad respecto de otros diccionarios es que *WordReference* incorpora un foro, en el que los usuarios del diccionario pueden participar haciendo consultas o respondiéndolas. Otra aportación de este diccionario es la opción de búsqueda *in context*, que vincula el diccionario con el buscador de *Google*, y la opción *images*, que permite buscar imágenes en *Google*. Cuando se utilizan cualquiera de las dos opciones, el diccionario abre una nueva ventana donde el usuario consulta los resultados directamente en *Google*.

**Wordsmyth** es un diccionario monolingüe en inglés, desarrollado por Wordsmyth Organization. La interfaz antigua de este diccionario (todavía disponible para la consulta) permite realizar cuatro tipos de búsqueda. Como en la mayoría de diccionarios descritos, la búsqueda exacta (*exact*) recupera la palabra precisa que se está buscando. En este diccionario existe también la posibilidad de utilizar el asterisco (\*) o el signo de porcentaje (%) para sustituir a una secuencia de caracteres ilimitada, y el punto (.) o la barra baja (\_) para sustituir a un único carácter. La búsqueda ampliada (*broad*) recupera definiciones, sinónimos, ejemplos y palabras relacionadas del término introducido. Otro tipo de

---

<sup>36</sup> En la obra de Fellbaum (1998) se utilizan distintas denominaciones.

búsqueda, *spelled-like*, recupera palabras que se escriben o se pronuncian de una manera similar a la palabra introducida. La búsqueda avanzada (*advanced search*) permite recuperar las entradas del diccionario que contienen las palabras de la búsqueda dentro del campo de definición, ejemplos, sinónimos y relaciones entre palabras. Además, permite restringir por categoría gramatical. En la nueva interfaz, las opciones de búsqueda avanzada se encuentran en *Search Tools>Reverse Search*.

Asimismo, en la nueva interfaz es posible buscar en el índice alfabético o por consulta directa en el *Beginner's Dictionary*, *Children's Dictionary* y *Advanced Dictionary*. Además de la entrada que se corresponde con la consulta introducida, se recuperan palabras compuestas (*Multi-Word Results*) y similares (*Similar Spellings*). También contiene opciones de búsqueda de anagramas (*Anagram Solver*, *Crossword Solver*).

Los nuevos artículos del diccionario incluyen información sobre pronunciación, imágenes, definiciones y palabras relacionadas para cada acepción. Dentro de los artículos, la opción *Word Explorer* permite acceder a distintas relaciones de una palabra.

A partir de este análisis hemos elaborado una tabla comparativa (véase anexo 1) en la que incluimos las distintas funciones de búsqueda que hemos descrito y señalamos en qué diccionarios puede utilizarse cada una. En el eje vertical de la tabla se encuentran las funciones de búsqueda y en el eje horizontal los diccionarios analizados. Asimismo, añadimos comentarios sobre algunas funciones de búsqueda en cada diccionario.

### **3.2.4 Clasificación de técnicas de búsqueda en diccionarios electrónicos**

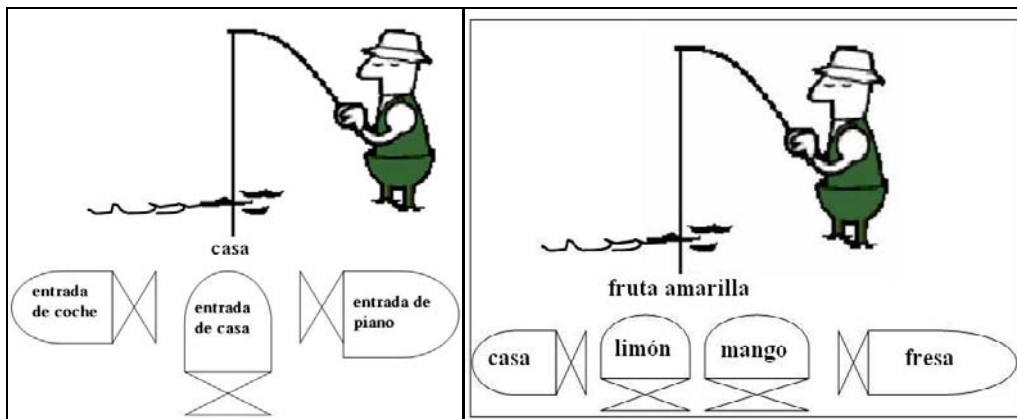
Con nuestra revisión de las técnicas de búsqueda propuestas por los distintos autores y el análisis de técnicas de búsqueda en diccionarios electrónicos, hemos visto que las búsquedas que pueden realizarse en los diccionarios son muy variadas.

En este sentido, creemos como han sugerido también otros autores que es necesario sistematizar las técnicas de búsqueda que se han implementado en los diccionarios electrónicos hasta el momento. Por ello, en este apartado proponemos una clasificación de técnicas de búsqueda en diccionarios electrónicos. Nuestra clasificación de técnicas de búsqueda está dividida en función de los tres elementos que hemos visto que intervienen en una búsqueda: la consulta, el instrumento y el resultado. La distinción entre consulta, instrumento y resultado nos permitirá reflejar de una forma clara y ordenada todas las posibilidades de búsqueda de los diccionarios.



La *consulta* es la palabra o expresión que introduce el usuario en la interfaz del diccionario. El *instrumento* es el diccionario o parte del diccionario en el que se busca la palabra o expresión de búsqueda. El *resultado* de la búsqueda es el elemento al que se accede cuando se consulta un diccionario.

En un diccionario, si utilizamos como consulta una palabra exacta, por ejemplo *casa*, y la buscamos en el campo de entradas de un diccionario (instrumento), obtendremos como resultado la ficha del diccionario correspondiente a la palabra *casa* que contiene información sobre esa palabra. Por otro lado, si introducimos como consulta una combinación de palabras, por ejemplo *fruta y amarilla* y las buscamos en el *campo de definiciones* del diccionario (instrumento), obtendremos como resultado una lista de palabras que contienen en las definiciones de sus fichas las palabras que hemos utilizado en la consulta, por tanto recuperaremos las entradas de *limón* y *mango*, pero no de *casa* y *fresa*.



**Ilustración 18. Representación de dos técnicas de búsqueda en un diccionario.**

A continuación, explicamos con más detalle las técnicas de búsqueda de nuestra clasificación y aportamos ejemplos concretos de cómo se implementan estas técnicas de búsqueda en los distintos diccionarios electrónicos que hemos analizado.

### 3.2.4.1 El primer elemento: la consulta

La consulta es la expresión que el usuario introduce en el diccionario cuando realiza una búsqueda. En la mayoría de los casos se corresponde con una palabra exacta que se busca en el diccionario. En algunos diccionarios puede introducirse una palabra truncada, o sea, un fragmento de una palabra. En otros diccionarios puede utilizarse una palabra aproximada. En ocasiones también es posible introducir un anagrama, un conjunto de

caracteres que pueden formar una palabra o no. Algunos diccionarios también permiten introducir dos o más palabras combinadas.

En combinación con la expresión de búsqueda, puede introducirse otra información con el objetivo de restringir o precisar el resultado que se desea obtener. Para ello, pueden utilizarse filtros. Un filtro condiciona que la expresión cumpla algún criterio en relación con su categoría gramatical o con el área temática a la que pertenece. Por ejemplo, un filtro de la palabra *play* podría ser la categoría gramatical sustantivo, o la palabra *pasta* podría restringirse al ámbito de la construcción.

Así, en relación con la consulta, podemos encontrar en los diccionarios diferentes tipos y subtipos de técnicas: 1) palabra exacta, 2) palabra truncada, 3) expresión aproximada (forma flexiva y similitud fonética u ortográfica), 4) anagrama, 5) combinación de dos o más palabras.

#### 3.2.4.1.1 Búsqueda mediante una palabra exacta

La búsqueda de una palabra exacta consiste en introducir en el diccionario una palabra completa que se incluye en el diccionario tal cual se ha introducido. Puede utilizarse para obtener la ficha del diccionario que contiene información sobre la palabra introducida (una definición, un ejemplo, información gramatical, etc.). Esta función la ofrecen todos los diccionarios. Por ejemplo, podemos consultar en el diccionario la palabra *casa* en la lista de entradas del diccionario y encontrar su definición, etimología, etc. En la búsqueda de una palabra exacta, pueden distinguirse o no acentos y mayúsculas o minúsculas. Normalmente la búsqueda de una palabra exacta en un diccionario deberá ser el lema de dicha palabra, ya que las entradas de los diccionarios se corresponden a los lemas de las palabras (sin morfemas flexivos).

#### 3.2.4.1.2 Búsqueda mediante una palabra truncada

Una palabra truncada es una palabra incompleta. La parte que falta en una palabra truncada puede ser el principio, el centro o el final de la palabra. Esta parte de la palabra que se omite, se sustituye con un carácter comodín. Los comodines más utilizados son el asterisco (\*) y el signo de interrogación (?).

La interrogación (?) suele sustituir a un único carácter de la búsqueda. Por ejemplo, *niñ?* recuperará ***niño*** y ***niña***.

El asterisco (\*) suele sustituir varios caracteres. Por ejemplo, *niñ\** recuperará ***niño***, ***niña***, ***niños***, ***niñera***, etc.

De los diccionarios analizados, estos dos comodines pueden utilizarse por ejemplo en el *CED*, *DUE*, *OED* y *OneLook Reverse Dictionary*. En algunos diccionarios se utilizan otros comodines. El diccionario *Ultralingua* utiliza el asterisco (\*) para reemplazar cero o más caracteres, y el signo de interrogación (?) o de suma (+) para reemplazar a uno o más caracteres. El diccionario *Wordsmyth* utiliza el asterisco (\*) o el signo de porcentaje (%) para sustituir a una secuencia de caracteres ilimitada, y el punto (.) o la barra baja (\_) para sustituir a un único carácter. El *DRAE* utiliza la interrogación (?) que sustituye a cualquier carácter, el asterisco (\*) que sustituye a cero caracteres o una secuencia de caracteres, el signo de suma (+) que sustituye a uno o más caracteres, [...] que busca cualquiera de las letras incluidas entre corchetes, [!...] que busca cualquiera letra que no esté incluida detrás de la exclamación entre corchetes, la arroba (@) que sustituye a una vocal y la almohadilla (#) que sustituye a una consonante. La opción *substring* en el diccionario *Lexical Freenet* busca palabras que contengan la cadena de caracteres de la búsqueda, por ejemplo, la palabra *reach* recupera ***reach***, ***reachable***, ***preach***, ***unreachable***, ***overreached***, etc. Asimismo, en el *Diccionari de la llengua catalana* si buscamos palabras que contengan *casa*, recuperamos 22 entradas del diccionario, algunas de ellas son *acasat*, *casa*, *casaca*, *casada*, *casador*, *casal*, etc.

Además, la mayoría de diccionarios electrónicos que hemos analizado incluyen una función de autocompletar, lo que significa que cuando el usuario empieza a introducir una palabra en el diccionario, aparece una lista desplegable en la que se sugieren entradas que se incluyen en el diccionario que empiezan con las letras que el usuario está introduciendo. Por ejemplo, en el *Macmillan English Dictionary*, si empezamos a introducir *ima*, el diccionario sugiere las palabras *image*, *imagery*, *imaginable*, *imaginary*, *imagination*, etc.

#### 3.2.4.1.3 Búsqueda mediante una expresión aproximada

Entendemos por una expresión aproximada la introducción de una palabra o una secuencia de caracteres en forma de palabra que se asemejan a alguna palabra exacta que está incluida en las entradas del diccionario. La expresión aproximada puede corresponderse con una **forma flexiva** de una palabra, o una secuencia de caracteres a modo de palabra que **suena o se escribe de una forma similar** a una palabra que se encuentra en el diccionario. Esta técnica de búsqueda puede emplearse para obtener una lista de denominaciones de palabras exactas, entradas del diccionario, que se asemejan por alguno de los criterios anteriores a la expresión aproximada que se ha introducido. A continuación describimos esta búsqueda con más detalle.

### 3.2.4.1.3.1 Búsqueda de una forma flexiva

Una forma flexiva es una palabra con morfemas flexivos, por ejemplo un sustantivo en plural o un verbo conjugado. Cuando el usuario introduce una forma flexiva en el diccionario, algunos diccionarios recuperan la forma canónica de esa palabra, que se corresponde con la entrada del diccionario. Por ejemplo, en los diccionarios *Goodrae* o *Ultralingua*, si se inserta un verbo conjugado, nos muestran el infinitivo, o si se busca un nombre o un adjetivo en plural, los diccionarios devuelven el nombre o adjetivo en singular.

Por medio de técnicas de lematización o *stemming* algunos diccionarios al introducir una palabra de consulta, buscan todas sus variantes morfológicas. Por ejemplo, en *Dirae* si introducimos *golpearon*, se accede a artículos del diccionario que contienen en el campo de entrada o de definición las variantes *golpear*, *golpe*, *golpes*. Así, se muestran los artículos *golpear*, *golpe*, *somatón*, *percutir*, *trompazo*, *zapotazo*, *cacharrazo*, etc.

### 3.2.4.1.3.2 Búsqueda de una expresión con similitud ortográfica o fonética a otra palabra

Esta técnica consiste en introducir una secuencia de caracteres, que pueden constituir una palabra o no, que se asemeja fonéticamente u ortográficamente a alguna palabra que se incluye en el diccionario. Por ejemplo, en el *DUE*, la opción *búsqueda en las entradas* busca palabras por aproximación de escritura. Como vemos en la Ilustración 19, si introducimos *perr*, el diccionario propone una lista de palabras que se asemejan por su escritura a la expresión aproximada que hemos introducido (*peor*, *pero*, *perro*, *pera*).

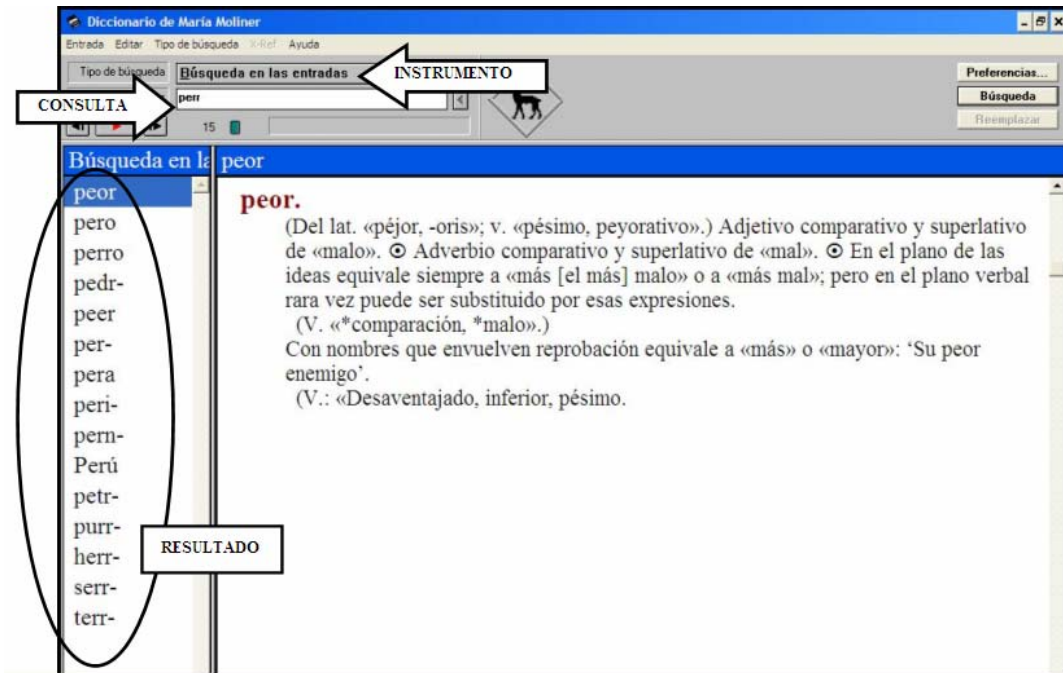


Ilustración 19. Búsqueda de una expresión con similitud ortográfica en el diccionario *DUE*.

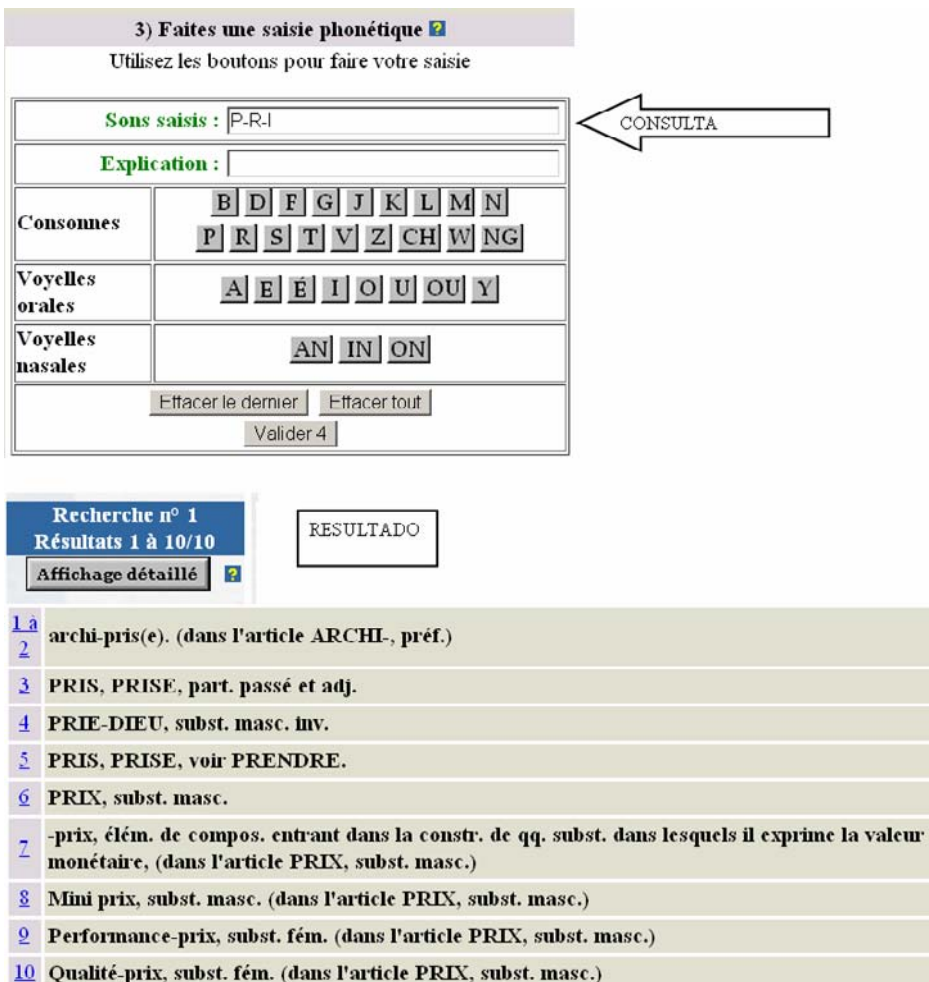


Ilustración 20. Búsqueda de una expresión con similitud fonética en el *TLFi*.

El diccionario *TLFi* permite buscar por similitud fonética, con la opción *sons saisis* cada carácter que se introduce en la casilla de búsqueda no corresponde a una letra, sino a un fonema. Por ejemplo, como observamos en la Ilustración 20, al introducir en la casilla de búsqueda fonética los fonemas P-R-I, la lista de resultados contiene las palabras *pris*, *prix*, etc.

El diccionario *Wordsmyth* incluye una opción denominada *spelled-like*, con la que también se recuperan palabras similares ortográficamente a la expresión aproximada que introduzcamos. Por ejemplo, si se introduce la expresión aproximada *setelment*, el diccionario propone las palabras *battlement*, *sediment*, *seducement*, *settlement*, *statement* y *supplement*.

El diccionario *Lexical Freenet* presenta dos opciones de búsqueda de expresiones aproximadas. La primera función, *spell check*, recupera palabras que se asemejan por escritura a la expresión introducida. Por ejemplo, la introducción de la palabra *broad* recupera *broad*, *abroad*, *byroad*, *road*, etc. La segunda opción, *sounds-like* y *rhyme*, recupera aquellas palabras que se asemejan fonéticamente a la palabra que se introduce. Por ejemplo, si introducimos la palabra *congratulations*, algunos de los resultados son: *nations*, *stations*, *creations*, *crustaceans*, *conjugations*, etc.

Además de estas opciones específicas de algunos diccionarios que buscan palabras similares fonética u ortográficamente, la mayoría de diccionarios que hemos analizado incluyen una función en la que se sugiere al usuario palabras similares a la que el usuario ha introducido en la consulta, cuando la búsqueda no ha generado resultados. Por ejemplo, en el *Macmillan English Dictionary*, si introducimos *wron*, el diccionario no muestra resultados, pero sugiere otras palabras similares que sí que se incluyen en el diccionario, como *wrong*, *wren*, *pron*, *iron*, *won*, *wrote*, etc.

#### 3.2.4.1.4 Búsqueda mediante un anagrama

Un anagrama es una secuencia de letras, que constituyen una palabra o no, cuya transposición resulta en una o más palabras completas. Los anagramas pueden utilizarse para obtener una lista de palabras del diccionario que contienen todas las letras del anagrama, en el mismo o diferente orden. Algunos diccionarios también pueden añadir o descartar de forma aleatoria un número de letras del anagrama, que el usuario especifique. Estas técnicas de búsqueda son útiles para encontrar una palabra si conocemos las letras que la componen.

Podemos realizar búsquedas de anagramas, por ejemplo en *DRAE*, *DUE*, *CED* y *Wordsmyth*. Si introducimos en el *DUE* la secuencia de letras *malo* y seleccionamos la opción de búsqueda *anagramas*, el resultado es la lista de palabras: *almo*, *amol*, *loma*, *malo*, *mola* y *olma*. Si introducimos en el *CED* la secuencia de letras *bowle* y seleccionamos la opción *anagrams*, se recupera una lista de palabras formadas con dichas letras: *bowel*, *below* y *elbow*. Además, el *DRAE* permite añadir o descartar letras a las introducidas por el usuario en la consulta. Por ejemplo, si introducimos *casa*, y seleccionamos la opción de generación de palabras añadiendo hasta tres letras más, recuperamos una lista larga de palabras que incluyen: *asca*, *casa*, *actas*, *casal*, *caseta*, *casuca*, *casería*, *caserna*, etc.

#### 3.2.4.1.5 Búsqueda mediante combinación de dos o más palabras

La búsqueda mediante dos o más palabras combinadas consiste en introducir al mismo tiempo en el diccionario dos o más palabras. Hemos visto cinco formas de combinar palabras en una consulta: 1) presencia de todas las palabras introducidas, 2) presencia de alguna de las palabras introducidas, 3) presencia de una palabra y ausencia de otra, 4) presencia de la secuencia exacta de palabras introducida, 5) introducción de palabras en forma de pregunta en lenguaje natural. Normalmente en los diccionarios se combinan las palabras utilizando operadores.

En los diccionarios que hemos analizado, el uso de la consulta mediante palabras combinadas suele utilizarse con la búsqueda en los campos de contenido del diccionario, por ejemplo para obtener una lista de denominaciones que contienen en sus definiciones las palabras que se hayan introducido como consulta.

##### 3.2.4.1.5.1 Presencia de todas las palabras introducidas

En esta técnica de búsqueda, todas las palabras introducidas en la consulta tienen que encontrarse en el instrumento de búsqueda. Los operadores utilizados para indicar que deben estar presentes todas las palabras de la búsqueda pueden cambiar de un diccionario a otro. En el *DUE* se utiliza el operador *&*, que se combina con dos palabras, una a cada lado del símbolo (*palabra1 & palabra2*). En este caso, el diccionario recupera las entradas que contienen en los campos de contenido ambas palabras de la búsqueda. Otros diccionarios que permiten la presencia de todas las palabras de búsqueda son *CED*, *Goodrae* y *Wordsmyth*. El *CED* utiliza el operador *AND*, en *Goodrae* basta con introducir más de una

palabra en la casilla de búsqueda sin necesidad de operadores y el diccionario *Wordsmyth* requiere el uso de la opción *all of the words*.

Por ejemplo, se combinan las palabras *agreement* y *lawsuit* (sin necesidad de operadores) en el *OneLook Reverse Dictionary*. El resultado es una lista de palabras, entre las que se incluyen: *settle*, *accord*, *contract*, *champerty*, etc. Todas las palabras que propone el diccionario están relacionadas con los dos términos de la consulta. En la Ilustración 21 puede observarse un extracto de la entrada de *settle* en el diccionario, en el que se indica que *settle* es un verbo que se utiliza en el contexto judicial para designar a un acuerdo entre las partes.



**Ilustración 21. Presencia de todas las palabras en *OneLook Reverse Dictionary*.**

### 3.2.4.1.5.2 Presencia de alguna de las palabras introducidas

Los operadores utilizados para indicar que debe estar presente alguna de las palabras de la consulta también difieren de un diccionario a otro. En el *DUE* se utiliza el operador |, que se combina con dos palabras, una a cada lado del símbolo (palabra1 | palabra2). En esta ocasión, el diccionario recupera las entradas que incluyen únicamente la primera o la segunda palabra de la búsqueda, o ambas al mismo tiempo. En el *CED* se emplea el operador OR. En el diccionario *Wordsmyth*, la opción *any of the words* recupera las entradas que contienen alguna de las palabras que introduzcamos.

En la siguiente imagen podemos ver un ejemplo de esta técnica de búsqueda en el diccionario *Wordsmyth*. Con la opción de búsqueda *reverse search*, se introduce en la casilla *word(s)* una consulta de tres palabras, *transport*, *carry* y *arrows*, para buscarlas en las definiciones del diccionario, lo que significa que el diccionario recuperará las denominaciones que contengan en su definición cualquiera de las palabras introducidas. El resultado es un listado de más de cien palabras. En dicha lista encontramos, por ejemplo, la palabra *quiver*, que se define como *a case designed to hold and transport arrows, often*



*strapped to the back or waist*. Esta definición contiene dos de las tres palabras de la búsqueda. También se incluyen otras palabras que solo contienen en su definición una de las palabras de la búsqueda, como por ejemplo *achieve*. Una de las definiciones de *achieve* es *to reach or carry through successfully; accomplish*. El criterio que utiliza el diccionario para ordenar la lista de palabras recuperadas es alfabético. Este diccionario no prioriza aquellas denominaciones que tienen un mayor número de palabras de búsqueda en sus definiciones.

The screenshot shows the Wordsmyth website interface. At the top, there is a navigation bar with links for 'Return to old site', 'Search Tools', 'Puzzle Solvers', 'Glossary Maker', and 'Quiz Makers'. The main content area is titled 'Reverse Search' and explains that it allows searching within the full text of dictionary entries. A search box contains the text 'transport carry arrows' and is labeled 'CONSULTA'. Below the search box, there are four radio button options: 'Word(s)', 'Word(s) + Forms', 'All Word(s)', and 'Text String'. Each option has a brief explanation and an example. To the right of these options, a box labeled 'Presencia de alguna de las palabras introducidas' points to the search results. Below the search options, there are checkboxes for 'Definition', 'Synonym', 'Similar Word', 'Example', 'Antonym', and 'Related Word'. A 'Perform Search' button is located at the bottom of the search section. Below the search section, a large arrow points down to a box labeled 'RESULTADO'. This box contains a list of search results, including 'achieve', 'quiver<sup>2</sup>', and 'sleigh'. Below the search results, there is a section for the 'Advanced Dictionary' entry for 'quiver<sup>2</sup>'. This section includes the pronunciation 'kwɪ ver', the part of speech 'noun', and two definitions. The first definition is 'a case designed to hold and transport arrows, often strapped to the back or waist. He adjusted the position of his quiver so that his arrows were in easy reach.' The second definition is 'enough arrows to fill such a case.'

Ilustración 22. Presencia de alguna de las palabras en el diccionario Wordsmyth.

### 3.2.4.1.5.3 Presencia de una palabra y ausencia de otra

Esta técnica de búsqueda consiste en que alguna palabra de las introducidas debe estar presente en el contenido de las entradas del diccionario y otras deben estar ausentes. De los diccionarios analizados, *CED*, *DUE*, *OED* y *Termium Plus* permiten esta técnica de búsqueda. El *DUE* utiliza el operador **!**, que se emplea combinado con dos palabras, una a cada lado del símbolo (palabra1 ! palabra2). Este comando recupera las entradas que presentan en su definición la primera palabra, con la condición de que no aparezca la segunda palabra. El *CED* y *OED* utilizan el operador **NOT** y *Termium Plus* el operador **AND NOT**.

En la Ilustración 23 mostramos un ejemplo de esta búsqueda en el *CED*. Combinamos las palabras *feline NOT domestic*, que se buscan en los campos de definición del diccionario. Recuperamos entradas que contienen en su definición la palabra *feline*, pero no la palabra *domestic*. La lista de resultados incluye denominaciones de felinos: *bobcat*, *caracal*, *cheetah*, *feline*, *jaguar*, *jaguarondi*, *leopard*, *lion*, *lynx*, etc. Las definiciones de estas palabras incluyen *feline* pero no *domestic*, tal y como podemos observar en la definición de *leopard*. En la lista de resultados no se ha recuperado la palabra *gato* porque se define en el diccionario como *domestic feline*.

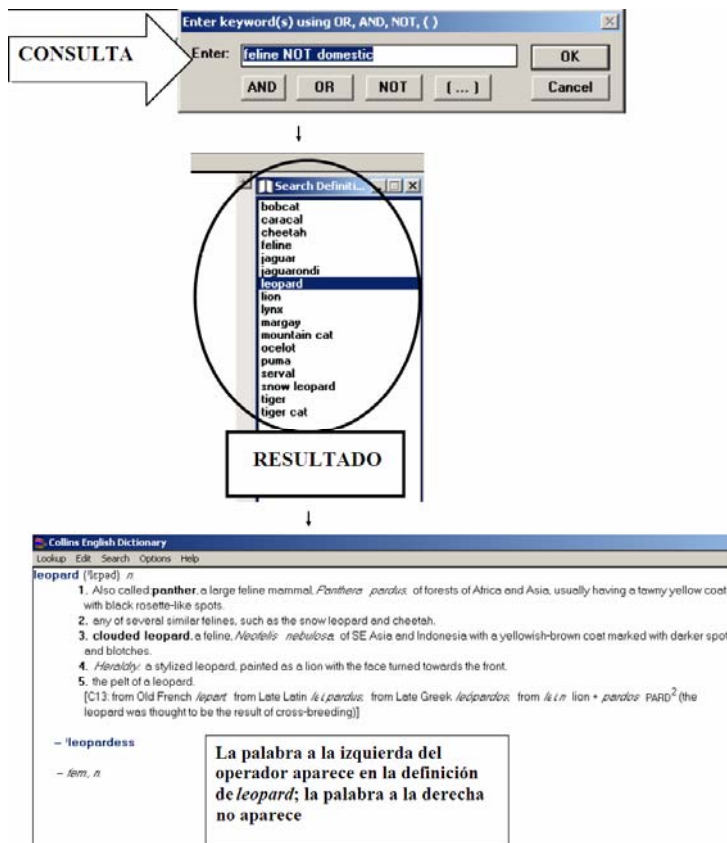


Ilustración 23. Presencia de una palabra y ausencia de otra palabra utilizando operadores en el *CED*.

Las búsquedas mediante combinación de dos o más palabras en las definiciones del diccionario pueden ser problemáticas porque requieren que el usuario adivine la palabra que aparece en la definición de la denominación que busca. Aunque la búsqueda de *feline NOT domestic* recupera una larga lista de felinos salvajes, *cat NOT house* no ofrece los mismos resultados.

#### 3.2.4.1.5.4 Presencia de la secuencia exacta de palabras introducida

De los diccionarios analizados, el *DUE*, *Dirae*, *Goodrae* y *Wordsmyth* permiten la búsqueda de una secuencia exacta de palabras en el campo de definiciones. En la mayoría de diccionarios la secuencia de palabras debe aparecer entrecomillada, aunque existen excepciones como en *Wordsmyth*, donde se introduce la secuencia en la casilla *text string*.

A continuación incluimos un ejemplo de búsqueda de una secuencia exacta de palabras en el *DUE*. Se indica al diccionario que busque palabras que contengan en su definición la expresión exacta *felino doméstico*, el resultado es la palabra *gato*. Como podemos observar en la Ilustración 24, la secuencia de palabras *felino doméstico* aparece en la definición de *gato*.

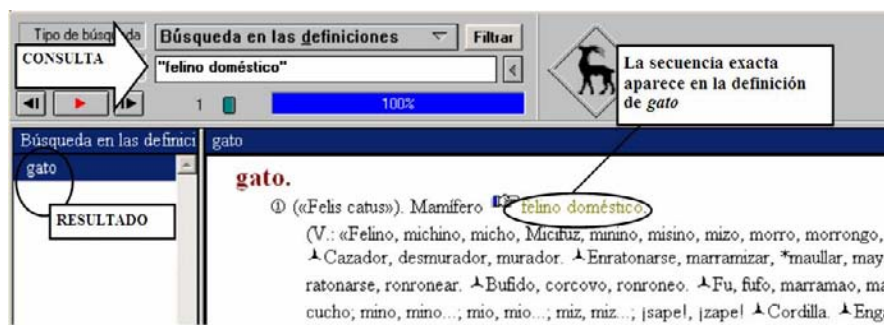


Ilustración 24. Presencia de una secuencia exacta de palabras utilizando operadores en el *DUE*.

#### 3.2.4.1.5.5 Búsqueda mediante consultas en lenguaje natural

El último subtipo de técnica de búsqueda mediante combinación de dos o más palabras es la introducción de una consulta en lenguaje natural. Esta búsqueda puede utilizarse para obtener una lista de denominaciones que pueden ser la respuesta a la pregunta introducida.

De los diccionarios que hemos analizado, *OneLook Reverse Dictionary* permite la introducción de consultas en lenguaje natural. No existe restricción con respecto al tipo de consultas que se pueden introducir. También pueden introducirse preguntas del tipo *wh-* (por ejemplo que comiencen por *what is* o *who is*).

A modo de ejemplo, si formulamos la consulta *a big building*, el diccionario propone una lista de palabras, entre las que se incluyen *hall, block, barn, tower, castle, pile, termites, court, basilica, mausoleum*, etc. La palabra *castle* está definida como *large building*. En este mismo diccionario, si formulamos la consulta *a small building*, la lista de palabras resultante incluye *chapel, turret, shed, summerhouse, lodge, shop, dentil, coach house, portakabin, cottage*, etc. La palabra *lodge* está definida como *small (rustic) house*.

### OneLook Reverse Dictionary

Describe concept  a big building

1. <a href="#">hall</a>	6. <a href="#">pile</a>
2. <a href="#">block</a>	7. <a href="#">termites</a>
3. <a href="#">barn</a>	8. <a href="#">court</a>
4. <a href="#">tower</a>	9. <a href="#">basilica</a>
5. <a href="#">castle</a>	10. <a href="#">mausoleum</a>

Quick definitions  
(*castle*)

► **noun:** a large building formerly occupied by a ruler and fortified against attack

### OneLook Reverse Dictionary

Describe concept  a small building

1. <a href="#">chapel</a>	6. <a href="#">shop</a>
2. <a href="#">turret</a>	7. <a href="#">dentil</a>
3. <a href="#">shed</a>	8. <a href="#">coach</a>
4. <a href="#">house</a>	
<a href="#">summerhouse</a>	9. <a href="#">portakabin</a>
5. <a href="#">lodge</a>	10. <a href="#">cottage</a>

Quick definitions  
(*lodge*)

► **noun:** a small (rustic) house used as a temporary shelter

**Ilustración 25. Búsqueda de consultas en lenguaje natural en el *OneLook Reverse Dictionary*.**

La búsqueda mediante consultas en lenguaje natural no deja de ser una búsqueda de palabras combinadas, ya que el diccionario extrae las palabras clave de la pregunta. El diccionario también detecta algunos elementos más complejos, como negaciones inmediatamente delante de alguna palabra clave. Por ejemplo, si introducimos en el *OneLook Reverse Dictionary* la consulta *country which has sea* el primer resultado de la búsqueda es *seaside (a place by the sea)*. Por el contrario, si introducimos *country which has no sea* el primer resultado de búsqueda es *landlocked (almost or entirely surrounded by land)*. Sin embargo, cuando la negación no aparece inmediatamente delante de la palabra clave, el sistema no extrae correctamente el significado de la pregunta. Por ejemplo, con *country that does not have sea* la respuesta que se obtiene son palabras como *short* o *solid*, que no guardan relación con la consulta introducida. Los sistemas de búsqueda todavía no permiten la introducción de preguntas más complejas.

### 3.2.4.1.6 Búsqueda mediante el historial de consultas

Como hemos visto, en la mayoría de diccionarios electrónicos, la función de historial permite al usuario acceder a consultas que ha efectuado anteriormente y volver a efectuarlas en el diccionario.

### 3.2.4.1.7 Filtros

Los filtros se emplean para añadir una condición de búsqueda a la consulta que se introduce, por ejemplo una restricción de **categoría gramatical**, **área temática** o **idioma** (en diccionarios bilingües y multilingües). Los filtros pueden combinarse con la búsqueda de una palabra exacta, una palabra truncada, una expresión aproximada o una combinación de dos o más palabras. Son útiles para restringir los resultados de la búsqueda a palabras de una categoría gramatical o un área temática concretas. Algunos diccionarios también permiten buscar dentro de los resultados de una búsqueda anterior (refinar la búsqueda). A continuación mostramos ejemplos búsqueda con filtros.

#### 3.2.4.1.7.1 Filtro de categoría gramatical

Un ejemplo de filtro de categoría gramatical sería buscar una palabra solo cuando funciona como sustantivo. El diccionario *Lexical Freenet* permite restringir la búsqueda por categoría gramatical (sustantivo, adjetivo, verbo y adverbio). Como se puede observar en la Ilustración 26, si se restringe la búsqueda de la palabra *play* a la categoría de sustantivo, el diccionario recupera únicamente la palabra *play* como sustantivo y descarta en los resultados de la búsqueda aquellas acepciones en las que *play* actúa como verbo.

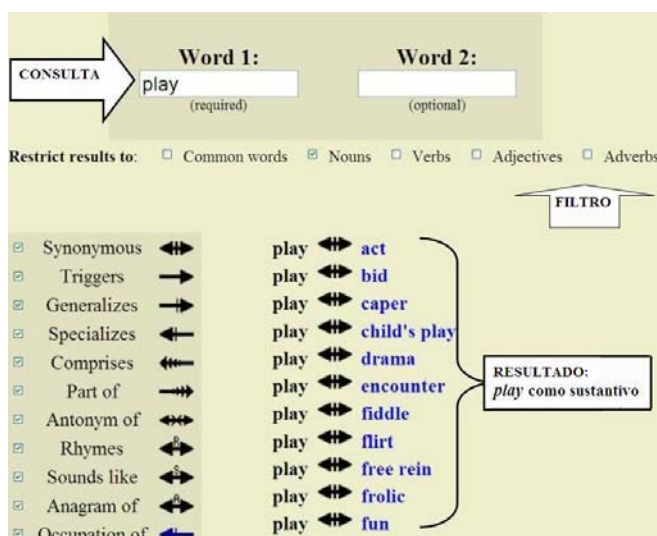


Ilustración 26. Uso de filtros de categoría gramatical en el diccionario *Lexical Freenet*.



### 3.2.4.1.7.2 Filtro de área temática

Un ejemplo de filtro de área temática podemos observarlo en la base de datos *Cercaterm*. En este diccionario introducimos la palabra *puente* y restringimos el área temática a *Electrónica*. El resultado son términos o expresiones en las que *puente* se refiere a un tipo de conexión eléctrica (*establir un pont*, *pont d'inversió*, etc.).

Si realizamos la consulta de *puente* limitando los resultados al área temática de *Transportes*, los resultados se refieren a términos o expresiones en las que *puente* es un tipo de construcción (*pont*, *aproximació a un pas a nivell o a un pont mòbil*, *pont davanter*, etc.).

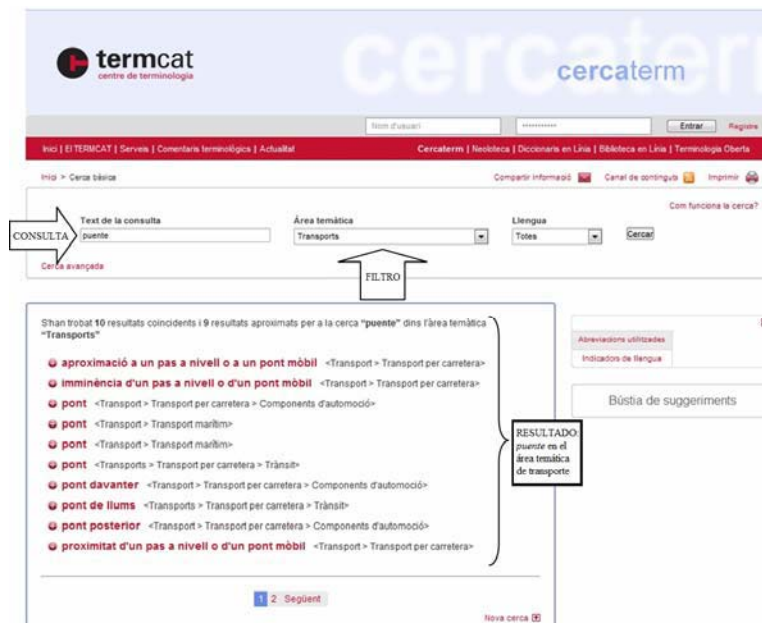


Ilustración 27. Uso de filtros de área temática en *Cercaterm*.

### 3.2.4.1.7.3 Filtro de idioma

Un ejemplo de filtro de idioma lo encontramos en la base de datos multilingüe *IATE*, donde podemos seleccionar una lengua de partida y una o más lenguas de llegada. En el ejemplo de la Ilustración 28 buscamos *drought* en inglés como lengua de partida y seleccionamos como lenguas de llegada alemán, italiano, francés y español.

The screenshot shows the IATE search interface. At the top, there's a logo for 'IATE InterActive Terminology for Europe' and a language dropdown set to 'español (es)'. Below this are links for 'Preferencias de búsqueda', 'Eliminar las preferencias de búsqueda', and 'Ayuda'. The main search area is titled 'Criterios de búsqueda' and contains a search box with 'drought' and a 'Buscar' button. A callout 'CONSULTA' points to the search box. Below the search box, there's a dropdown for 'Lengua de partida' set to 'en - Inglés'. Underneath, there's a grid of checkboxes for 'Lenguas de llegada' with 'de', 'es', 'fr', and 'it' selected. A callout 'FILTRO' points to this grid. Below the search criteria, there's a section for 'Criterios facultativos' with a dropdown for 'Elija una área temática para su consulta' set to 'Cualquier área temática' and radio buttons for 'Tipo de búsqueda' set to 'Cualquiera'. Below the search criteria, there's a box labeled 'RESULTADO:' containing the text 'Equivalentes de drought en alemán, español, francés e italiano'. Below this, there are two tables of results. The first table is for 'Ciencias de la tierra [COM]' and lists 'drought' in EN, 'dry spell', 'Trockenperiode' in DE, 'Dürre' in ES, 'sécheresse' in FR, and 'siccità' in IT. The second table is for 'AGRICULTURA, SILVICULTURA Y PESCA, Ciencias biológicas [COM]' and lists 'drought' in EN, 'Dürre' in DE, 'aridez' in ES, 'andité' in FR, and 'siccità' in IT. Each result row includes a star rating, a magnifying glass icon, and a document icon.

Ciencias de la tierra [COM]		Ficha completa	
EN	drought	★★★★	🔍
	dry spell	★★★★	🔍
	Trockenperiode	★★★★	🔍
DE	Dürreperiode	★★★★	🔍
	Dürre	★★★★	🔍
ES	sequía	★★★★	🔍
FR	sécheresse	★★★★	🔍
IT	siccità	★★★★	🔍

AGRICULTURA, SILVICULTURA Y PESCA, Ciencias biológicas [COM]		Ficha completa	
EN	drought	★★★★	🔍
DE	Dürre	★★★★	🔍
ES	aridez	★★★★	🔍
FR	andité	★★★★	🔍
IT	siccità	★★★★	🔍

Ilustración 28. Uso de filtros de idioma en *IATE*.

### 3.2.4.2 El segundo elemento: el recurso o instrumento

En un diccionario electrónico la información aparece estructurada en secciones o campos. Cada uno de esos campos sería susceptible de recibir consultas. El instrumento de la búsqueda puede corresponderse con los distintos campos o fragmentos de información que pueden aparecer en un diccionario. Esta característica de los diccionarios electrónicos difiere de los diccionarios en papel, donde la información aparece ordenada alfabéticamente y la única posibilidad es buscar por orden alfabético una palabra que se encuentre en la lista. Los diccionarios electrónicos posibilitan otros tipos de búsqueda: en las entradas del diccionario, en los campos de contenido (definiciones, ejemplos,

relaciones, foro o corpus internos), en el índice temático y en las conexiones externas (a un buscador o a otros diccionarios).

En general, el instrumento o los campos concretos que contienen la información pueden ser de cuatro tipos: 1) campo de entrada, 2) campo de contenido o texto, 3) campos de índices, y 4) campos de acceso a recursos externos.

#### 3.2.4.2.1 Búsqueda en las entradas

Las entradas de un diccionario son cada una de las palabras que encabezan las fichas del diccionario y sobre las que versa su contenido. El uso de cualquier consulta (palabra exacta, palabra truncada, etc.) en este campo permitirá acceder a las entradas que coinciden con la consulta introducida. La utilidad de la búsqueda en las entradas es obtener información sobre una palabra que se encuentra en dichas entradas, por ejemplo una definición, información gramatical, un ejemplo de uso, etc. Una entrada también puede contener subentradas, y cada subentrada contiene su propio campo de entrada. Por ejemplo, una entrada es la entrada principal (que puede contener variantes ortográficas, formas derivadas, formas compuestas y oraciones en las que suele aparecer la palabra, etc.), además de una o más subentradas que contienen expresiones idiomáticas, también *phrasal verbs* en el caso del inglés, etc.

En las entradas, se puede buscar una palabra por orden alfabético. Por ejemplo, en el *OED*, la función *word look-up* ordena la lista de entradas del diccionario por la palabra o secuencia de caracteres que introduce el usuario. Si introducimos la secuencia *hous*, la lista empieza con la palabra *housage*, que es la primera palabra de la lista de entradas del diccionario que empieza por *hous*.

En algunos diccionarios también es posible ordenar la lista de entradas de forma inversa, por la última letra introducida. Por ejemplo, en el *DRAE* si introducimos con la opción de *diccionario inverso* la secuencia de caracteres *cción*, la lista de entradas se ordena por la palabra *acción* porque es la primera palabra de la lista que alfabéticamente termina por *cción*. A esta palabra le siguen otras como *redacción*, *reacción*, *facción*, etc.



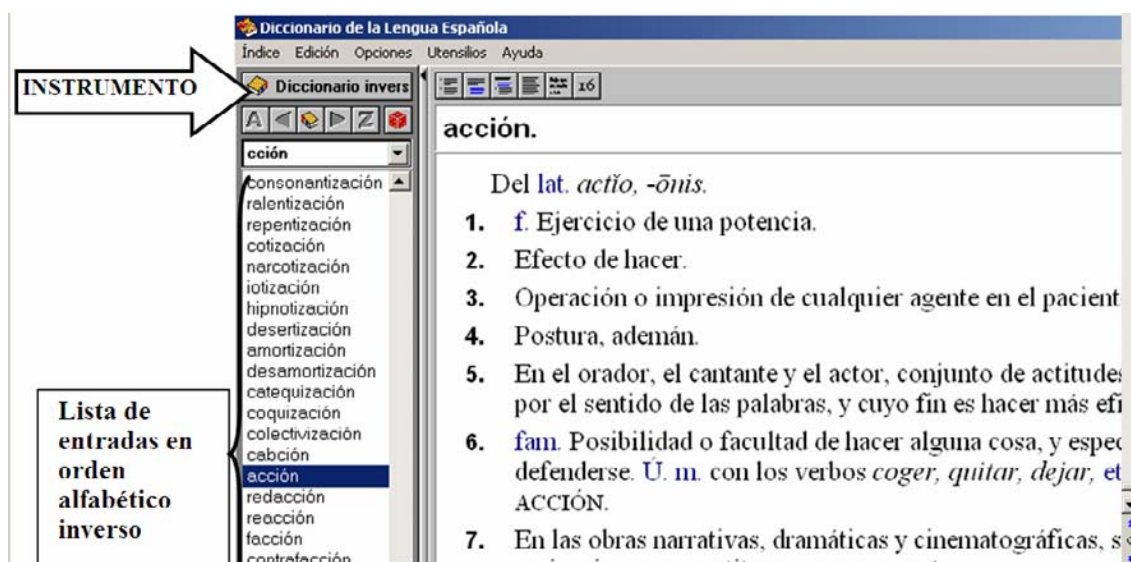


Ilustración 29. Búsqueda alfabética inversa en la lista de entradas del *DRAE*.

Algunos diccionarios permiten, dada la consulta de una palabra en el campo de entradas, recuperar tanto aquellas entradas que coinciden exactamente con la consulta introducida, como entradas de formas compuestas que contienen la palabra introducida. Por ejemplo, en *Wordsmyth* con la consulta de *house* se recupera la entrada de *house* y se sugieren también otras entradas *apartment house*, *boardinghouse*, *clean house*, etc.

### 3.2.4.2.2 Búsqueda en los campos de contenido

Los campos de contenido contienen información en formato texto respecto de cada una de las entradas. La información puede ser de distintos tipos: definición, ejemplos, relaciones léxicas, comentarios de un foro y concordancias de corpus. En estos campos de contenido se pueden introducir consultas para encontrar fichas cuyo contenido coincida con la consulta. A continuación incluimos ejemplos de búsqueda en los campos de contenido mencionados.

#### 3.2.4.2.2.1 Búsqueda en los campos de definiciones

En el ejemplo siguiente, buscamos las palabras *fruit* y *yellow* en el **campo de definiciones** del diccionario *Wordsmyth*. El diccionario recupera una lista de denominaciones *agrimony*, *apple*, *apricot*, *cherry*, *chinaberry*, *citron*, *Golden Delicious*, *grapefruit*, *Japanese quince*, *jujube* y *lemon*. En las definiciones que se encuentran a la derecha de estas denominaciones aparecen las palabras de la búsqueda *fruit* y *yellow*.

En el *Grand dictionnaire terminologique* si consultamos las palabras *instrument* *measure atmospheric pressure* en las definiciones, se recupera el término *barometer*, que

aparece definido como «Instrument used for measuring atmospheric pressure». Para que se recuperen resultados, el usuario tiene que introducir las palabras exactas que se incluyen en la definición. Por ejemplo, si introducimos *device measure air pressure* no obtendremos resultados.

*Termium* también permite la búsqueda en los campos de definiciones, contextos, ejemplos de uso, observaciones y fraseología. Por ejemplo, al buscar las palabras *fabricar Y moneda*, *Termium* recupera la ficha de *acuñar* porque la definición de la entrada es *fabricar moneda metálica*. Sin embargo *moneda Y fabricar* no ofrece ningún resultado.

The screenshot shows the Termium search interface. At the top, there are navigation links: 'Historial de la búsqueda', 'Claves de búsqueda', and '[Palabras españolas - pruebas textuales]'. The search bar contains 'fabricar Y moneda' with a search button and a 'Borrar' button. Below the search bar, there are options for 'Resultados por campo' and 'Todos los campos [1 ficha]'. The main content is a table with three columns: Inglés, Francés, and Español. The table lists search results for 'fabricar Y moneda' and 'acuñar'. The results include the word, its part of speech, and a definition in the respective language. The Spanish result for 'acuñar' is highlighted in red.

Inglés	Francés	Español
1. Subject Field(s) - Coining - Coins and Bank Notes	Domaine(s) - Monnayage - Pièces de monnaie et billets de banque	Campo(s) temático(s) - Acuñación de moneda - Monedas y billetes
coin [Fuente] CORRECTO, VERBO	frapper [Fuente] CORRECTO	acuñar [Fuente] CORRECTO
mint [Fuente] CORRECTO, VERBO	monnayer [Fuente] CORRECTO	acuñar la moneda [Fuente] CORRECTO
	monétiser [Fuente] CORRECTO	
	battre monnaie [Fuente] CORRECTO, OBSOLETO	
DEF - To strike a blank between two dies to make a coin; to use a metal to make coins. [Fuente]	DEF - Presser deux coins portant un motif sur un flan de métal pour en faire une pièce de monnaie; utiliser un ou des métaux pour en faire des pièces de monnaie. [Fuente]	DEF - <b>Fabricar moneda metálica.</b> [Fuente]
OBS - "To coin or mint money" and "to mint coins" mean to strike coins; "to coin or mint gold or silver" means to make money out of gold or silver. [Fuente]	OBS - On dit couramment «frapper des pièces de monnaie», «frapper de la monnaie». «Frapper, monnayer ou monétiser de l'or ou de l'argent» signifie utiliser de l'or ou de l'argent pour fabriquer des pièces de monnaie. «Battre monnaie» remonte à la fabrication de la monnaie par la frappe au marteau. [Fuente]	CONT - Acuñar significa batir, troquelar, imprimir por medio de cuño o troquel, es decir, darle forma de relieve por medio de presión o de un golpe, al trozo de metal que se pone entre la sufridera y el troquel. [Fuente]

Ilustración 30. Ejemplo de búsqueda en los campos de definiciones en *Termium*.

Otros diccionarios que permiten realizar búsquedas en el campo de definiciones son el *CED*, *Dirae*, *DRAE* en versión CD-ROM, *DUE*, *Goodrae*, *OED*, *Diccionari de la llengua catalana* y *OneLook Reverse Dictionary*.

### 3.2.4.2.2 Búsqueda en los campos de relaciones léxicas

Algunos diccionarios incorporan información sobre **relaciones semánticas, conceptuales o de significado** (por ejemplo relaciones de tipo paradigmático, de sinonimia, antonimia, hiperonimia, etc.) y también pueden incorporar **relaciones sintagmáticas** (por ejemplo colocaciones) entre sus palabras o términos. En muchos casos,

permiten acceder a esta información por exploración y navegación a través de una estructura de enlaces, o también por consulta directa con palabras clave.

La mayoría de diccionarios electrónicos permite una búsqueda por navegación de unas entradas a otras que se encuentran relacionadas por una estructura de hipervínculos. También es posible la exploración en una estructura de conceptos a través de un tesoro o una ontología. En la consulta por palabras clave, el diccionario busca en los campos de relaciones (por ejemplo de sinónimos) la palabra que se ha introducido como consulta y el resultado es una lista de entradas que contienen dicha palabra en sus campos de relaciones. A continuación incluimos ejemplos de estas técnicas de búsqueda.

Un ejemplo de diccionario que permite la búsqueda por navegación de términos relacionados semánticamente es *WordNet*. Este recurso permite acceder a palabras relacionadas con el término introducido en la búsqueda, por ejemplo sinónimos, hiperónimos, hipónimos, etc. Como se observa en la Ilustración 31, en la entrada de la palabra *transport*, los hipónimos que aparecen son *air transport* (si el transporte es por vía aérea) *navigation* (si el transporte es marítimo), *hauling* o *trucking* (si el transporte es en camión), etc.



**Ilustración 31. Navegación en las relaciones semánticas en *WordNet*.**

*Genoma* ofrece en las fichas del banco de datos acceso a una ontología en la que pueden observarse relaciones de hiperonimia, hiponimia, cohiponimia o de otro tipo, por ejemplo en la entrada de *genoma* se puede explorar en la ontología el concepto GENOME, donde se encuentran las relaciones de hiponimia EUKARYOTE-GENOME y PROKARYOTE-GENOME.

The image shows a screenshot of the Genoma website. At the top, there is a search bar with the text "Términos de búsqueda: genoma (Español) Condición de búsqueda: Lema exacto". Below the search bar, the results are displayed under the heading "Resultado de la Búsqueda | Términos encontrados (1)". The main result is "genoma (GENOME)". To the right of the search results, there is a section titled "Ontología" with a sub-heading "Relaciones de hiponimia". Under this section, there is a diagram showing "GENOME" at the top, with two arrows pointing down to "EUKARYOTE-GENOME" and "PROKARYOTE-GENOME". A large bracket on the right side of the diagram is labeled "RELACIONES".

**Ilustración 32. Relación de hiponimia del concepto GENOME en Genoma.**

En *OncoTerm* podemos explorar en las relaciones conceptuales de *CELL* a *BLOOD-CELL*, dentro de este concepto se incluyen los términos *blood cell* en inglés y *célula de la sangre* y el sinónimo *glóbulo sanguíneo* en español. Se incluye información gramatical y definición de cada uno de los términos. A partir de las relaciones conceptuales de *BLOOD-CELL* podemos acceder también a subclases como *ERYTHROCYTE*, *LEUKOCYTE* y *PLATELET*.

Generated by [OncoTerm](#) Updated: 23/03/2002 3:23:33

**BLOOD-CELL**

---

subject field: medicine, body-part  
 origination date: 01/06/2002  
 originator: Carlos  
 check date: 24/06/2002  
 checker: Pamela

---

Conceptual Structures	
<b>ISA</b>	CELL
<b>SUBCLASSES</b>	ERYTHROCYTE LEUKOCYTE PLATELET
<b>DESCENDANTS</b>	ACRAHULOCYTE B-CELL BASOPHIL CYTOTOXIC-T-CELL EOSINOPHIL ERYTHROCYTE GRANULOCYTE HISTIOCYTE KILLER-CELL KUFFERS-CELL LEUKOCYTE LYMPHOBLAST LYMPHOCYTE LYMPHOKINE ACTIVATED KILLER CELL MACROPHAGE MONOCYTE NATURAL-KILLER-CELL NEUTROPHIL PHAGOCYTE PLASMA-CELL PLATELET PRONYCOCYTE RETICULOCYTE T-CELL TUMOR-INFILTRATING LYMPHOCYTE
<b>ANCESTORS</b>	ALL ANIMAL-PART ANIMATE ANIMATE-PART BODY-PART CELL INTERNAL-ANIMAL-PART OBJECT PHYSICAL-OBJECT SEPARABLE ENTITY STRUCTURAL-PART

RELACIONES

**Spanish**

<b>célula de la sangre</b>	<b>definition:</b> cada uno de los distintos tipos de células que flotan en la sangre, con funciones diferentes: transporte de sangre, defensiva, coagulación, etc. (es) <b>gender:</b> feminine <b>number:</b> singular <b>part of speech:</b> noun <b>reliability code:</b> 10 <b>term type:</b> main entry term
<b>glóbulo sanguíneo</b>	<b>gender:</b> masculine <b>term type:</b> synonym <b>number:</b> singular <b>part of speech:</b> noun <b>reliability code:</b> 10 <b>definition:</b> ver célula de la sangre (es)

**English**

<b>blood cell</b>	<b>term type:</b> main entry term <b>number:</b> singular <b>part of speech:</b> noun <b>reliability code:</b> 10 <b>definition:</b> Any of three main types of cell in the blood stream: the red cell, which carries oxygen; the white cell, which fights infections; and the platelet, which helps prevent bleeding. (en)
-------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

**Ilustración 33. Relaciones conceptuales en OncoTerm.**

Un ejemplo de consulta directa de relaciones conceptuales lo encontramos en diccionarios de sinónimos. *WordReference* contiene una opción de búsqueda de sinónimos en español. Este diccionario también incluye enlaces a entradas relacionadas. El diccionario *Ultralingua* contiene diccionarios de sinónimos en inglés y francés.

Otro ejemplo de consulta por palabras clave en los campos de relaciones semánticas puede realizarse en el diccionario *Wordsmyth*. En la Ilustración 34 se busca la palabra



*agreement* en el campo de sinónimos. El diccionario muestra una lista de 27 sinónimos de *agreement*, entre los que se encuentran *accession*, *accord*, *alliance*, etc.

The screenshot shows the Wordsmyth dictionary interface. At the top, there is a search bar with the text "1. Find all words whose include(s)" and a dropdown menu set to "ALL OF THE WORDS/PHRASES". The search term "agreement" is entered. To the right of the search bar, there are two arrows: one pointing left labeled "CONSULTA" and one pointing right labeled "INSTRUMENTO". Below the search bar, a box says "RESULTADO: lista de palabras relacionadas". The results are displayed in two columns. The left column lists words with their definitions, and the right column lists words with their definitions. The words listed are: *accession*, *accord*, *alliance*, *arrangement*, *assent*, *bargain*, *compact*<sup>2</sup>, *concord*, *consent*, *contract*, *convention*, *cooperation*, *covenant*, *deal*<sup>1</sup>, *harmony*, *negotiation*, *obligation*, and *settlement*.

Word	Definition
<i>accession</i>	the act of coming into possession of an office, title, or the like. [4 more definition(s)]
<i>accord</i>	balanced interrelationship; proper proportion; harmony. [5 more definition(s)]
<i>alliance</i>	the act of allying or condition of being allied. [2 more definition(s)]
<i>arrangement</i>	the act of putting things in order. [4 more definition(s)]
<i>assent</i>	to agree or give approval; concur (usu. fol. by to). [2 more definition(s)]
<i>bargain</i>	an agreement that contains the terms of a transaction or exchange. [5 more definition(s)]
<i>compact</i> <sup>2</sup>	an agreement or contract; pact.
<i>concord</i>	a state of agreement or harmony between persons or things. [2 more definition(s)]
<i>consent</i>	to give agreement, approval, or permission (often fol. by to). [2 more definition(s)]
<i>contract</i>	an agreement, usu. legally binding, entered into by two or more parties that specifies things that should or should not be done. [7 more definition(s)]
<i>convention</i>	a formal meeting or assembly called together to discuss or pursue common interests. [3 more definition(s)]
<i>cooperation</i>	the process of cooperating. [2 more definition(s)]
<i>covenant</i>	a usu. formal agreement between two or more parties to engage in or refrain from something. [5 more definition(s)]
<i>deal</i> <sup>1</sup>	to be concerned or to handle (usu. fol. by with or in). [12 more definition(s)]
<i>harmony</i>	a state of agreement or accord; unity. [2 more definition(s)]
<i>negotiation</i>	(often pl.) mutual discussions intended to produce an agreement. [1 more definition(s)]
<i>obligation</i>	the act of binding oneself legally or morally to do or to refrain from doing something; contract. [4 more definition(s)]
<i>settlement</i>	the process or act of settling, or the condition of being settled. [3 more definition(s)]

**Ilustración 34. Búsqueda directa de relaciones semánticas en el diccionario *Wordsmyth*.**

El *OneLook Reverse Dictionary* permite introducir en el motor de búsqueda la descripción de un concepto para buscar palabras relacionadas, por ejemplo al consultar *agreement* recupera, entre otros, *contract*, *treaty*, *accord*, *covenant*, *bargain*, *assent*, *concord*, etc.

El diccionario *Lexical Freenet* permite realizar dos tipos de consulta de relaciones semánticas. En primer lugar, es posible buscar las palabras relacionadas con un término. Por ejemplo, la búsqueda de la palabra *falcon* con la opción *show related*, recupera los términos relacionados *pigeons*, *nest*, *hawk*, *hunt*, *american kestrel*, *caracara*, *falco columbarius*, *falco peregrinus*. Al lado de cada palabra, se indica el tipo de relación con el término *falcon*: relación de sinonimia, generalización, especialización, etc.

Asimismo, este diccionario permite acceder a palabras que estén relacionadas al mismo tiempo con dos palabras de búsqueda (opción *connection*). Como se puede observar en la Ilustración 35, una palabra que está relacionada al mismo tiempo con las palabras *agreement* y *lawsuit* es *settlement*.

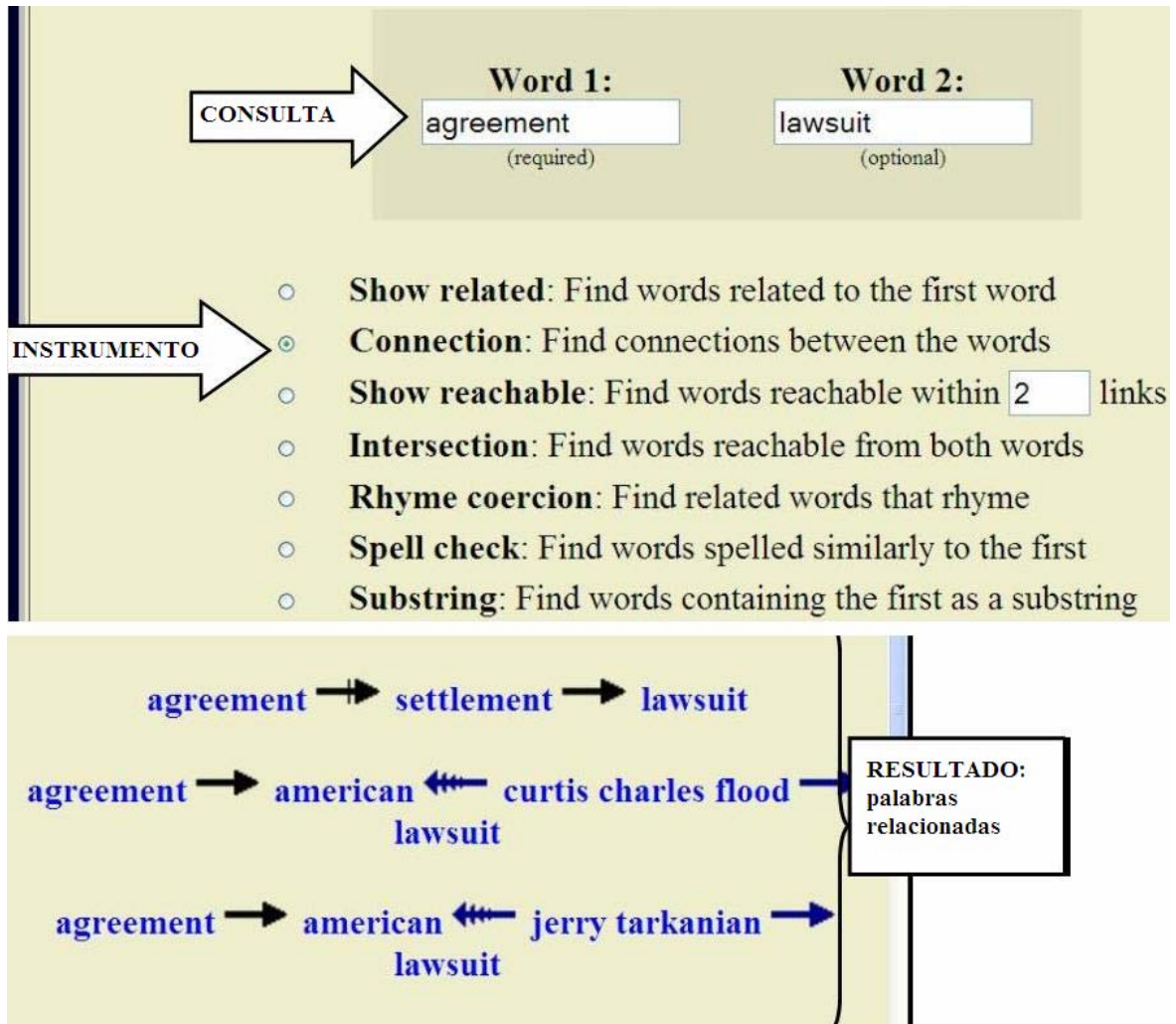


Ilustración 35. Búsqueda directa de relaciones semánticas en *Lexical Freenet*.

Otros diccionarios permiten la búsqueda de relaciones sintagmáticas, entre las que se incluyen las colocaciones de un término. Así, encontramos diccionarios de colocaciones. Por ejemplo, en la Ilustración 36, en el diccionario de informática *DiCoInfo* introducimos el verbo *envoyer* y lo buscamos en el campo *lien lexical*. Con ello, accedemos a entradas que contienen en el campo de relaciones léxicas colocaciones en francés del verbo *envoyer*. En los resultados de la búsqueda, la entrada *courriel<sub>1</sub>* contiene en el campo de relaciones léxicas la colocación *envoyer un courriel*, y la entrada *spam<sub>1</sub>* contiene la colocación *envoyer un spam à*.

**DiCoInfo**  
Le dictionnaire fondamental de l'informatique et de l'Internet

Recherche dans les fiches du DiCoInfo

Mode:   Afficher les équivalences

Langue:

Précision:

Rechercher:  **CONSULTA**

Lancer la recherche

**Résultats de la recherche trouvés dans 36 entrées.** RESULTADO

**Sens voisins**

- [rediriger<sub>1</sub>](#) : *envoyer<sub>2</sub>* (≈)
- [rediriger<sub>2</sub>](#) : *envoyer<sub>2</sub>* (≈)
- [transmettre<sub>1</sub>](#) : *envoyer<sub>2</sub>* (≈)

**Sens contraires**

- [recevoir<sub>1</sub>](#) : *envoyer<sub>2</sub>* (Conversif)

**Autres parties du discours et dérivés**

- [envoi<sub>1</sub>](#) : *envoyer<sub>2</sub>* (Verbe)

**Combinatoire**

- [adresse<sub>1</sub>](#) : *envoyer<sub>2</sub>* ... à une ~ ( L'utilisateur utilise une a. pour intervenir sur le message )
- [bavardoir<sub>1</sub>](#) : *envoyer<sub>2</sub>* ... à ... dans un ~ ( L'internaute utilise un b. pour intervenir sur des messages avec d'autres internautes )
- [canular<sub>1</sub>](#) : *envoyer<sub>2</sub>* un ~ à ... ( L'expéditeur transmet un c. au destinataire )
- (→ en: [hoax<sub>1</sub>](#) : *send a - to ...* ( The sender sends a h. to the recipient ))
- [client<sub>1</sub>](#) : *le ~ envoie<sub>1</sub>* ... à ... ( Le c. intervient sur les ressources avec le serveur )
- [client<sub>1</sub>](#) : *le ~ envoie<sub>1</sub>* ... à ... ( Le c. intervient sur la requête avec le serveur )
- (→ en: [client<sub>1</sub>](#) : *the ~ sends ... to ...* ( The c. acts on the query avec le serveur ))
- [code<sub>1</sub>](#) : *envoyer<sub>1</sub>* un ~ ( Le logiciel ou l'ordinateur fait fonctionner un c. )
- [commande<sub>1</sub>](#) : *envoyer<sub>1</sub>* une ~ à ... ( L'utilisateur prépare une c. afin de la faire fonctionner )
- (→ en: [command<sub>1</sub>](#) : *issue a -* ( The user prepares a c. to it can work ))
- [copie conforme<sub>1</sub>](#) : *envoyer<sub>2</sub>* ... à ... en ~ ( L'expéditeur transmet le message au destinataire dans la forme d'une c. )
- [corbeille<sub>1</sub>](#) : *envoyer ... à la ~* ( L'utilisateur met le fichier dans la c. )
- (→ en: [recycle bin<sub>1</sub>](#) : *place ... in the ~* ( The user places the file in the r. ))
- [courriel<sub>1</sub>](#) : *envoyer<sub>2</sub>* un ~ ( L'expéditeur transmet un c. au destinataire )
- [courriel<sub>2</sub>](#) : *envoyer<sub>2</sub>* du ~ ( L'expéditeur transmet le c. au destinataire )
- (→ en: [email<sub>1</sub>](#) : *send -* ( The sender sends e. to the recipient ))
- [courrier électronique<sub>1</sub>](#) : *envoyer<sub>2</sub>* ... par ~ ( L'utilisateur utilise le c. pour intervenir sur le message )
- (→ en: [email<sub>1</sub>](#) : *send ... by ~* ( The user uses e. to act on the message ))
- [groupe de discussion<sub>1</sub>](#) : *envoyer<sub>2</sub>* ... à un ~ ( L'internaute met le message dans un g. )
- (→ en: [newsgroup<sub>1</sub>](#) : *post ... to a ~* ( The user places the message in a n. ))
- [message<sub>2</sub>](#) : *envoyer<sub>2</sub>* un ~ à ... ( L'expéditeur transmet un m. au destinataire )
- (→ en: [message<sub>2</sub>](#) : *send a - to ...* ( The user transmits a m. to the recipient ))
- (→ en: [message<sub>2</sub>](#) : *transmit a - to ...* ( The user transmits a m. to the recipient ))
- [paquet<sub>1</sub>](#) : *envoyer<sub>1</sub>* un ~ ( L'utilisateur transmet un p. à l'ordinateur )
- (→ en: [packet<sub>1</sub>](#) : *to send a - to ...* ( The user sends a p. to the computer ))
- (→ en: [packet<sub>1</sub>](#) : *forward a -* ( The user sends a p. to the computer ))
- [pièce jointe<sub>1</sub>](#) : *envoyer<sub>2</sub>* ... en ~ ( L'expéditeur transmet le fichier au destinataire sous la forme d'une p. )
- (→ en: [attachment<sub>1</sub>](#) : *send ... as an ~* ( The sender sends the file to the recipient as an a. ))
- [pourriel<sub>1</sub>](#) : *envoyer<sub>2</sub>* un ~ à ... ( L'utilisateur transmet un p. au destinataire )
- [pourriel<sub>2</sub>](#) : *envoyer<sub>2</sub>* du ~ à ... ( L'expéditeur transmet le p. au destinataire )
- [requête<sub>2</sub>](#) : *envoyer<sub>1</sub>* une ~ à ... ( L'utilisateur ou l'ordinateur commence à utiliser une r. )
- (→ en: [request<sub>1</sub>](#) : *direct a -* ( The user, the computer or the program starts using a r. ))
- (→ en: [request<sub>1</sub>](#) : *make a -* ( The user, the computer or the program starts using a r. ))
- (→ en: [request<sub>1</sub>](#) : *send a -* ( The user, the computer or the program starts using a r. ))
- [signal<sub>1</sub>](#) : *envoyer<sub>1</sub>* un ~ à ... ( Le matériel ou le programme transmet un s. au matériel ou au programme )
- (→ en: [signal<sub>1</sub>](#) : *send a -* ( The device or the program transmits a s. to another device or program ))
- [spam<sub>1</sub>](#) : *envoyer<sub>2</sub>* un ~ à ... ( L'expéditeur transmet un s. au destinataire )
- (→ en: [spam<sub>1</sub>](#) : *send a - to ...* ( The expéditeur transmits a s. to the recipient ))
- [spam<sub>2</sub>](#) : *envoyer<sub>2</sub>* du ~ à ... ( L'expéditeur transmet du s. au destinataire )
- [spammeur<sub>1</sub>](#) : *le ~ envoie<sub>2</sub>* de ... à ... ( Le s. transmet le pourriel au destinataire )
- [téléphone<sub>1</sub>](#) : *envoyer<sub>1</sub>* une ~ ( Le commutateur transmet une t. à l'ordinateur )
- (→ en: [frame<sub>1</sub>](#) : *send a -* ( The hardware transmits a f. to the computer ))
- (→ en: [frame<sub>1</sub>](#) : *transmit a -* ( The hardware transmits a f. to the computer ))

Ilustración 36. Búsqueda de relaciones sintagmáticas (colocaciones) en el diccionario *DiCoInfo*.



El último avance del DiCoInfo ha sido organizar las colocaciones por significados, de manera que el usuario pueda acceder a las mismas de forma onomasiológica. Así, en la búsqueda del término *fichier*, la ficha recuperada contiene en el campo de *combinatoire lexicale* las colocaciones clasificadas por los siguientes significados: «créer ou supprimer», «transformer», «utiliser / ne pas utiliser», «mettre quelque part», «identifier» y «autres». Como vemos en la Ilustración 37, en el significado de «créer ou supprimer» accedemos a las colocaciones verbales de *fichier*: *créer<sub>1</sub> un ~*, *générer<sub>1</sub> un ~*.

### Recherche dans les fiches du DiCoInfo

Mode :

Langue :

Précision :

Rechercher :

Afficher les rôles actanciels

Afficher les équivalences

CONSULTA

---

**Résultats de la recherche trouvés dans 1 entrée.**

Apparaît comme entrée dans l'article :  
[fichier<sub>1</sub>](#) (↔ en: [file<sub>1</sub>](#), es: [archivo<sub>1</sub>](#))

---

**fichier<sub>1</sub>**, n. m.

Structure actancielle : un fichier : ~ créé par ([utilisateur<sub>1</sub>](#)) pour intervenir sur ([données<sub>1</sub>](#))

[Réalizations linguistiques des actants](#)

[Synonyme\(s\)](#) : fichier informatique

[Contextes](#)

[Liens lexicaux](#)

[Combinatoire lexicale](#)

[Créer ou supprimer](#)

[Créer / Faire apparaître](#)

Colocaciones clasificadas por significados

L'utilisateur crée un f	<a href="#">créer<sub>1</sub> un ~</a>
L'utilisateur crée un f. à partir de paramètres prédéfinis	<a href="#">générer<sub>1</sub> un ~</a>

[Supprimer / Détruire](#)

L'utilisateur détruit un f	<a href="#">supprimer<sub>1</sub> un ~</a>
L'utilisateur détruit un f	<a href="#">effacer<sub>1</sub> un ~</a>
Nominalisation de "L'utilisateur détruit un f "	<a href="#">suppression<sub>1</sub> d'un ~</a>
Nominalisation de "L'utilisateur détruit un f "	<a href="#">effacement<sub>1</sub> d'un ~</a>

[Transformer](#)

[Utiliser / Ne pas utiliser](#)

[Mettre quelque part](#)

[Identifier](#)

[Autres](#)

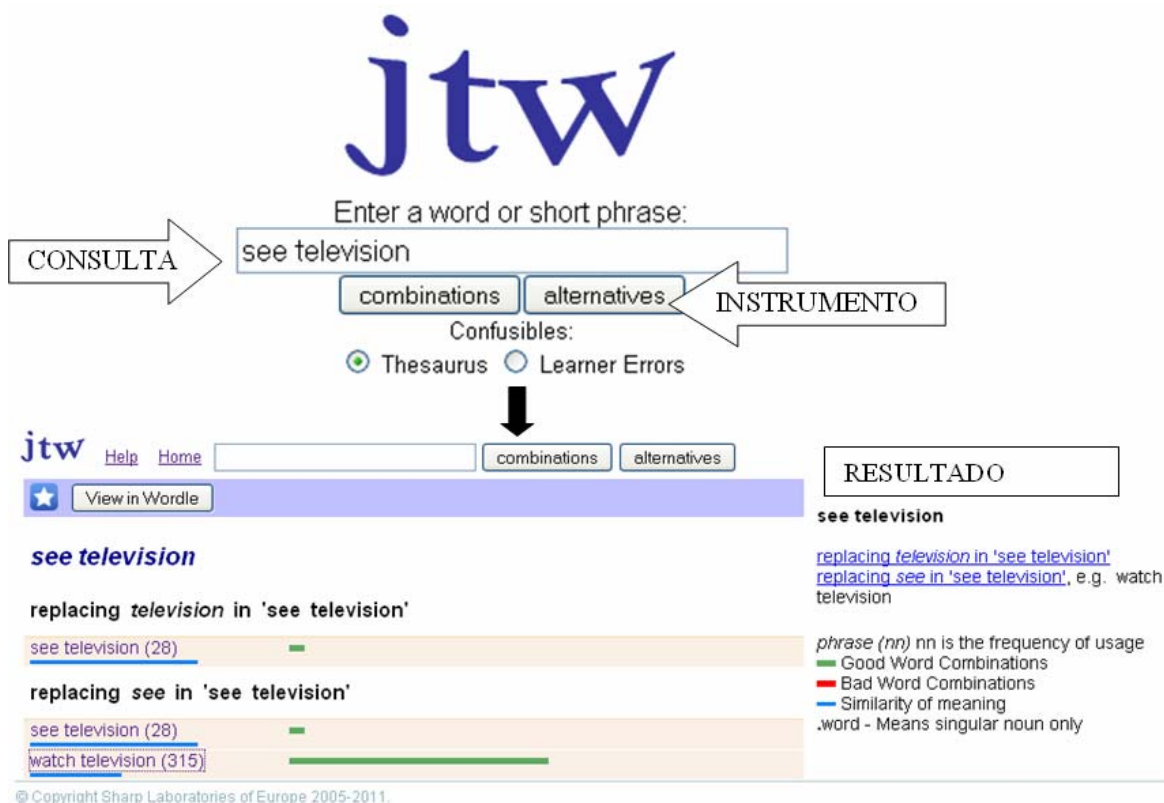
Ilustración 37. Colocaciones clasificadas por significados en el *DiCoInfo*.

En *Just The Word* si introducimos la palabra *house*, con la opción *combinations* accedemos a sus colocaciones más probables clasificadas por la categoría gramatical de la palabra de búsqueda y la relación sintáctica de las colocaciones. Por ejemplo, los primeros resultados que se muestran en la Ilustración 38 son verbos que adoptan como objeto directo el sustantivo *house* (*bring down house*, *bring from house*, *get in house*, etc.), pero también se recuperan otras combinaciones en las que *house* aparece tanto como sustantivo como verbo.

The screenshot shows the 'jtw' website interface. At the top, there is a search bar with the text 'Enter a word or short phrase:' and the word 'house' entered. To the right of the search bar is a button labeled 'CONSULTA'. Below the search bar are two buttons: 'combinations' (selected) and 'alternatives'. Below these buttons are radio buttons for 'Confusibles': 'Thesaurus' (selected) and 'Learner Errors'. An arrow labeled 'INSTRUMENTO' points to the search bar, and another arrow labeled 'CONSULTA' points to the search button. Below the search bar is a navigation bar with 'jtw', 'Help', 'Home', and buttons for 'combinations' and 'alternatives'. Below the navigation bar is a 'View in Wordle' button. The main content area is titled 'houses (Noun)'. Below this title is a section 'V obj "houses"' which contains a list of phrases with horizontal bars indicating frequency. The phrases and their frequencies are: cluster 1, bring down house (31), bring from house (12), bring into house (13), bring to house (57), carry into house (14), carry to house (14), drive to house (59), follow into house (17), follow to house (15), get in house (44), get into house (45), get to house (72), go from house (26), go in house (49), go into house (158), go through house (24), go to house (378), go round house (29), hurry to house (11), lead into house (19), and leave house (425). To the right of the list is a section 'houses (N)' with various grammatical categories and examples, and a section 'houses (V)' with various grammatical categories and examples. At the bottom right, there is a legend for 'phrase (nn) nn is the frequency of usage' and a green bar representing 'Good Word Combinations'.

**Ilustración 38. Búsqueda de relaciones sintagmáticas (colocaciones) en *Just The Word*.**

Con la opción *Alternatives*, *Just The Word* permite comprobar la idoneidad de la combinación de dos o más palabras y, en el caso de que el resultado indique que dichas palabras no se combinan habitualmente propone otras combinaciones más frecuentes. Por ejemplo, en la consulta de *see* y *television* identifica 28 resultados de esta combinación en el corpus, mientras que si reemplaza *see* por *watch*, el resultado es 315 (véase la Ilustración 39).



**Ilustración 39. Búsqueda de alternativas a relaciones sintagmáticas (colocaciones) en *Just The Word*.**

### 3.2.4.2.3 Búsqueda en un foro interno

Algunos diccionarios incorporan foros en el contenido de las fichas. En estos foros el usuario puede hacer consultas que estén relacionadas con la entrada de una ficha o consultar preguntas y respuestas de otros usuarios. Estos foros pueden ser útiles cuando la información que contiene la ficha del diccionario no satisface las dudas del usuario. Por ejemplo en el diccionario *WordReference*, las entradas *agreement* y *contract* contienen un comentario del foro «contract for agreement» en el que se explica la diferencia que existe en inglés entre estas dos palabras (Ilustración 40).

The screenshot shows a WordReference forum thread. At the top, there's a header for the word 'agreement' with links for 'In Spanish', 'In French', 'In Italian', 'In context', and 'Images'. Below that, there's a 'Listen' button with a speaker icon and 'US - UK' accents. A small text line says 'Adapted From: WordNet 2.0 Copyright 2003 by Princeton University. All rights reserved.' The main content is the dictionary entry for 'agreement', which includes a category tree (abstraction -> attribute -> quality -> characteristic -> compatibility -> agreement, correspondence) and two numbered definitions. Below the definitions, there's a section titled 'Forum discussions with the word(s) 'agreement' in the title:' followed by a list of discussion titles. A large white arrow with the word 'INSTRUMENTO' inside points to the first discussion title, 'expressions regarding agreement'. Below the list, there's a 'contract for agreement' section with a 'Hello,' message and a discussion about the difference between 'contract' and 'agreement'.

**Ilustración 40. Búsqueda en el foro a través de la entrada de *agreement* en *WordReference*.**

Los foros pueden ser útiles tanto para los usuarios como para los creadores del diccionario ya que pueden utilizar las dudas de los usuarios y apoyarse en las respuestas obtenidas para mejorar el diccionario. De esta forma, el foro se convierte en un mecanismo de control y evaluación que permite mejorar el diccionario adaptándolo a las necesidades de los usuarios. Sin embargo, esta técnica de búsqueda hay que utilizarla con cuidado ya que las respuestas que se ofrecen en un foro no están controladas por lo que se recomienda evaluar la fiabilidad de las respuestas.

#### 3.2.4.2.4 Búsqueda en un corpus interno

Algunos diccionarios incorporan en el contenido de las fichas enlaces a un corpus, en el que el usuario puede acceder a concordancias de cada entrada del diccionario. Un ejemplo de este tipo de búsqueda lo encontramos en la base de datos *EOHS Term*. En la Ilustración 41 vemos una ficha de esta base de datos. En la parte superior de la entrada *employee* aparece el botón *Concordances*. Si se presiona dicho botón se accede a 333

concordancias del término *employee*. Pueden consultarse cada una de esas concordancias, así como acceder al texto completo en el que aparecen cada uno de los contextos, presionando en el icono que hay a la derecha de cada concordancia.

Concordancias

←

INSTRUMENTO

---

**Employee**

**Domain:** Health and safety at work => Professional figures and company committees

**Grammatical category:** Noun

**Grammar notes:** Countable

**Definition:** An individual who works under a contract of employment.

**Employee's duties** It shall be the duty of every employee while at work 1) to take reasonable care for the health and safety of himself and of other persons who may be affected by his acts or omissions at work. Therefore, employees must use all work items (eg machinery, equipment, dangerous substance, transport equipment, means of production or safety device) provided by their employer correctly, in accordance with their training and the instructions they received to use them safely.

---

Results: 333

RESULTADO: concordancias

---

employer in compliance with regulation 4(1)(b)(ii) of these Regulations.	<b>employee</b>	while at work shall make full and proper use of any system of work provided for	
of work equipment shall apply to such equipment provided for use or used by an	<b>employee</b>	of his at work. (3) The requirements imposed by these Regulations on an employer	
including risks from wheels or tracks.	<b>employee</b>	is carried by mobile work equipment unless - (a) it is suitable for	

**Ilustración 41. Búsqueda en un corpus interno en la base de datos *EOHS Term*.**

*Just The Word* también permite acceder a las concordancias del corpus *BNC* del cual se han extraído las colocaciones propuestas. Pongamos como ejemplo la consulta de *see* y *television*. A partir de las concordancias deducimos que estas palabras no se combinan para expresar el significado de «ver la televisión», como muestran las siguientes concordancias: *I never saw a televisión, They should not see television as a panacea for all financial ills*. La idea de «ver la televisión» se formula con la combinación de *watch* y *television*, como puede observarse en las concordancias de la Ilustración 42: *After dinner the people watch the television [...], He'd nip out while they were watching television*.



The screenshot shows the JTW website interface. At the top left is the logo 'jtw' with links for 'Help' and 'Home'. To the right are buttons for 'combinations' and 'alternatives', and a search bar. The main content area displays a list of entries, each with a code (e.g., A06 - 000002) and a sentence where the phrase 'watch television' is highlighted in red. The entries illustrate various contexts of the phrase, such as watching TV in a restaurant, at home, or in a classroom.

**Ilustración 42. Búsqueda en un corpus interno en *Just The Word*.**

En *Genoma* a partir de la entrada de un término en el banco de datos, pueden consultarse los contextos en el corpus.

### 3.2.4.2.5 Búsqueda en el índice temático

Un índice temático es una lista de áreas ordenadas jerárquicamente que el usuario puede explorar y seleccionar el ítem que quiere consultar. Este tipo de campo se utiliza frecuentemente para mostrar un mapa de áreas temáticas. Dentro de estas áreas temáticas aparecen clasificadas las entradas del diccionario, que pueden ser palabras pero también imágenes. Existen dos tipos de búsqueda en el índice temático. En la búsqueda por exploración el usuario desciende en la estructura jerarquizada de áreas temáticas. En la consulta directa el usuario introduce en el diccionario una palabra clave que se corresponde con un área temática. A continuación incluimos ejemplos de estos dos tipos de búsqueda.

En el ejemplo de la Ilustración 43, buscamos por exploración en el índice temático del *DRAE*. Descendemos en la estructura jerárquica de *profesiones* y *disciplinas*, a continuación *ciencia* y *técnica*, luego *matemáticas* y finalmente *álgebra*. El resultado es una lista de entradas del diccionario que están clasificadas dentro del área temática seleccionada, entre las que se incluyen *binomio*, *cociente*, *coeficiente*, *combinación*, etc.

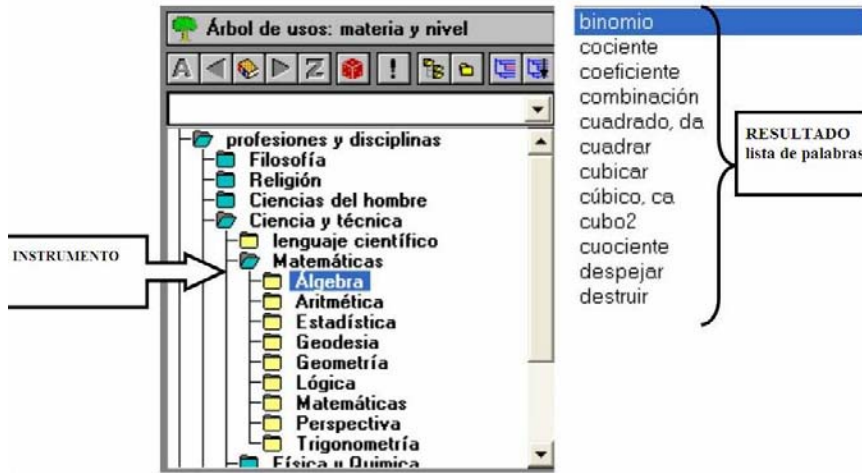


Ilustración 43. Búsqueda por exploración en el índice temático del *DRAE*.

También podemos buscar por exploración en la estructura de dominios de *EcoLexicon*, en función de si el concepto que buscamos puede clasificarse en la ontología como un *agente*, un *paciente*, un *proceso* o una *descripción*, y dentro de estos campos de clasificación encontramos subcampos. Por ejemplo, podemos acceder a la estructura *agente*, después *agente natural*, a continuación *agente biológico* y finalmente *plancton*. Además, se abre una red de conceptos, en la que unos conceptos se encuentran relacionados con otros.

En el ejemplo de la Ilustración 44 realizamos una consulta por palabras clave en el índice temático del diccionario visual de *Merriam Webster*. Introducimos la palabra clave *flower*. El diccionario propone como áreas temáticas *flower* y *flowering*. Si seleccionamos *flower* aparecen entradas de imágenes relacionadas con este ámbito como *pleasure garden*, *examples of flowers*, *structure of a flower*, *structure of a plant*, etc. En este diccionario también es posible acceder por exploración a través del índice temático completo.

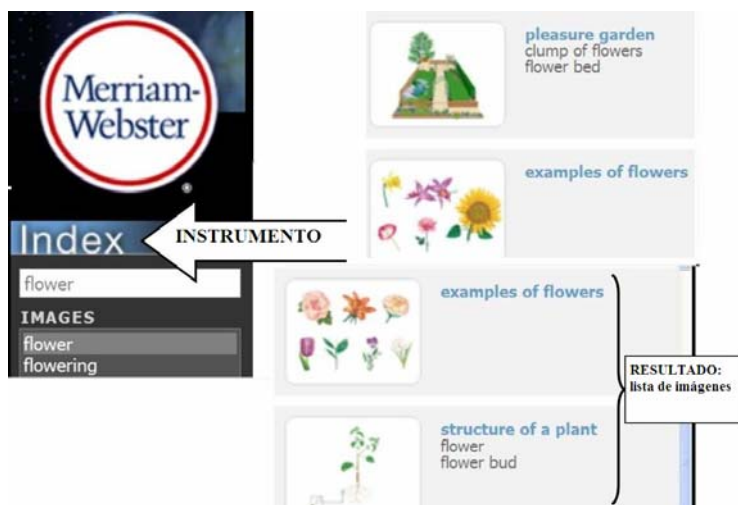


Ilustración 44. Consulta directa con palabras clave en el índice del diccionario visual *Merriam Webster*.

### 3.2.4.2.6 Búsqueda en conexiones externas

Este tipo de campo ofrece enlaces a recursos externos al propio diccionario. Algunos diccionarios ofrecen acceso a buscadores de Internet u otros diccionarios.

Por ejemplo, en el diccionario *WordReference*, la opción *in context* busca la consulta introducida en el buscador de *Google*. Asimismo, la opción *images* busca la consulta introducida en las imágenes de *Google*.

Un ejemplo de búsqueda en diccionarios externos la encontramos en el diccionario *OneLook Reverse Dictionary*. En la entrada de la palabra *hall* encontramos enlaces a otros diccionarios en línea de carácter general, como *American Heritage Dictionary of the English Language* o *Cambridge International Dictionary of English*, así como otros diccionarios específicos, como por ejemplo de arte, economía, informática, medicina, etc.

We found 43 dictionaries with English definitions that include the word *hall*:  
Click on the first link on a line below to go directly to a page where "hall" is defined.

➔ **General** (29 matching dictionaries)

1. [hall](#): American Heritage Dictionary of the English Language [[home](#), [info](#)]
  2. [hall](#): Encarta® World English Dictionary, North American Edition [[home](#), [info](#)]
  3. [hall](#): Merriam-Webster's Online Dictionary, 11th Edition [[home](#), [info](#)]
  4. [hall](#): Cambridge International Dictionary of English [[home](#), [info](#)]
- [...]

ENLACES  
EXTERNOS

➔ **Art** (1 matching dictionary)

30. HALL: Shakespeare Glossary [[home](#), [info](#)]

➔ **Business** (2 matching dictionaries)

31. HALL: Bouvier's Law Dictionary 1856 Edition [[home](#), [info](#)]
32. [hall](#): Legal dictionary [[home](#), [info](#)]

**Ilustración 45. Búsqueda en conexiones externas a otros diccionarios en línea en el *OneLook Reverse Dictionary*.**

### 3.2.4.3 El tercer elemento: el resultado

El resultado de la búsqueda es la información que se obtiene cuando se realiza una consulta en el diccionario. Normalmente, el resultado de la búsqueda en un diccionario es una ficha con información completa sobre una palabra (significado, información gramatical, información etimológica, información sobre uso en contexto, información sobre equivalencias en otras lenguas, colocaciones y palabras relacionadas, etc.). En otras



ocasiones el resultado es una representación gráfica de las relaciones semánticas y también sintagmáticas de una palabra que puede ser a modo de grafo, tabla o de otro tipo. También es posible obtener como resultado una denominación o un listado de denominaciones que se corresponden con entradas del diccionario. Por último, también es posible que el resultado sea una imagen o un conjunto de imágenes, y los resultados también pueden incluir archivos de audio. La generación de estos resultados depende del tipo de diccionario y de las funciones que tiene incorporadas. A continuación vamos a explicar los resultados más novedosos desde el punto de vista de la implementación en diccionarios electrónicos.

#### 3.2.4.3.1 Contextos

En primer lugar explicamos el resultado en el que se obtiene **información sobre el uso en contexto** de una palabra. Algunos diccionarios incluyen en sus fichas información contextual. El diccionario *DiCoInfo* contiene un campo contextual en el que aparece una muestra de concordancias extraídas del corpus utilizado para crear el diccionario. Por ejemplo, en la entrada de *backup* aparecen tres contextos con sus referencias al fichero del corpus del que se han obtenido.

*EcoLexicon* también prevé una ventana de resultados donde pueden incluirse contextos, concordancias extraídas de un corpus, fraseología, notas, etc.

Las entradas en *FrameNet* igualmente incluyen información contextual. En la Ilustración 46 mostramos la entrada del verbo *play* en el marco *performers\_and\_roles*. Todos los elementos del marco que se han anotado en los contextos de *play* se incluyen en una tabla. A cada elemento del marco se le asigna un color, que permite al usuario identificar cada elemento del marco en un contexto extraído de un corpus.

Asimismo es posible obtener información contextual de las entradas del diccionario en enlaces a un corpus interno del diccionario, como hemos visto en *EOHS Term*.

En *Just The Word* las colocaciones se extraen del *BNC*, por lo que haciendo clic en cada una de ellas es posible acceder a los contextos del corpus de los que se han extraído.

## Lexical Entry

[Annotation](#) [Performers\\_and\\_roles](#)

### play.v

Frame: Performers\_and\_roles

Definition:

COD: represent (a character) in a play or film

#### Frame Elements and Their Syntactic Realizations

The Frame Elements for this word sense are (with realizations):

Frame Element	Number Annotated	Realization(s)
Audience	(2)	INI-- (2)
Event_description	(2)	PP[like].Dep (1) VPing Dep (1)
Manner	(4)	AVP Dep (3) PP[like].Dep (1)
Medium	(2)	INI-- (2)
Performance	(16)	INI-- (2) PP[in].Dep (12) NP Obj (1)
Performer	(36)	NP.Ext (31) PP[by].Dep (5)
Performer1	(1)	NP.Ext (1)
Performer2	(1)	PP[with].Dep (1)
Place	(3)	PP[for].Dep (1) PP[at].Dep (2)
Role	(36)	NP Obj (29) PP[as].Dep (1) NP.Ext (5) PP[by].Dep (1)
Score	(1)	INI-- (1)
Script	(1)	INI-- (1)
Time	(6)	PP[before].Dep (1) AVP Dep (3) PP[in].Dep (1) PP[at].Dep (1)
Type	(1)	INI-- (1)

[Clear Sentences](#) [Turn Colors Off](#)

- [X] While this was going on Marcus took off his jacket and his tie and handed them to Irina, who was now visibly PLAYING the part of the magician's assistant.
- [X] The part of Joy Davidman is very well PLAYED by Kay Gallo, though it is difficult to take to the character in the first act of the play.
- [X] Lucy Alcorn as Mrs Montgomery was beautifully PLAYED, with telling scenes.
- [X] He would PLAY all the characters like a one-man theatre.

Información de uso en contexto

Ilustración 46. Resultados de uso en contexto en *FrameNet*.

### 3.2.4.3.2 Colocaciones y palabras relacionadas

El resultado de la búsqueda puede ser una ficha del diccionario con información sobre palabras relacionadas paradigmáticamente (por ejemplo, sinónimos, antónimos, etc.) y palabras relacionadas sintagmáticamente (colocaciones). En la versión francesa del *DiCoInfo*, cuando una unidad léxica es la base de colocaciones que expresan diferentes significados, estas colocaciones aparecen agrupadas en función del significado. Por ejemplo, la unidad léxica *fichier* es la base de dos colocaciones que expresan la idea de «suprimir un archivo»: *supprimer* o *effacer un fichier*, por eso estas dos colocaciones aparecen juntas. En el *DiCE* la opción *ayuda a la redacción* permite al usuario recuperar las colocaciones que expresan la idea de «amor verdadero»: *amor acendrado*, *verdadero*, *único*.

### 3.2.4.3.3 Representación gráfica de las relaciones de una palabra

En algunos diccionarios los resultados de las relaciones semánticas y sintagmáticas de una palabra aparecen representados de forma gráfica por medio de grafos o tablas. Un grafo es una estructura de datos consistente en un conjunto de nodos y un conjunto de arcos que establecen las relaciones entre los nodos. En el caso del *Visual DiCoInfo*, si consultamos el verbo *cliquer* podemos acceder a la representación gráfica de las relaciones sinonímicas (*double cliquer* y *double-cliquer*), cuasiantonímicas (*relâcher*), derivados (*cliquable*), actantes (*fichier, utilisateur*) y vecinos (*activer, enfoncer, appuyer*). Haciendo clic en cada uno de los nodos del grafo puede accederse a un nuevo conjunto de relaciones de la palabra seleccionada.

También en *EcoLexicon* las palabras consultadas aparecen representadas a modo de grafo, como por ejemplo *sequía* «atributo de» *sequedad*, «tipo de» *sequía hidrológica*, *fenómeno atmosférico*, «afecta a» *temperatura*, *lluvia*, *terreno*, *lago temporal*, *caudal*, «resultado de» *estiaje*, y equivalentes en diferentes idiomas. Como en el anterior diccionario, haciendo clic en uno de los nodos del grafo, accedemos a una nueva representación de la palabra que seleccionemos (véase la Ilustración 47).



Ilustración 47. Representación gráfica de las relaciones de *sequía* en *EcoLexicon*.

#### 3.2.4.3.4 Denominación o lista de denominaciones

Otro resultado que puede obtenerse cuando se consulta un diccionario es una **denominación o lista de denominaciones**. Algunas técnicas de búsqueda, como la búsqueda en campos de contenido o la búsqueda en el índice temático pueden generar un listado de denominaciones de entradas como resultado. Por ejemplo, si buscamos en el campo de definiciones del diccionario *Wordsmyth* la palabra *vehicle*, obtenemos una lista de denominaciones que contienen dicha palabra en su definición, como *aircraft*, *airflow*, *ambulance*, *aquaplane*, *ATV*, *automobile*, etc.

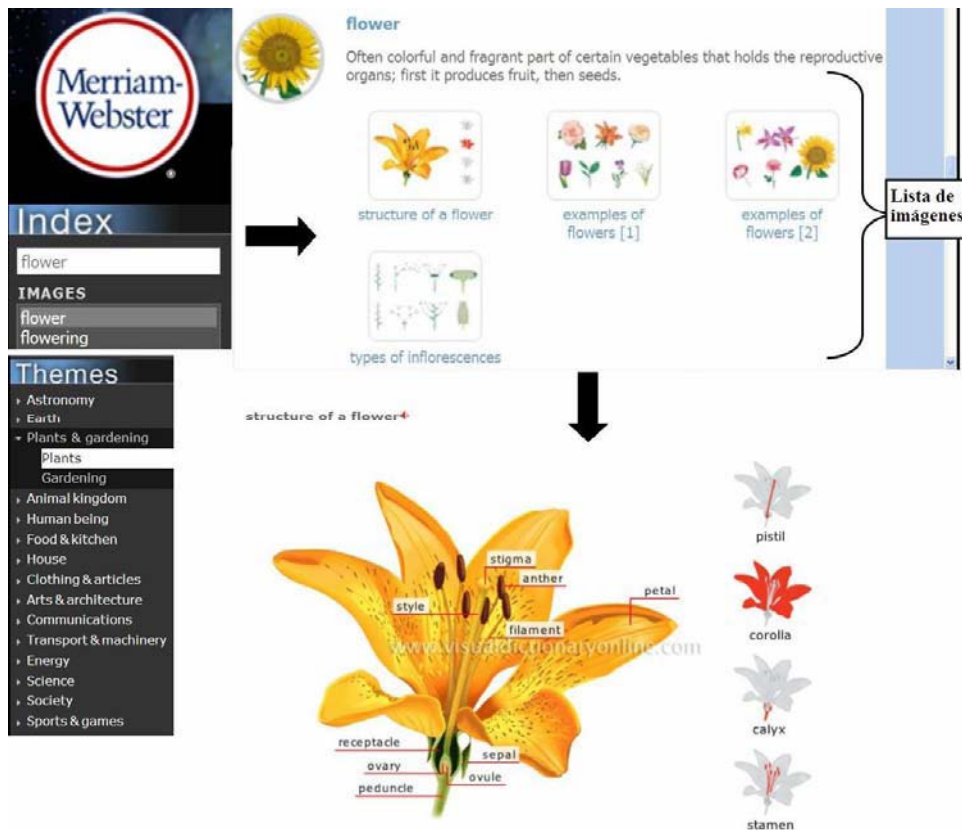
Asimismo, si buscamos en el índice temático del *DRAE* el área temática *Lógica* dentro de *Filosofía* y, a su vez, dentro de *profesiones y disciplinas*, el resultado es un listado de denominaciones que están clasificadas en este ámbito, como *a contrariis*, *ad hómīnem*, *antecedente*, *a pari*, *apodíctico*, *argumento*, etc.

En la mayoría de diccionarios analizados, los listados de denominaciones se muestran en orden alfabético. Una excepción es *Dirae*, en el que los resultados pueden ordenarse *alfabéticamente*, por *relevancia* en función de la búsqueda del usuario, por orden de *frecuencia* en un corpus, o por *longitud* de las palabras.

Sería útil que los resultados se presentaran al usuario en orden de relevancia en función de la consulta. Por ejemplo, si el usuario ha introducido *agency*, la lista resultante de denominaciones no debería visualizarse en el orden en el que aparecen en el diccionario, sino por orden de relevancia, de manera que *agency* debería aparecer en la lista delante de *adoption agency*.

#### 3.2.4.3.5 Imagen o lista de imágenes

Algunos diccionarios visuales recuperan imágenes que suelen estar clasificadas en un índice temático. Por ejemplo, como en la Ilustración 48, en el diccionario visual de *Merriam Webster*, si accedemos en el índice (*themes*) a *Plants & gardening*, luego *Plants*, después *Flower* y finalmente *Structure of a flower*, encontramos una imagen de una flor con el nombre de todas sus partes.



**Ilustración 48.** Resultado de una imagen en el diccionario visual de *Merriam Webster*.

En *EcoLexicon* se incorporan imágenes que complementan la información aportada en las definiciones. Por ejemplo, en la entrada de *desertificación* se incluye una imagen que ilustra el significado de *desertificación*, y además se incluye un enlace a la página web de la que se ha obtenido la imagen donde hay disponible más información sobre el concepto de *desertificación*.

### 3.2.4.3.6 Archivos de audio

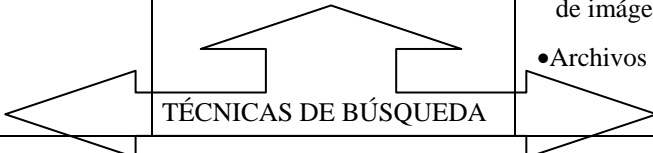
Cada vez es más frecuente que los diccionarios incluyan archivos de audio en las entradas de algunos términos. En muchos diccionarios estos archivos de audio proporcionan la pronunciación de palabras. Otros archivos de audio pueden ayudar al usuario a comprender el significado de una palabra, de la misma forma que una definición o una imagen. Por ejemplo, en el *Macmillan English Dictionary* algunas entradas incluyen archivos de audio que clarifican el significado de las palabras por medio de un sonido. En la entrada de la palabra *applause*, podemos escuchar el sonido de un aplauso, por lo que se facilita al usuario la comprensión de la palabra.

### 3.2.5 Recapitulación de nuestra propuesta de clasificación de técnicas de búsqueda en diccionarios electrónicos

En los apartados anteriores, a partir de las aportaciones de otros estudios, por un lado, y del análisis de una serie de diccionarios electrónicos, por otro, hemos sintetizado las técnicas de búsqueda en diccionarios electrónicos en función de tres elementos: la consulta, el instrumento y el resultado. La combinación de estos tres elementos nos ha permitido representar todas las posibilidades de búsqueda que hemos observado en los diccionarios electrónicos, y que hemos resumido en la Tabla 14.

Una clasificación de este tipo puede ayudar, gracias a la unificación de nomenclaturas, a resolver los problemas que se han detectado, como la enseñanza del uso de los diccionarios a los usuarios y el diseño de nuevos diccionarios. Además, está preparada para la inclusión de nuevos elementos en función de los avances que puedan producirse en el futuro.

CONSULTA	RECURSO O INSTRUMENTO	RESULTADO
<ul style="list-style-type: none"> <li>•Una palabra exacta (lema: perro)</li> <li>•Una palabra truncada (perr*)</li> <li>•Una palabra aproximada (forma flexiva: <i>played</i>; secuencia de caracteres a modo de palabra: <i>setelment</i>)</li> <li>•Un anagrama (<i>lamo</i>)</li> <li>•Una combinación de dos o más palabras (<i>fruta, amarillo y ácido</i>)</li> <li>•Historial de consultas</li> <li>•Filtros               <ul style="list-style-type: none"> <li>-Categoría gramatical (sustantivo, verbo, adjetivo, etc.)</li> <li>-Área temática (electrónica, transportes)</li> <li>-Idioma (español, francés)</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>•Las entradas (índice alfabético)</li> <li>•Los campos de contenido (definiciones, ejemplos, relaciones, foros internos, corpus internos, etc.)</li> <li>•El índice temático</li> <li>•Las conexiones externas</li> </ul>	<ul style="list-style-type: none"> <li>•Información sobre el significado</li> <li>•Información gramatical</li> <li>•Información etimológica</li> <li>•Información sobre equivalencias</li> <li>•Información sobre uso en contexto</li> <li>•Información sobre colocaciones y palabras relacionadas</li> <li>•Representación gráfica de las relaciones de una palabra</li> <li>•Una denominación o un listado de denominaciones</li> <li>•Una imagen o un conjunto de imágenes.</li> <li>•Archivos de sonido</li> </ul>



TÉCNICAS DE BÚSQUEDA

Tabla 14. Resumen de la clasificación de técnicas de búsqueda en diccionarios electrónicos.

### **3.3 Corpus**

Un corpus puede definirse como un «conjunto suficientemente grande de datos reales de la lengua (escrita o hablada) que se quiere investigar y está en un formato procesable por ordenador. El conjunto de textos del corpus debe ser representativo de la lengua o variedad de lengua que se quiere analizar, ya sea un dialecto o variedad lingüística profesional, para que los resultados del estudio puedan extrapolarse a dicha variedad de lengua» (Sánchez-Gijón 2003a: 188). Otra definición de corpus la aporta Corpas Pastor (2004: 226) que afirma que «se entiende por corpus una colección más o menos extensa de (fragmentos de) textos en soporte electrónico, cuya principal característica es la representatividad del estado o sector de la lengua seleccionado».

En esta parte de nuestro estudio hemos analizado las técnicas de búsqueda en corpus. Para ello, hemos examinado corpus disponibles en línea que incorporan sus propias interfaces de consulta, programas de análisis de corpus que permiten cargar un corpus proporcionado por el usuario o descargar un corpus de Internet y, por último, programas de traducción asistida para la consulta de memorias de traducción. A partir de este estudio, hemos elaborado una clasificación de técnicas de búsqueda en corpus con los mismos criterios de categorización que la clasificación de técnicas de búsqueda en diccionarios.

En los siguientes subapartados nos referimos, en primer lugar, a la utilidad de los corpus como recurso de búsqueda (apartado 3.3.1). A continuación, en el apartado 3.3.2, efectuamos una revisión de los tipos de corpus disponibles, de las ventajas y desventajas que presentan y sus posibilidades de búsqueda. En el apartado 3.3.3 procedemos a analizar un conjunto de corpus para observar los tipos de consulta que se han implementado en ellos. A partir de este análisis empírico, diseñamos una clasificación de técnicas de búsqueda en corpus, que presentamos en el apartado 3.3.4. Concluimos en el apartado 3.3.5 con el resumen y algunas reflexiones de nuestra clasificación de técnicas de búsqueda en corpus.

#### **3.3.1 Utilidad de los corpus como recurso de búsqueda terminológica, según la bibliografía**

Diversos estudios han revelado que los diccionarios no resuelven todas las dudas terminológicas de los traductores. Los diccionarios muestran carencias por la falta de información que presentan, por la escasa actualización de sus contenidos y por sus limitadas funciones para acceder a esos contenidos (Bowker 1998b: 634; Bowker y

Pearson 2002: 15-16; Fraser 1999: 26; Nesi 1999: 58, entre otros). Gallardo San Salvador y de Irazazábal (2002: 194) afirman que «[...] la terminología que el traductor necesita debería contener, además de los términos equivalentes en otras lenguas, los contextos que le faciliten información sobre cómo y dónde utilizar ese término y datos sobre el concepto que expresa, a fin de que pueda garantizar que está utilizando la forma precisa que corresponde a un determinado contenido». Todo ello fomenta que los traductores utilicen cada vez más recursos como Internet y los corpus para acceder a la terminología que necesitan.

Actualmente parece indiscutible que la elaboración de recursos léxicos se apoya en el uso de corpus para la extracción de los datos lingüísticos (Ahmad y Rogers 1997; Bowker y Pearson 2002; Cabré 1993; Meyer y Mackintosh 1996; Vargas Sierra 2006). Asimismo, cada vez son más las voces que abogan por incluir herramientas de análisis de corpus como parte de los diccionarios electrónicos (De Schryver 2003: 167). Esto mismo se ha sugerido en trabajos como los de Atkins (1996), Dickens y Salkie (1996: 556), Grefenstette (1998: 39), Kalliokuusi y Varantola (1998: 606), Nesi (1999: 62), Alonso Ramos (2009: 1194), Rogers y Ahmad (1998: 193), Varantola (1994: 606), etc. Varantola (2002b: 35), por ejemplo, aboga por un diccionario electrónico que incluya acceso a recursos lingüísticos, como corpus y programas de análisis textual e incluso memorias de traducción; o en la dirección inversa, corpus etiquetados sintácticamente y semánticamente que estén vinculados a entradas de un diccionario y que permitan un acceso bidireccional entre corpus y diccionarios (Varantola 2006: 224).

In my vision, the future dictionary is rather an integrated tool or a number of tools in a professional user's toolbox where it coexists with other language technology products such as encyclopedic sources of reference, different types of corpora, corpus analysis tools, such as WordSmith tools or Word Sketches type software, as well as corpus compilation software, translation memory systems, etc. (Cf. the respective Web pages) An integrated dictionary would be compatible with other similar tools. It should thus be easy to move from one tool to another and also to customize the dictionary to match the user profile and individual preferences.

(Varantola 2002b: 35)

Desde la aparición de los corpus electrónicos los estudios han evidenciado la necesidad de estos recursos en traducción, no solo para enseñar a traducir, sino también como recurso de búsqueda de los traductores. En este sentido, se ha señalado que el traductor precisa el uso de nuevos recursos terminológicos como los corpus (Bowker 1998b: 648; Bowker 2000: 21; Hull 2001: 225; Montero Martínez y Faber 2008: 162), que complementen a diccionarios y bases de datos (Faber et al. 2001: 167; Kraif 2008: 84;



Shreve 1997: 778) y satisfagan problemas terminológicos puntuales de una forma más rápida y fiable.

En esta línea, Zanettin (1998: 617) señala que el uso de corpus para la enseñanza del traductor ha sido habitual incluso antes de que aparecieran los corpus electrónicos. Snell-Hornby (1988) y Schäffner (1998), por ejemplo, indican que estudiando textos del mismo género pueden identificarse rasgos prototípicos que ayuden al traductor en la producción del texto meta.

No obstante, hay que señalar que, en un principio, los corpus no están pensados para el traductor. Otros autores ya han subrayado anteriormente que los corpus no están diseñados específicamente para traductores, ni evaluados por traductores, por lo que no cubren todas las necesidades de búsqueda de los traductores, ni son lo suficientemente intuitivos para que un traductor pueda utilizarlos sin haber recibido una instrucción previa (Wilkinson 2006b).

En este sentido, como veíamos anteriormente con los diccionarios, también resulta necesario dar a conocer los corpus y, especialmente, enseñar a utilizar los corpus a los traductores, desde sus niveles más iniciales de formación, para que los adopten como una herramienta más y necesaria dentro de su abanico de recursos de búsqueda terminológica. Alcina (2003: 80), por ejemplo, aboga por que los estudiantes de traducción reciban una buena formación dentro de las aulas respecto al uso de nuevas herramientas y recursos electrónicos de manera que den respuesta a la demanda de las empresas e instituciones y realiza una propuesta didáctica dividida en cuatro niveles de especialización en terminológica, en los que incluye la consulta de corpus en línea u otros soportes, así como el uso de programas de gestión de corpus (Alcina 2003: 84-87).

Muchos otros autores que investigan sobre el uso de corpus por traductores, también sugieren y recomiendan que se incluyan en los currículos de los estudios de traducción la enseñanza de corpus, para que los traductores aprendan a utilizarlos eficazmente sobre todo en la traducción especializada (Aston 2000: 28; Bernardini et al. 2003: 2; Bowker 2006: 185; Corpas Pastor y Seghiri 2009: 77; Friedbichler y Friedbichler 2000: 113; Kübler 2003: 41; Laviosa 2003a: 112 y 115; Maia 2003: 52; Philip 2009: 72; Rodríguez Inés 2009: 129; Sánchez-Gijón 2009: 109; Varantola 2002a: 184, 2003a: 56; Wilkinson 2005b; Zanettin 2002a, entre otros). Beeby et al. (2009: 1) señalan que es durante su formación cuando los traductores disponen del tiempo suficiente para aprender a utilizar los corpus ya que los traductores profesionales siempre se encuentran bajo presión de plazos. Sánchez-Gijón (2009: 115) además indica que las habilidades de documentación

con corpus se adquieren mejor en clase de traducción, utilizándolos para resolver problemas de traducción. Varantola (2006: 224) añade que antes de empezar a enseñar estas nuevas herramientas en los programas de traducción, es necesario que los docentes también se formen y entiendan el verdadero potencial de estos nuevos recursos.

Actualmente, los traductores deben adaptarse a los corpus con el fin de obtener de ellos el máximo beneficio, y esperar a que con el tiempo estos recursos vayan adaptándose a sus necesidades (Vargas Sierra 2002). Entre tanto, los estudios abogan por una mayor cooperación en las tareas de investigación y desarrollo entre los diseñadores de herramientas, en este caso los corpus, y los usuarios finales de dichas herramientas, los traductores (Jääskeläinen y Mauranen 2004: 53). Varantola (2002a: 183; 2006: 223-224), por ejemplo, reclama *intelligence in corpora*, herramientas de análisis de corpus que automaticen operaciones de búsqueda utilizadas frecuentemente por los traductores, que permitan organizar y sistematizar los corpus *ad hoc*, que guarden automáticamente los textos en un formato adecuado para su posterior análisis, que incluyan listas de palabras de campos especializados para utilizar con la función de palabras clave, que contengan lematizadores para lenguas diferentes, etc.

En resumen, en este apartado hemos visto la necesidad de los corpus en traducción, como un recurso complementario a los diccionarios que resulta útil tanto para la enseñanza de la traducción como para la búsqueda de terminología en traducción. Por ello, resulta interesante observar y clasificar primero los tipos de corpus que existen, así como las posibilidades de búsqueda que ofrecen este tipo de herramientas, antes de profundizar en el uso específico de los corpus por parte de los traductores (tema que trataremos en el capítulo 4).

### **3.3.2 Tipología de corpus y las posibilidades de búsqueda, según la bibliografía**

En este apartado trataremos de determinar qué tipos de corpus electrónicos existen actualmente para realizar búsquedas terminológicas y qué ventajas e inconvenientes aportan cada uno. Diversos trabajos han abordado la tipología de corpus (Berber Sardinha 2004; Colominas 2004; Granger 2003; Sánchez-Gijón 2003a, 2004a; Sinclair 1996; Zanettin 2002a).

Los corpus pueden clasificarse en función de distintos criterios como la modalidad (corpus oral o escrito), aspectos cronológicos (sincrónico si representa el estado actual de la lengua, diacrónico si muestra su evolución y periódico o cronológico cuando se refiere a

un momento histórico particular), en cuanto a los aspectos temáticos (corpus general o especializado)<sup>37</sup>. A continuación profundizaremos en los corpus en función del modo de compilación y de su naturaleza puntual o estable (corpus estables y corpus *ad hoc*), del marcado (bruto o etiquetado), del número de lenguas (monolingüe, bilingüe o multilingüe), y de la relación que se establece entre los textos (comparable, paralelo y alineado). Una vez hayamos identificado cuáles son los tipos de corpus y explicado sucintamente las utilidades de cada uno para los traductores, pasaremos a establecer el análisis de las técnicas de búsqueda.

### 3.3.2.1 Según la compilación y la naturaleza puntual o estable

En primer lugar, los corpus pueden diferenciarse en función del modo de compilación del corpus y su naturaleza puntual o estable. Los **corpus estables** (*robust, proper, stable o ready-made corpus*) (Corpas Pastor 2004: 226) suelen ser amplios, su diseño y tamaño son fijos, aunque puedan variar dependiendo de las necesidades de los usuarios y están preparados para ser manipulados por un programa de gestión de corpus; incluso algunos contienen su propio sistema de recuperación de la información. Son corpus prefabricados, que el usuario encuentra ya compilados. Algunos ejemplos de corpus estables son el *British National Corpus (BNC)*, *Corpus de Referencia del Español Actual (CREA)*, *Corpus Diacrónico del Español (CORDE)*, *Hellenic National Corpus (HNC)*, etc.

Según algunos autores, existen distintos niveles de estabilidad (*robustness*) en función del equilibrio y la representatividad de los textos seleccionados, de que el corpus proporcione tanto información lingüística, como extralingüística (que el corpus esté anotado) y los medios de consulta de que disponga para acceder a la información del corpus (McEnery y Wilson 1996).

Algunos corpus estables se consideran **corpus de referencia**, si proporcionan información global o exhaustiva de una lengua y son lo suficientemente grandes para representar todas las variantes lingüísticas y las características del vocabulario de una lengua (Sinclair 1996). Colominas (2004: 364) explica que los corpus de referencia incluyen un «número muy elevado de textos escritos en una misma lengua con diferentes registros, estilos, géneros, etc. Pretenden ser una muestra representativa de esa lengua». Esta autora considera a los corpus de referencia un recurso complementario de los

---

<sup>37</sup> Véase Sánchez-Gijón (2004a: 74).

diccionarios monolingües, tanto en la lengua de partida, como en la de llegada. Por ejemplo, en un corpus de referencia en lengua origen el traductor puede encontrar información de los contextos de uso o el registro de una palabra y en un corpus de referencia de la lengua meta se puede comprobar el carácter colocacional de una combinación de palabras.

Por oposición a los corpus estables (que como hemos visto suelen ser grandes y se encuentran ya compilados), encontramos un tipo de corpus, normalmente de tamaño más reducido, menos estable (Zanettin 2002a: 12) y con una finalidad específica (Cabré 2006: 95). Este corpus ha recibido distintas denominaciones: *corpus ad hoc* (Aston 1999; Corpas Pastor 2001a, 2002; Sánchez-Gijón 2004a, 2004b), *special purpose corpus* (Pearson 1998), *corpus specials* (Sánchez-Gijón 2003a) porque es un corpus con una intención determinada, normalmente representa una parte de la lengua, por ejemplo, puede estar formado por documentos de un determinado género textual, como los seguros de viaje (Corpas Pastor y Seghiri 2009); y también suele llamarse *do-it-yourself (DIY) corpus* (Maia 1997; Zanettin 2002b), *customized corpus* (Austermühl 2001) o *virtual corpus* (Ahmad et al. 1994) porque no se encuentra ya compilado, sino que el traductor debe compilarlo por sí mismo. En palabras de Corpas Pastor (2004: 236) un corpus *ad hoc* «se compila puntualmente para la realización de un determinado encargo de traducción en cualquier dirección (directa, inversa o indirecta)». Wright & Wright (1997: 147) explican la importancia de la gestión terminológica *ad hoc* por parte del traductor: «Translators are frequently called upon to deal with texts of every possible length taken at random from greater contexts of which they have no prior knowledge. [...] The critical factor affecting this kind of “one-off” text production is the fact that such translators must practice *ad hoc terminology management*». Por tanto, el traductor compila su propio corpus virtual atendiendo a las necesidades concretas de su traducción (Corpas Pastor y Seghiri 2009: 78). Precisamente este carácter puntual o efímero lo convierte también en un *disposable corpus* (Varantola 2000, 2003a) porque no forma parte de un corpus estable permanente sino que se crea para una traducción determinada y puede descartarse una vez la traducción ha finalizado, por lo que tampoco plantea problemas de copyright (Zanettin 2002b: 242).

Este tipo de corpus suele compilarse a partir de textos extraídos de Internet, que es la fuente para la creación de corpus más habitual de los traductores, ya que siempre se encuentra disponible, es la más completa, y facilita la tarea de compilación de corpus (Zanettin 2002a: 12). No suelen existir corpus de temáticas específicas disponibles en línea y, si existe alguno, normalmente no resulta representativo y no cubre todas las necesidades

del traductor. Por este motivo, numerosos autores señalan a los corpus *ad hoc* como recursos útiles para una traducción especializada (Bernardini y Zanettin 2000; Corpas Pastor 2001a, 2004; Corpas Pastor y Seghiri 2009; Pearson 1998; Sánchez-Gijón 2003a, 2003b; Zanettin 2002b). En este punto cabe señalar que, aunque la mayoría de corpus *ad hoc* se compilan con textos descargados desde Internet, se recomienda que su consulta se realice con programas de gestión y explotación de corpus (Montero Martínez y Faber 2008: 171; Zanettin 2002a: 12), que analizaremos en el siguiente apartado. Estos programas según apunta Zanettin son intuitivos y fáciles de utilizar, aunque no necesariamente estén pensados para satisfacer las necesidades de los traductores.

El problema de la compilación de corpus *ad hoc* es que este proceso requiere la dedicación por parte del traductor de un tiempo muy valioso, que probablemente necesite para otras etapas del proceso de traducción. Por ello, son muchos los estudios dirigidos a proporcionar métodos para la recuperación de corpus virtuales de la forma más rápida y eficaz posible (Maia 1997), tanto multilingües paralelos y comparables (Bowker y Pearson 2002: 95-102; Corpas Pastor y Seghiri 2009: 79-92; Gasiglia y Paulussen 2008; Resnik 1998, 1999; Resnik y Smith 2003), como monolingües sobre todo en lengua meta y especializados (Corpas Pastor 2001a, 2004; Hurskainen 2008; Maia 1997; Montero Martínez y Faber 2008: 171-180; Nazar et al. 2008; Sánchez-Gijón 2009: 115-119; Sharoff 2006; Wilkinson 2006a, 2010).

Harvey Ciampi (2004), por ejemplo, presenta la herramienta *WWW Search Interfaces for Translators* que facilita a los traductores la realización de consultas eficaces para la recuperación de textos paralelos y glosarios en los buscadores de Internet. En esta misma línea de trabajo también se encuentran los trabajos de Baroni y Bernardini (2004) que han diseñado un conjunto de aplicaciones en Perl para recuperar corpus y términos especializados de Internet<sup>38</sup>. Estas aplicaciones requieren su descarga e instalación en el ordenador, por lo que es necesario que el ordenador cumpla con una serie de requisitos técnicos (como tener instalado Perl) y que el usuario tenga conocimientos de informática. Por ello, para facilitar el uso de estas aplicaciones, se desarrolló la interfaz de búsqueda *WebBootCat* (Baroni et al. 2006), que se encuentra disponible en línea y facilita la descarga de los textos de Internet simplemente con la selección de una serie de opciones de la interfaz. Actualmente *WebBootCat* se encuentra alojada en una interfaz de análisis de

---

<sup>38</sup> Estas aplicaciones están disponibles para su descarga en <http://sslmit.unibo.it/~baroni/bootcat.html>.

corpus, también disponible en línea, llamada *Sketch Engine* (Kilgarriff et al. 2004). Con ello, se ha facilitado además que, una vez descargados los textos de Internet con *WebBootCat*, puedan analizarse inmediatamente con el uso de *Sketch Engine*.

Además de *WebBootCat* y *Sketch Engine*, la mayoría de herramientas de gestión y análisis de corpus que se diseñan actualmente están encaminadas a que se permita descargar directamente los textos de Internet y analizarlos sin necesidad de convertir archivos ni ninguna otra tarea que ralentice el proceso de búsqueda, ejemplo de ello son *WordSmith Tools 4*, *Corpógrafo* (Sarmiento et al. 2004), *Jaguar* (Nazar et al. 2008), o *TerminoWeb* (Barrière 2009), entre otros, que también analizaremos en el siguiente apartado. Todos estos programas (la mayoría disponibles además en línea, excepto *WordSmith*) facilitan y agilizan la tarea de compilación de corpus virtuales especializados, ya que descargan los textos directamente desde Internet y los analizan con el uso de la misma herramienta. Algunos incorporan a las búsquedas de textos en línea criterios de calidad, como que incluyan expresiones de tipo semántico (Barrière 2009).

Se ha sugerido que la compilación de un corpus a partir de páginas de Internet presenta ventajas respecto al uso de toda la Web como un corpus. Sánchez-Gijón (2009: 114), por ejemplo, comenta que la búsqueda de terminología directamente en Internet puede provocar inconsistencias en la traducción, ya que cada solución a un problema de traducción se obtiene de un contexto diferente. Dicho de otra manera, Internet como corpus resulta en ocasiones demasiado amplio, los textos pertenecen a temáticas muy diferentes y no están compilados para cubrir una necesidad de traducción y una temática concreta. En cambio, la compilación de un corpus virtual permite restringir el corpus a una determinada temática, así como la consulta a través de programas específicamente diseñados para la consulta de corpus, que ofrecen mayores posibilidades de búsqueda que los buscadores de Internet.

Asimismo, en un corpus *ad hoc*, no se aplican los criterios de calidad de forma absoluta, sino que más bien hay que tener en cuenta el valor relativo del conjunto del corpus en función del texto origen y del encargo de traducción (Sánchez-Gijón 2004a: 79-81).

Aún así, no se puede negar que la búsqueda directa en Internet ofrece la ventaja de la inmediatez de la información. En los últimos años, para mejorar las búsquedas de terminología en Internet, están surgiendo nuevas herramientas de explotación de la Web como corpus, que tratan de imitar las funciones de búsqueda de los programas de explotación de corpus para la consulta en Internet. Ofrecen concordancias en formato

KWiC, e incluso algunos disponen de otras funciones como la generación de listas de palabras y colocaciones. Más adelante hacemos una revisión de estas herramientas<sup>39</sup>.

### 3.3.2.2 Según el mercado

Un corpus puede encontrarse en bruto o sin etiquetar, o incluir algún tipo de anotación o marcado. En un **corpus anotado o etiquetado** pueden marcarse distintos elementos: lingüísticos, morfosintácticos, lexicosemánticos, también elementos macroestructurales y microestructurales, como los párrafos y las oraciones de los textos, así como elementos extralingüísticos, como el autor o la fecha de un texto (Zanettin 2000: 114). Cada forma de etiquetado permitirá un tipo de consulta (Sánchez-Gijón 2003a: 188).

En la actividad de traducción, como sugiere Colominas (2004: 369), un corpus anotado o etiquetado permite restringir la búsqueda mediante información lingüística (lemas, categorías gramaticales y sintaxis). Esta autora explica que la búsqueda en un corpus anotado es especialmente útil cuando en la traducción existen diferencias estructurales entre las lenguas. Por ejemplo, pueden conocerse, a través de las categorías gramaticales, las palabras más frecuentes en los textos paralelos en lengua de llegada. Una comparación de esta información con la correspondiente en la lengua de partida puede ayudar al traductor a evitar calcos y obtener una traducción más adecuada estilísticamente.

Bowker y Pearson (2002: 88) señalan algunas consultas que pueden ofrecer los corpus etiquetados gramaticalmente, por ejemplo la distinción de homógrafos (*process*, *result* y *experiment* como sustantivos o como verbos, e incluso *cross* como sustantivo, verbo o adjetivo), en la generación de una lista de palabras de un corpus sin etiquetar *process* aparecerá una sola vez, por lo que el usuario deberá acceder a todos los contextos del corpus si quiere identificar cuándo se emplea como sustantivo o como verbo; en un corpus etiquetado la lista de palabras incluiría tanto *process\_verb* como *process\_noun*. También pueden introducirse categorías gramaticales en lugar de palabras como expresiones de búsqueda, por ejemplo, pueden recuperarse términos compuestos de un corpus que cumplan los siguientes patrones gramaticales: adjetivo+nombre, nombre+nombre.

Además de la anotación de categorías gramaticales, cada vez más habitual en la mayoría de los corpus, existen otros tipos de anotaciones, aunque no tan frecuentes en la

---

<sup>39</sup> Véase el apartado 3.4.3.2.

actualidad, como la anotación sintáctica (*parsing*), la anotación semántica y la anotación de relaciones anafóricas (McEnery y Wilson 1993).

Según estos autores la anotación sintáctica identifica en el corpus información sobre la estructura sintáctica de los textos (oraciones de relativo, sintagmas nominales, preposicionales y verbales). Este tipo de anotación permite la alineación de textos a nivel de los constituyentes sintácticos lo que supone un avance respecto a la alineación a nivel de oración, por ejemplo, para la traducción automática. Sin embargo, autores como Jakubíček et al. (2010) señalan que la anotación sintáctica todavía no se encuentra suficientemente desarrollada y supone mucho tiempo y esfuerzo, particularmente en corpus muy extensos. Por ese motivo, los lenguajes de consulta de corpus, Corpus Query Language (CQL), que incluyen los sistemas de gestión de corpus como Manatee (Rychlý 2007) o IMS CorpusWorkbench (Christ et al. 1999), utilizados con algunos corpus que analizaremos a continuación, permiten realizar búsquedas sintácticas en corpus anotados gramaticalmente y no sintácticamente. Por ejemplo para recuperar sintagmas nominales u oraciones de relativo introducidas por *that*, se utiliza el comando **within <structure>**.

```
[tag="N.*"]+ within <s/>
```

```
[lemma="that"] within [tag="AT0"]* [tag="AJ0"] [tag="N.*"] [0,20] [tag="V.*"] within <s/>
```

(Jakubíček et al. 2010)

Recursos basados en estos sistemas, como *Sketch Engine*, permiten asimismo generar un *word sketch*, un esquema del comportamiento gramatical y colocacional de una palabra en un corpus (Kilgarriff et al. 2004).

Otro tipo de anotación, la anotación de relaciones anafóricas (McEnery y Wilson 1993: 6) puede resultar de gran utilidad para la traducción automática en la traducción de pronombres, ya que el género y número del antecedente determina también el del pronombre, por ejemplo con el pronombre *they* en inglés, cuya traducción al francés dependerá del antecedente al que se refiere; en este caso, si fuera *the girls*, se traduciría por *elles* en lugar de *ils*.

Finalmente, la anotación semántica, mencionada también por estos autores, permite la clasificación de las relaciones semánticas del corpus como en un tesoro o estructuras del tipo agente/paciente. En algunos corpus como los desarrollados por Davies (2005) se incluyen bases de datos relacionadas que incluyen, entre otros datos, sinónimos que



permiten realizar búsquedas por relaciones semánticas en el corpus, como veremos en nuestro análisis.

En los últimos años se está mostrando un creciente interés por otros tipos de anotaciones de corpus, como la anotación sintáctico-semántica de la estructura actancial de las unidades léxicas de un corpus a través del etiquetado de roles semánticos (Agente, Paciente, Destino, Instrumento, etc.). Este tipo de anotaciones de corpus se están desarrollando actualmente en proyectos<sup>40</sup> como *FrameNet* (Lowe et al. 1997) basado en la Semántica de Marcos o *DiCoInfo* y *DicoEnviro* (Hadouche et al. 2009) basados en la Lexicología Explicativa y Combinatoria.

Varantola (2006: 223-224) señala que un corpus en bruto, sin anotar o anotado solo gramaticalmente es un corpus de una *inteligencia superficial*, por lo que sería deseable avanzar en el desarrollo de corpus de *inteligencia profunda* que ofrezcan un mayor potencial, como los corpus etiquetados sintáctica y semánticamente; para ello, serán necesarias herramientas sofisticadas de compilación de corpus, que incorporen etiquetadores y analizadores sintácticos automáticos.

### 3.3.2.3 Según el número de lenguas y la relación que se establece entre los textos

Aunque habitualmente en la investigación se ha producido una divergencia terminológica al designar a este tipo de corpus, Sánchez-Gijón (2003a: 190) y Zanettin (2002a: 10), entre otros, siguiendo la tipología textual de EAGLES, elaborada por Sinclair (1996), clasifican los corpus según el número de lenguas en **corpus monolingües** y **corpus multilingües** (incluidos los **corpus bilingües**) y, dentro de los corpus multilingües, según la relación que se establece entre los textos de las distintas lenguas, distinguen entre **corpus comparables** y **corpus paralelos**.

Dentro de los corpus monolingües, Granger (2003: 20) diferencia dos subtipos de corpus monolingües comparables, uno que contiene textos *originales* y *traducidos* en la misma lengua, denominado **corpus comparable monolingüe**, *Monolingual Comparable Corpus* por Baker (1993), y otro compuesto por textos escritos en una misma lengua por un nativo y por un estudiante de la lengua (*native and learner texts*). Según Zanettin (2002a: 10), el corpus monolingüe comparable es interesante desde el punto de vista de la

---

<sup>40</sup> Véase el apartado 3.2.3 donde se analizan *FrameNet* y *DiCoInfo*.

investigación teórica en traducción y de la enseñanza de la traducción, pero resulta menos relevante para traductores profesionales que utilizan el corpus como recurso terminológico en un encargo de traducción. Otro corpus que también resulta útil desde el punto de vista pedagógico e investigador, según Bowker y Bennison (2003), son corpus de traducciones de estudiantes de traducción.

Para la práctica de la traducción, sin embargo, como señala Zanettin (2002a: 10), son de mayor interés los corpus monolingües (compuestos por textos originalmente escritos tanto en lengua origen como en lengua meta) y los corpus multilingües (tanto comparables como paralelos).

En lo que respecta a los corpus monolingües, en palabras de Zanettin (2002a: 11), el corpus monolingüe puede complementar al diccionario monolingüe: «Translators can use target monolingual corpora alongside target monolingual dictionaries to check the meaning and usage of translation candidates in the target contexts. Like source language dictionaries, source language corpora can be consulted for source text analysis and understanding». Los corpus monolingües en lengua de partida pueden ser útiles para resolver tareas de comprensión del texto origen, tal y como también se señala en Sánchez-Gijón (2004b). Por ejemplo, en un corpus monolingüe de la lengua de partida encontramos información acerca del significado de una palabra de la lengua origen, los contextos de uso o el registro.

Además, los corpus monolingües en lengua meta pueden utilizarse para acceder a equivalentes de traducción. Sin embargo, como indica Zanettin (2002a: 10), en la búsqueda de equivalentes los corpus con los que más familiarizados están los traductores son los corpus bilingües, especialmente los corpus paralelos.

Los **corpus bilingües** contienen textos en dos lenguas y los **multilingües** presentan documentos como mínimo en tres lenguas diferentes. Los **corpus paralelos**, *translation or parallel corpora* según la terminología empleada por Granger (2003: 20), son aquellos compuestos por una colección de textos originales con sus traducciones a una o más lenguas (Colominas 2004: 364; Granger 2003: 20). Un corpus paralelo puede ser **unidireccional** (de la lengua A a la lengua B, C o D), **bidireccional** (de la lengua A a la lengua B, C o D y de la lengua B, C o D a la lengua A) y **multidireccional** (de la lengua A, B, C o D a cualquiera de las otras lenguas) (Zanettin 1998: 616).

Los corpus paralelos bidireccionales o recíprocos (Aston 1999), textos originales en lengua A alineados a sus traducciones en lengua B, y textos originales en lengua B alineados a sus traducciones en lengua A, comprenden en un solo corpus las características

de distintos tipos de corpus, un corpus paralelo unidireccional (que es el tipo de corpus paralelo más estudiado en la bibliografía), un corpus comparable bilingüe con textos originales comparables en dos lenguas o con textos traducidos comparables en dos lenguas, así como también un corpus comparable monolingüe (que tratábamos al principio de este subapartado) con textos originales y traducidos en la misma lengua.

En este punto, es importante indicar que los pares de textos en un corpus paralelo no tienen por qué ser siempre traducciones directas los unos de los otros, sino que pueden ser traducciones de terceros textos. Por ejemplo, un corpus paralelo francés-alemán puede contener textos en francés y alemán que sean traducciones de textos en inglés, por lo que ni el francés ni el alemán serían textos originales. Bowker y Pearson (2002: 93) señalan que esto sucede con frecuencia en contextos multilingües, como en las traducciones de la Unión Europea.

Según explica Sánchez-Gijón (2003a: 191) el corpus paralelo además de contener textos equivalentes en una o más lenguas, suele estar alineado, normalmente a nivel de frase, con lo que cada oración en una lengua está vinculada a la oración equivalente en otra o en otras lenguas. Este tipo de corpus ha recibido también el nombre de **corpus alineado**.

Según Bowker y Pearson (2002: 93), la principal característica de los corpus paralelos alineados es que permiten la generación de concordancias bilingües. El usuario introduce una palabra o secuencia de palabras en una lengua y el programa de análisis recupera todos los segmentos del corpus en los que aparece dicha palabra o secuencia junto con sus segmentos correspondientes en otra lengua, y normalmente los segmentos equivalentes se muestran uno al lado del otro (por ejemplo en dos columnas).

Zanettin (2002a: 11) establece un paralelismo entre los corpus paralelos y los diccionarios bilingües, en cuanto que ambos recursos pueden utilizarse para acceder a equivalencias. Sin embargo, este autor establece algunas diferencias: los diccionarios bilingües son repertorios de equivalentes léxicos (en los diccionarios generales) o términos (en los diccionarios especializados) creados por terminólogos. Sin embargo, los corpus paralelos son repertorios de equivalentes de traducción generados por traductores. Además, como señalan Zanettin (2002a: 11) y otros autores como Laviosa (2002: 104) o King (2003: 165-166), los corpus paralelos ofrecen ventajas en el acceso a los equivalentes, frente a los diccionarios bilingües. Por ejemplo, los términos aparecen rodeados de un contexto y de un número elevado de ejemplos reales de uso; además ofrecen soluciones en casos de equivalencia cero. Además, Teubert & Čermáková (2004: 123) afirman que el diccionario bilingüe más grande presenta solo una pequeña porción de los equivalentes que

pueden encontrarse en un corpus paralelo que no sea demasiado reducido. Añaden que en el diccionario faltan unidades largas y complejas, así como los contextos en los que se utilizan dichas unidades. Cuando se busca un equivalente en un diccionario bilingüe, normalmente es tarea del traductor averiguar en qué contextos se emplea dicho equivalente, ya que esta información no suele proporcionarla el diccionario. En los corpus paralelos, los equivalentes entre lenguas están inmersos en un contexto.

Los corpus paralelos pueden resultar de especial utilidad al traductor para observar traducciones anteriores de secuencias de palabras o términos técnicos (Somers 2003: 26) especialmente con el uso de memorias de traducción (Somers 2003: 31-47).

Asimismo, existen herramientas específicamente diseñadas para que los traductores se ayuden de los corpus paralelos o memorias de traducción en el ejercicio de su profesión, como son los programas de traducción asistida (por ejemplo *SDL Trados*, *Transit*, *Déjà Vu*, etc.) (Danielsson y Ridings 1996: 57), cuya eficacia está en cierto modo avalada por su uso cada vez más generalizado por parte de los traductores profesionales. Para Zanettin (2002a: 10) las memorias de traducción funcionan como bancos de datos en los que se recuperan de forma automática fragmentos de traducciones anteriores, pero también pueden consultarse manualmente en busca de términos y patrones específicos.

Las memorias de traducción entrañan una serie de ventajas para el traductor, especialmente en lo que concierne a la búsqueda de terminología. Por ejemplo, la búsqueda automática de terminología en las bases de datos anexas a la memoria garantiza que la consistencia terminológica, el almacenamiento y la recuperación de terminología sea flexible y rápida de recuperar y de actualizar por medios electrónicos; otra ventaja es la posibilidad de pegar términos directamente de la base de datos a la traducción, sin necesidad de teclearlos. Asimismo, los programas de traducción asistida permiten el intercambio de información entre traductores que pueden acceder y contribuir a una misma base de datos cuando trabajan en un mismo proyecto, lo que también garantiza la consistencia terminológica. También se permite la transferencia de información a clientes y otros traductores, siempre y cuando se garantice la importación y exportación a formatos estándar. Por último, las funciones de extracción terminológica de un programa de traducción asistida permiten al traductor la generación de listas de terminología monolingüe o bilingüe para la creación de nuevas bases de datos (Bowker 2003b: 58-60).

No obstante, algunos autores sugieren una serie de inconvenientes de los corpus paralelos: las diferencias lingüísticas entre los textos originales y los textos traducidos, la edición de algunos textos originales antes de ser traducidos, la necesidad de tener

conocimientos especializados para decidir si la traducción recuperada es adecuada en un contexto determinado, la dudosa calidad de las traducciones proporcionadas y de las competencias del traductor, que normalmente se desconocen, especialmente por el hecho de que hasta los mejores traductores no siempre generan traducciones perfectas (Friedbichler y Friedbichler 2000: 110).

Los **corpus comparables bilingües** contienen textos (normalmente originales) similares y comparables en dos o más lenguas (Colominas 2004: 364; Sinclair 1996), que comparten contenido, temática o función comunicativa (Zanettin 1998: 617), tipo de texto, período en el que se escribieron los textos, grado de especialidad (Bowker y Pearson 2002: 93). Por ejemplo, pueden ser textos del mismo género en distintas lenguas (Granger 2003: 20).

En la bibliografía a menudo se produce una confusión terminológica entre *paralelo* y *comparable*, ya que se consideran *textos paralelos* a los textos en lengua origen y en lengua meta que comparten las mismas características y que se han producido en situaciones comunicativas similares (Schäffner 1998; Williams 1996).

En cuanto a las aplicaciones en traducción de los corpus comparables, Granger (2003: 19) señala que los corpus comparables presentan la ventaja de que están compuestos por textos originales, por lo que el lenguaje que se utiliza no está influenciado, en principio, por otras lenguas, por oposición a los corpus paralelos que, como hemos visto, uno de sus principales problemas es que los textos en lengua meta están inevitablemente influenciados por la lengua origen, por lo que no siempre son fiables en la elección de un equivalente en lengua meta. Además, no siempre es posible encontrar corpus paralelos de cualquier temática.

Sin embargo, el acceso a equivalencias en los corpus paralelos es más sencilla y rápida que en los comparables. El principal inconveniente de los corpus comparables, según Granger (2003: 19), es que resulta más difícil establecer una comparación entre los textos de diferentes lenguas y culturas, especialmente en textos con rasgos culturales muy marcados, en los que el establecimiento de equivalencias exactas entre lenguas, en ocasiones, no es posible.

Por eso, la bibliografía deja constancia de que al igual que existen herramientas para analizar corpus monolingües y corpus paralelos, se hace necesaria más investigación en programas capaces de buscar automáticamente en corpus bilingües comparables (Laviosa 2002: 103; Zanettin 1998: 12-13). En este sentido, Friedbichler y Friedbichler (2000: 110) señalan que una manera de resolver el inconveniente de la falta de vínculo entre lenguas

(*missing interlingual link*) sería alinear de forma indirecta los corpus comparables con la ayuda de una lista bilingüe de palabras clave específicas del área especializada. Podemos ver un ejemplo de proyecto de búsqueda terminológica semiautomática en un corpus comparable en Kilgarriff et al. (2011: 125-128).

### 3.3.3 Análisis empírico de técnicas de búsqueda en corpus

En el estudio de las técnicas de búsqueda en corpus analizamos, en primer lugar, una selección<sup>41</sup> de corpus estables (monolingües y bilingües o multilingües, generales y especializados) disponibles en línea<sup>42</sup>, que incorporan sus propias interfaces de consulta.

El análisis se efectúa desde el punto de vista de conocer las funciones de búsqueda que ofrecen las distintas interfaces de los siguientes corpus. El *Corpus de referencia del español actual* y *Corpus diacrónico del español (CREA y CORDE)* que comparten la misma interfaz de búsqueda. Los corpus de Brigham Young University (BYU) con una interfaz igual: *el Corpus del español*, *Corpus of Contemporary American English*, *BYU-British National Corpus*, *TIME Corpus*, *BYU-OED Oxford English Dictionary* y *Corpus do Português*. El *British National Corpus (BNC)* y otras interfaces para la consulta del BNC como *BNCweb (CQP-Edition)* y *Phrases in English*. También el *Hellenic National Corpus (HNC)*, los corpus del IULA en *Bwananet*, los corpus del OLST en *Le Migou*, el corpus *Scientext*, *CLUVI*, *COMPARA*, *EUR-Lex* y *TAUS Search*.

Asimismo, hemos analizado un conjunto de herramientas de gestión de corpus que incorporan novedades en cuanto a la consulta de corpus y, sobre todo, en cuanto a la descarga y análisis de corpus virtuales. Veremos *Corpógrafo*, *Jaguar*, *Sketch Engine*, *TerminoWeb* y *WordSmith Tools 4*. Todos estos programas tienen en común que permiten cargar y analizar corpus ya compilados y proporcionados por el usuario<sup>43</sup>, pero también, en

---

<sup>41</sup> Somos conscientes de la imposibilidad de analizar todos los corpus disponibles actualmente en línea. Esta es una selección de algunos corpus que por accesibilidad, contenido u opciones de búsqueda hemos considerado interesantes. En la red existen numerosos catálogos de corpus en línea, como por ejemplo el siguiente de la Universidad de Essex: [http://www.essex.ac.uk/linguistics/external/clmt/w3c/corpus\\_ling/content/corpora/list/index2.html](http://www.essex.ac.uk/linguistics/external/clmt/w3c/corpus_ling/content/corpora/list/index2.html), o el de la Eastern Michigan University: <http://linguistlist.org/sp/GetWRLlistings.cfm?WRabbrev=Texts>.

<sup>42</sup> Última revisión del análisis con fecha de 20 de abril de 2012. Es posible que con posterioridad a esta fecha algunas funciones de los corpus hayan cambiado.

<sup>43</sup> Podemos encontrar colecciones de corpus ya compilados, disponibles gratuitamente en Internet, tanto monolingües, como bilingües paralelos, por ejemplo corpus publicados por organismos públicos, grupos de investigación de universidades, etc. Véanse los siguientes a modo de ejemplo:

- Corpus *BioMed Central*: <http://www.biomedcentral.com/info/about/datamining/>

- Corpus monolingües disponibles en la Oxford University Computing Services: <http://ota.ox.ac.uk/catalogue/index-id.html>

el caso de que el usuario no disponga de un corpus compilado, facilitan la descarga de un corpus desde Internet.

Por último, hemos examinado las funciones de búsqueda de algunos programas de traducción asistida, desde el punto de vista de la consulta de memorias de traducción, un tipo de corpus paralelo muy extendido en traducción. Algunos de los programas a los que hemos tenido acceso son *Déjà Vu*, *OmegaT* y *SDL Trados*.

Por medio del análisis de todas estas herramientas, hemos procurado presentar un panorama lo más amplio posible sobre posibilidades de búsqueda en los corpus. En este análisis observamos las funciones de cada una de las interfaces, que permiten realizar distintas búsquedas. A continuación describimos brevemente cada uno de los corpus analizados destacando las peculiaridades de cada uno. Además, podrán observarse ejemplos concretos de búsqueda en nuestra clasificación de técnicas de búsqueda en corpus en el apartado 3.3.4.

### 3.3.3.1 Corpus en línea que incorporan sus propias interfaces de búsqueda

La Real Academia Española cuenta con dos corpus monolingües de español disponibles en línea. El *Corpus de referencia del español (CREA)* contiene textos que reflejan el estado actual de la lengua española (desde 1975 a 2004). El *Corpus diacrónico del español (CORDE)* está compuesto por textos en español escritos hasta el año 1975 y permite estudiar el origen y evolución de las palabras y de sus significados. Ambos corpus incorporan la misma interfaz de búsqueda en la que pueden utilizarse criterios de distancia entre las palabras de búsqueda. También permite aplicar filtros como área temática, autor, obra, fecha, registro y área geográfica. Asimismo, se ofrece información estadística de los resultados de búsqueda, concordancias y agrupaciones de palabras.

En Brigham Young University (BYU), Davies diseñó una interfaz para la consulta en línea de un conjunto de corpus monolingües (para lo cual es necesario registrarse): el

- 
- Corpus monolingües y multilingües disponibles en la Universidad de Leeds (acceso a corpus de diferentes lenguas o a su consulta con la interfaz Leeds CQP): <http://corpus.leeds.ac.uk/list.html>
  - Corpus paralelo alineado inglés-francés *Hansards of the 36th Parliament of Canada*: <http://www.isi.edu/natural-language/download/hansard/>
  - Corpus paralelo en seis lenguas de la ONU: <http://www.uncorpora.org/>
  - *European Parliament Proceedings Parallel Corpus*: <http://www.statmt.org/europarl/>
  - *International Corpus of English*: <http://ice-corpora.net/ice/index.htm>
  - *JRC-Acquis Multilingual Parallel Corpus*: <http://langtech.jrc.it/JRC-Acquis.html>
  - *Opus open parallel corpus* (catálogo de corpus paralelos para la descarga o consulta en línea): <http://opus.lingfil.uu.se/index.php>

*Corpus del español* (español de los años 1200-1900), *Corpus of Contemporary American English* (inglés estadounidense de los años 1990-2008), *BYU-British National Corpus* (inglés británico de los años 1980-1993), *TIME Corpus* (inglés estadounidense de los años 1923-actualidad), *BYU-OED Oxford English Dictionary* (textos en inglés antiguo hasta 1990) y *Corpus do Português* (portugués de los años 1300-1900).

Esta interfaz de búsqueda permite buscar una o más formas, lemas o categorías gramaticales. Pueden aplicarse restricciones de categoría gramatical, género y período de tiempo. Una de las opciones de búsqueda permite acceder a concordancias que pueden ordenarse por la palabra de búsqueda o palabra pivote y alguna de las tres palabras inmediatamente a la izquierda o derecha de la palabra pivote. Otra función posibilita comparar la frecuencia de términos, secuencias de palabras o construcciones gramaticales por género o período de tiempo. Puede accederse a colocaciones de una palabra o comparar las colocaciones de dos palabras.

Otra de las características más peculiares de este corpus es que permite realizar búsquedas de distribución semántica. Esto significa que permite buscar sinónimos<sup>44</sup> de la palabra introducida en el corpus. Por ejemplo puede buscarse el adjetivo *ardiente* y sus sinónimos a una posición a la derecha del sustantivo *fuego* y sus sinónimos. En los contextos recuperados encontramos *llama ardiente*, *brasa ardiente*, *pasión ardiente*, *fuego encendido*, *brasa encendida*, *calor encendido*, *calor sofocante*, *fuego abrasador*, etc. Por último, pueden crearse listas personalizadas de palabras o frases para utilizarlas en las búsquedas. Por ejemplo, puede utilizarse una lista de colores y buscar en el corpus todos los colores de la lista [nombredeusuario:colores] que aparezcan junto a la palabra *mar*. Asimismo, es posible guardar los resultados de una búsqueda en una lista.

El *British National Corpus (BNC)* es un corpus monolingüe general de inglés británico actual. Contiene textos que representan tanto el lenguaje escrito como el hablado. Su interfaz en línea permite realizar una búsqueda simple de una palabra o secuencia en el corpus. Las búsquedas más complejas pueden realizarse directamente en la casilla de búsqueda en línea del *BNC* por medio del lenguaje de interrogación CQL o con el programa *SARA* o *XAIRA* que requiere su descarga e instalación en el ordenador.

También existe una interfaz que permite búsquedas más avanzadas en la versión en línea del *BNCweb (CQP-Edition)*. La versión en línea de *BNCweb (CQP-Edition)* requiere

---

<sup>44</sup> Para más información sobre las búsquedas semánticas se recomienda el artículo de Davies (2005: 326). Además, incluimos ejemplos de este tipo de búsqueda en nuestra clasificación de técnicas de búsqueda.



registrarse previamente. Permite utilizar el lenguaje CQL en las búsquedas, además de otras opciones que pueden utilizarse sin necesidad de conocer este lenguaje. Por ejemplo, genera listas de colocaciones de la palabra de búsqueda, en las que se puede restringir que las colocaciones sean de una determinada categoría gramatical. Indica la frecuencia de aparición y la distribución de una palabra o expresión en el corpus. Permite ordenar las concordancias en función de las palabras a izquierda y derecha de la palabra pivote o de búsqueda y es posible aplicar restricciones de categoría gramatical en esta función. Por ejemplo, se puede indicar a la herramienta que ordene las concordancias por los verbos que aparecen una posición a la derecha de la palabra pivote. También genera listas de palabras frecuentes del corpus, de formas o de lemas. Es posible especificar la categoría gramatical de las palabras de la lista o algún patrón lingüístico, por ejemplo, verbos que empiecen de una determinada forma, por ejemplo por *under*. También recupera listas de palabras clave, comparando, por ejemplo, la versión escrita y la versión oral del *BNC*. Es posible generar subcorpus de una determinada temática para su consulta. La herramienta busca los textos etiquetados en el corpus con un código temático determinado, por ejemplo, *humanities\_arts*. Además, contiene una función de historial de búsquedas y de almacenamiento de búsquedas.

Otra interfaz que permite la consulta del BNC es *Phrases in English* (desarrollada por Fletcher). Permite el acceso a concordancias, de palabras (pueden utilizarse comodines para introducir palabras truncadas) y también frases exactas o que contengan o no determinadas palabras con el uso de operadores booleanos. También puede indicarse que las palabras aparezcan en la frase a una distancia determinada, que una palabra aparezca a principio o final de frase, o en un orden determinado respecto a otras palabras; se puede buscar por forma o por lema y se puede restringir la búsqueda por categorías, por dominio, género y medio.

También accede a agrupaciones de palabras (*n-grams*), en las que puede indicarse que sigan un determinado patrón (añadiendo o excluyendo palabras o categorías gramaticales en determinada posición). La opción *phrase-frames* permite acceder a grupos de agrupaciones de palabras que son iguales excepto por una palabra, lo que ayuda a identificar patrones fraseológicos. La opción *PoS-Grams* genera listas de agrupaciones de categorías gramaticales. Los *chagrams* son agrupaciones de caracteres que se repiten con frecuencia dentro de las palabras del corpus, pueden filtrarse los resultados por la posición que deberán ocupar las agrupaciones de caracteres dentro de las palabras (cualquiera, principio, medio o final). La opción *Simple/RegEx Search Tabbed Output* presenta los

resultados en un archivo de solo texto separado por tabuladores que puede importarse a una base de datos. La función “*Drill down*” *RegEx search* accede a expresiones o frases más largas que incluyen las palabras de búsqueda e indica el porcentaje de aparición de determinadas frases dentro de frases más largas. Puede indicarse a la herramienta la dirección en la que queremos que expanda la búsqueda de frases (a derecha o izquierda de la palabra o expresión de búsqueda), por ejemplo frases que terminen por *floor*.

El *Hellenic National Corpus* (*HNC*) es un corpus monolingüe de griego, desarrollado por el Institute of Language and Speech Processing. Los textos del corpus representan el lenguaje escrito del griego moderno a partir del año 1990. Permite buscar hasta tres formas, lemas o categorías gramaticales que aparezcan en el texto con 0 a 5 posiciones de distancia<sup>45</sup>.

*Bwananet* es el programa de explotación de los corpus monolingües y paralelos especializados, en inglés, español y catalán, del Institut Universitari de Lingüística Aplicada (IULA) de la Universitat Pompeu Fabra. Los corpus incluyen textos de los ámbitos de informática, medio ambiente, derecho, medicina, genoma, economía y otros temas especializados.

La interfaz genera listas de formas, lemas o categorías del corpus seleccionado. El usuario puede buscar concordancias de una o más formas, lemas o categorías gramaticales (que pueden combinarse en la búsqueda). La restricción por categoría gramatical tiene dos opciones: (1) la búsqueda de secuencias variables de categorías (en las que una de las categorías de la secuencia puede aparecer una o más veces, de 0-9), por ejemplo, podemos buscar términos compuestos de dos o tres sustantivos seguidos de un verbo; (2) la restricción de una palabra que no sea de una determinada categoría. Otra función de restricción limita la búsqueda a una determinada parte del corpus (títulos, listas, tablas, resto del texto o cualquier parte).

Para realizar una búsqueda más compleja, el usuario debe conocer el lenguaje de interrogación del IMS Corpus Workbench, con el que se permiten búsquedas de un número ilimitado de unidades, de todo tipo de combinaciones de formas, lemas y categorías, interrogaciones multilingües con restricciones morfológicas en documentos paralelos, cálculo de frecuencias sobre formas, lemas o categorías, etc.

---

<sup>45</sup> Muchas interfaces de corpus únicamente permiten introducir dos elementos cuando se utilizan criterios de distancia entre ellos y solo permiten la combinación de tres o más elementos con el uso del lenguaje CQL.

*Le Migou* es una interfaz de consulta de los corpus compilados por el Observatoire de Linguistique Sens-Text (OLST) a partir de los cuales se han elaborado sus bases de datos. Permite consultar por una secuencia exacta o truncada una o dos palabras. En el caso de la consulta de dos palabras, puede especificarse que se busquen en el orden exacto introducido o con una distancia de máximo 10 palabras. Es posible buscar en todos los corpus al mismo tiempo o especificar uno. También puede limitarse el número de contextos recuperados.

*Scientext* contiene tres corpus monolingües de textos científicos, uno en francés de textos pluridisciplinarios de un poco menos de 5 millones de palabras, otro en inglés de textos de biología y medicina de unos 13 millones de palabras y otro en inglés de 1,1 millones de palabras compuesto por trabajos de estudiantes escritos en inglés como lengua extranjera (Tutin et al. 2009). Para la consulta, una vez seleccionado uno de los tres corpus, se muestra una interfaz donde podemos restringir la búsqueda a los tipos de documentos que nos interesen de dicho corpus. A partir de aquí se ofrecen tres tipos de búsqueda. La primera es una búsqueda semántica (*Recherche sémantique*) guiada donde podemos acceder a consultas preseleccionadas por los autores del corpus, como por ejemplo la búsqueda de adjetivos de opinión, con la que se recuperan *nous sommes persuadés, je suis conscient, on est convaincu*, etc. La segunda es la búsqueda libre (*Recherche libre*) donde pueden buscarse concordancias en el corpus de cualquier combinación de formas, lemas o categorías gramaticales, y también se puede restringir una forma o lema a una categoría gramatical determinada. Además pueden establecerse relaciones sintácticas entre las palabras de la búsqueda (por ejemplo, la palabra 1 es adverbio de la palabra 2 o de la palabra 3). La tercera es la búsqueda avanzada (*Recherche avancée*), con la que pueden efectuarse consultas más complejas con el uso de un lenguaje de interrogación de corpus con expresiones regulares a través de *ConcQuest*.

*CLUVI* es una interfaz que permite la consulta de concordancias de una serie de corpus paralelos alineados bilingües o multilingües (en gallego, catalán, euskera, español, francés, inglés, portugués), de diversas temáticas (corpus jurídicos, literarios, económicos, de localización, etc.). Permite una consulta a partir de cualquiera de las lenguas del corpus. Aparte de la búsqueda de palabras o frases exactas, también es posible efectuar consultas más complejas por medio de expresiones regulares.

*COMPARA* (Frankenberg-García 2002; Frankenberg-García y Santos 2003) es un corpus paralelo alineado y bidireccional de inglés y portugués, disponible en línea. Para consultar este corpus también se requiere el conocimiento del lenguaje de interrogación

utilizado por el IMS Corpus Workbench. La interfaz permite, además, restringir los resultados a variantes lingüísticas de inglés y portugués, fecha, autor, etc. También pueden modificarse los formatos de las concordancias, por ejemplo, pueden visualizarse las propiedades de la alineación o las etiquetas de categoría gramatical.

*EUR-Lex* es un corpus paralelo no alineado que contiene traducciones, a las 23 lenguas oficiales, de la legislación de la Unión Europea y otros documentos considerados de carácter público. Una vez seleccionado el idioma de consulta, la búsqueda simple permite consultar palabras en el título o en el texto de los documentos, número de documento, fecha, referencia del DO, código CELEX y otros. La búsqueda avanzada permite combinar todos estos criterios de restricción. Además, puede especificarse que estén presentes todas las palabras con el operador AND, alguna de las palabras con OR, o descartar palabras con NOT. El carácter (+) sustituye a 0 o más caracteres y (?) a 0 o 1. El uso del operador de proximidad (&se) entre palabras indica que dichas palabras pueden aparecer a una distancia máxima de 10 palabras y con (&pa) a una distancia máxima de 40. Los resultados pueden ordenarse por distintos criterios como la fecha o que se muestren todas las lenguas y formatos disponibles del documento. Esto permite acceder a todas las versiones de un documento en distintas lenguas.

*TAUS Search* es una herramienta de búsqueda disponible en línea que permite la consulta de un conjunto de memorias de traducción. Es necesario indicar una palabra o frase de búsqueda y la combinación de idiomas que se quiere consultar en las memorias. Asimismo, es posible restringir la búsqueda por *industry* (para seleccionar un área temática de la lista), *owner* (para seleccionar un autor de la memoria), *content type* (se escoge un tipo de documento), *search all word forms* (para lematizar las palabras de búsqueda) y *part of speech* (permite limitar a una categoría gramatical). Existe la posibilidad de subir memorias de traducción o descargarlas (para ello hay que registrarse gratuitamente y ganar puntos subiendo memorias o comprarlos).

### 3.3.3.2 Programas de gestión de corpus

*Corpógrafo* (Sarmiento et al. 2004) es un programa desarrollado por Linguateca, Faculdade de Letras Universidade do Porto. Este programa se encuentra disponible en línea y, para su uso, se precisa solicitar un nombre de usuario y una contraseña. Permite analizar corpus (tanto monolingües como bilingües paralelos) proporcionados por el usuario, y también es posible descargar los textos del corpus desde Internet. Para ello, hay que proporcionar al programa una URL de la que descargar los textos. Se pueden descargar

textos de cualquier formato, o de un formato específico, por ejemplo PDF, PS, HTML, DOC, RTF, TXT.

Una vez descargados los textos de Internet, se pueden anexar a un corpus, que puede ser monolingüe o bilingüe. Esta herramienta también permite alinear los corpus importados desde el ordenador o descargados desde Internet. Respecto a las funciones de análisis, encontramos la generación de concordancias con las que se recuperan frases en las que aparece la expresión de búsqueda o se presentan los resultados en formato KWIC (*Concordância Frase* y *Concordância KWIC*), también se puede especificar una expresión de búsqueda y el número de elementos que pretendemos que se muestren a izquierda y derecha, de manera que en los resultados cada elemento aparece en un cuadro y se pueden ordenar los resultados por cualquiera de ellos (*Concordância Janela*). También es posible generar n-gramas (*Estudo de N-Gramas*), ordenados por frecuencia o alfabéticamente. Asimismo, permite crear bases de datos.

*Jaguar* es un programa de análisis de corpus disponible en línea, que facilita la descarga de textos especializados desde Internet (Nazar et al. 2008). Para utilizarlo es necesario registrarse para recibir un nombre de usuario y una contraseña. La interfaz permite tres modos de recuperación de corpus: cargándolo directamente desde el ordenador (un corpus proporcionado por el usuario); descargándolo desde Internet (se introduce una palabra de búsqueda que indique el dominio del corpus, la lengua del corpus, el formato de los textos, el número total de páginas recuperadas, etc.); descargando un corpus especializado desde Internet con el módulo Wüska. Este módulo permite por medio de una serie de funciones restringir el corpus de Internet a una determinada área temática.

En la primera fase de la búsqueda, el usuario especifica una serie de palabras de búsqueda, el idioma, el número máximo de documentos, etc.

Una vez realizada la recuperación de páginas de Internet, el programa genera un grafo con terminología extraída del corpus recuperado. Los términos del corpus aparecen interconectados en el grafo. Por ejemplo, en la recuperación de un corpus sobre *asma bronquial*, el programa genera un grafo con formas como *sibilancias*, *alergia*, *pulmonar*, *rinitis*, etc.

El usuario puede pinchar en cada una de las formas del grafo para observar concordancias en las que dicha palabra aparece en el mismo contexto que *asma bronquial* en el corpus. Una vez observados los contextos, el usuario puede decidir si cada forma constituye un término del ámbito de especialidad del corpus que desea compilar y puede decidir *incluirla* o *excluirla* de la lista de palabras clave de dominio. Una vez incluidas y

excluidas las formas que el usuario considere, puede lanzarse una nueva búsqueda con los términos seleccionados. La nueva búsqueda generará a su vez otro grafo, que permitirá limitar más el corpus. Este proceso puede repetirse tantas veces como sea necesario.

En cuanto al análisis de los corpus, es posible generar concordancias determinando una o dos expresiones de búsqueda combinadas (es posible utilizar expresiones regulares en las consultas). También pueden generarse listas de n-gramas (de hasta 10 palabras) con las siguientes opciones: que incluyan un componente, que se ignoren los números, que contengan palabras con una extensión mínima o máxima de caracteres, con una determinada frecuencia en el corpus, especificando un idioma, utilizando una *stoplist*. La opción *distribution* obtiene listas de palabras en función de su distribución en el corpus. La opción *similarity* permite generar listas de términos similares (morfológicamente) a una palabra introducida, es decir, variantes de género, número o derivados, y puede comprobarse la similitud de los documentos de un corpus. Por último, es posible exportar el corpus recuperado para su posterior análisis con otra herramienta de gestión de corpus.

*Sketch Engine* (Kilgarriff et al. 2004) es un programa para el análisis de corpus, que se encuentra disponible en línea. Es posible solicitar un nombre de usuario y una contraseña para utilizar la interfaz durante 30 días. Incorpora una aplicación denominada *WebBootCat*, que permite descargar textos directamente desde Internet para analizarlos inmediatamente con el uso de *Sketch Engine*.

En la interfaz de *WebBootCat*, para la descarga de un corpus de Internet, el usuario debe indicar una referencia que identificará al corpus y la lengua del corpus. Para la descarga de los textos de Internet, el usuario puede especificar una serie de direcciones, o también puede optar por introducir palabras de búsqueda (*seed words*). Por ejemplo, si queremos compilar un corpus en español sobre cerámica, podemos introducir los términos *cerámica*, *baldosa*, *azulejo* y *arcilla*. El programa lanza a *Yahoo!* una serie de búsquedas combinando las palabras introducidas y, como resultado, recupera páginas que el usuario puede aceptar o descartar para el corpus. Algunas opciones de búsqueda avanzada permiten restringir las páginas del corpus a un formato determinado, limitar el número de páginas recuperadas, especificar el mínimo y máximo de palabras o el tamaño de las páginas que se van a recuperar, y añadir además palabras que deberán estar presentes o no en las páginas que se recuperen.

Una vez compilado el corpus, es posible su análisis inmediato con el programa *Sketch Engine*. Este programa presenta las funciones principales de un programa de

análisis textual: generación de concordancias, listas de palabras, listas de palabras clave, listas de agrupaciones de palabras, colocaciones, etc.

Además este programa contiene otras tres funciones más novedosas: *Word Sketch*, *Thesaurus* y *Sketch-Diff*. Estas opciones no están disponibles para la consulta en corpus compilados por el usuario con *WebBootCat*, pero sí que pueden utilizarse para consultar una lista de corpus que se proporcionan en el sistema (la mayoría descargados de la red, pero anotados lingüísticamente por los desarrolladores de la herramienta). Son corpus en diferentes idiomas: árabe, chino, holandés, francés, rumano, ruso, español, sueco, etc. Por ejemplo, uno de los corpus proporcionados para la consulta con *Sketch Engine* es el *British National Corpus*.

La opción *Word Sketch* genera listados de palabras que presentan relaciones sintácticas con una palabra de búsqueda (*object of*, *subject of*, *n modifier*, *modifies*, *y/o*). Por ejemplo, con el corpus en español ofrecido por el programa (*Spanish web corpus*), generamos un *Word Sketch* de la palabra *colegio*. En la lista de la relación *object of*, se recuperan verbos cuyo objeto directo es *colegio*. En la lista *subject of* aparecen verbos cuyo sujeto es *colegio*. La lista *n modifier* contiene adjetivos y nombres que modifican a *colegio*, como por ejemplo *abogados* (*colegio de abogados*) o *veterinario* (*colegio veterinario*). En la lista *modifies* aparecen palabras modificadas por la palabra *colegio*, como *patio* (*patio de colegio*). Por último la lista *y/o* contiene palabras que aparecen coordinadas con *colegio* por medio de las conjunciones *y* y *o*, como *universidad* (*colegio o universidad*). Además, en cada una de las palabras de la lista podemos acceder a los contextos en los que aparece combinada con la palabra *colegio*.

La opción *Thesaurus* recupera palabras con relaciones sintácticas y colocacionales similares a las de la palabra introducida. Por ejemplo, la introducción de *colegio* genera la lista de palabras: *escuela*, *hospital*, *academia*, *instituto*, *asociación*, *biblioteca*, etc. porque todas estas palabras suelen tener las mismas colocaciones y aparecen en las mismas construcciones sintácticas que *colegio*. Pinchando en cada una de estas palabras podemos observar una comparativa de las relaciones colocacionales y sintácticas de dicha palabra y de *colegio* (véase también la función *Sketch-Diff*).

La función *Sketch-Diff* permite comparar las relaciones sintácticas y colocacionales de dos palabras introducidas, por ejemplo, si introducimos las palabras *escuela* y *colegio*, observamos que ambas palabras aparecen en los contextos del corpus combinadas con los mismos verbos y modificando o siendo modificadas por los mismos sustantivos y adjetivos. Todas las palabras de las listas aparecen combinadas con ambas palabras de

búsqueda. Los colores muestran si, además de que una palabra coaparece con ambas palabras de búsqueda, lo hace con frecuencia similar en ambas palabras (color amarillo) o con frecuencia diferente en cada palabra, en cuyo caso, si la palabra coaparece en mayor medida con *colegio*, se destaca en color verde y, si la palabra coaparece con más frecuencia con *escuela*, se resalta en color rojo. Por ejemplo, en los contextos del corpus aparece con mayor frecuencia *escuela primaria* o *escuela secundaria*, que *colegio primario* o *colegio secundario*; pero es más frecuente la combinación *colegio profesional* que *escuela profesional*.

**TerminoWeb** (versión 2.0) es una herramienta desarrollada por Barrière (2009), en colaboración con NRC Institute for Information Technologies. Es un programa que puede utilizarse en línea, pero se precisa un nombre de usuario y una contraseña, que hay que solicitar. Permite la selección de un corpus y también la descarga de un corpus de textos especializados de Internet, en inglés y francés, para la búsqueda de concordancias, colocaciones y patrones semánticos. Para la obtención o creación del corpus, *TerminoWeb* dispone de varios métodos.

Una de las opciones es partir de un texto en inglés o francés, por ejemplo el texto que se quiere traducir o texto de origen. Este texto se introduce en la herramienta. *TerminoWeb* realiza una extracción automática de terminología de ese texto. Podemos especificar que se extraigan términos simples o compuestos, o ambos. El usuario, manualmente, puede validar o descartar los términos extraídos.

A continuación, se lanzan búsquedas a la Web utilizando los términos validados de la extracción como palabras clave, y el usuario puede establecer también palabras clave de dominio. Para realizar las búsquedas en Internet, *TerminoWeb* utiliza la API de Java de *Yahoo!*.

Otra opción de compilación del corpus consiste en que el usuario introduce manualmente las palabras clave y las palabras de dominio que se buscarán en Internet. También se puede realizar una búsqueda simple, con una sola palabra de búsqueda, así como descargar los textos de una única página web introducida por el usuario.

Los textos que se recuperan de estas búsquedas en Internet forman el corpus. La herramienta se queda con los 10 textos más *informativos*. Para elegir los textos más informativos, el programa tiene en cuenta una serie de parámetros, la especificidad temática, la presencia de patrones definicionales como patrones de conocimiento, el tamaño del texto, longitud de las frases, etc. Por ejemplo, la herramienta tiene en cuenta que en los textos aparezcan patrones de hiperonimia *is a kind of*, o de sinonimia *also*



*known as*. Asimismo, el usuario tiene acceso al link de cada página web, con lo que puede acceder a ellas y decidir si las quiere incluir en el corpus o no.

Una vez compilado el corpus, puede analizarse con *TerminoWeb*. A partir de una extracción terminológica del corpus compilado, podemos seleccionar un término o una serie de términos y generar concordancias. También es posible buscar colocaciones, y concordancias de las colocaciones. Otra opción, denominada *Term pair exploration*, busca concordancias en las que dos términos seleccionados por el usuario aparecen en un contexto cercano. Por último, es posible buscar concordancias utilizando patrones lingüísticos que indican relaciones entre los términos. Los patrones disponibles en *TerminoWeb* por defecto son de antonimia, causa, definición, función, hiperonimia, meronimia, similitud y sinonimia. Asimismo, es posible que el usuario defina nuevos tipos de relaciones y que introduzca patrones nuevos para las relaciones que ya existen y para las nuevas relaciones que haya creado.

*WordSmith Tools 4*<sup>46</sup> es otro programa de análisis textual para la gestión de corpus elaborado por Scott y comercializado por Oxford University Press. Esta herramienta se ha utilizado en numerosas investigaciones con corpus (Faber et al. 2001; Pérez Hernández 2002; Sánchez-Gijón 2003a). Para utilizar este programa, el usuario puede proporcionar a la herramienta un corpus propio. Asimismo, si el usuario no dispone de un corpus propio para el análisis, también puede descargar los textos de Internet directamente desde la interfaz del programa para utilizarlos en las búsquedas. En la interfaz de selección de textos, pinchamos en el botón con el icono del globo terráqueo para acceder a la pantalla de configuración de descarga de textos. Podemos configurar nuestra búsqueda de textos en Internet, seleccionar el buscador que queremos que *WordSmith* utilice para recuperar los textos (*Google*, *Altavista*, *Excite* o *Yahoo*), así como la lengua de los textos, por ejemplo español. También introducimos las palabras que queremos que estén presentes en los textos recuperados. Una vez descargados los textos desde Internet, el usuario puede acceder a ellos y cargarlos para la consulta.

Las tres funciones principales de análisis de corpus de este programa son Concord, Wordlist y Keywords.

**Concord** se utiliza para la búsqueda de concordancias. El programa muestra los resultados en un formato KWiC. Además, los contextos pueden ordenarse según su lugar

---

<sup>46</sup> Actualmente se encuentra disponible la versión 6 de este programa, a la que no hemos tenido acceso para el análisis.

de aparición en el corpus o por orden alfabético en función de la palabra situada a derecha o izquierda de la palabra pivote.

También es posible utilizar comodines, por ejemplo para buscar las variantes de un mismo lexema (con la búsqueda de *prens\**, se recupera *prensa*, *prensar*, *prensado*, etc.). *WordSmith* utiliza los siguientes caracteres de truncamiento: el comodín (\*), que sustituye a una cadena de caracteres. Puede utilizarse a principio o final de palabra, o entre palabras, por ejemplo *molienda \* vía* recupera *molienda por vía*. Otros comodines son: ? (que sustituye a un único carácter), # (que sustituye a un número del 0 al 9) y ^ (que sustituye a una letra del alfabeto).

Asimismo pueden realizarse búsquedas más complejas de un término o patrón de búsqueda que aparezca a una determinada distancia de otro término o patrón de búsqueda (posiciones a izquierda o derecha de la palabra pivote). Por ejemplo, si buscamos en un corpus sobre cerámica<sup>47</sup> en español *mol\** a una distancia de cinco posiciones a izquierda y derecha de *sec\**, entre las concordancias generadas encontramos *molienda por vía seca*, *molienda en seco* o *molturación en seco*.

Otra función es la de **Collocates**, o búsqueda de colocaciones. *WordSmith* puede ordenar las colocaciones alfabéticamente o por frecuencia de aparición en el corpus. Además se muestra la posición en la que aparece cada colocación respecto de la palabra de búsqueda. Siguiendo el ejemplo de búsqueda anterior, podemos acceder a las colocaciones del adjetivo *seco* en el corpus de la cerámica. Ordenando la lista alfabéticamente, las colocaciones que empiezan por *mol* son *molido*, *molienda*, *molturación*, *molturado*. Con un doble clic en las colocaciones de la lista accedemos a los contextos. Otra opción es la de **Cluster**, que se emplea para detectar agrupaciones de palabras que aparecen con frecuencia en el corpus. Por ejemplo, podemos generar agrupaciones de tres términos que incluyan la palabra *naranja*. Una de las agrupaciones es *piel de naranja*.

**WordList** genera listados de palabras en función de su frecuencia de aparición en el corpus. Con el uso de una *stoplist* se reduce el número de palabras vacías de significado que aparecen en la lista. Los listados se ordenan alfabéticamente o por frecuencia de aparición en el corpus. A partir de una palabra de la lista se puede acceder directamente a la herramienta Concord para la búsqueda de sus concordancias. **KeyWords** extrae listas de palabras clave mediante la comparación estadística de dos listas de palabras pertenecientes

---

<sup>47</sup> Por ejemplo en el corpus txtCeram 2.0, desarrollado por el grupo de investigación TecnoLeTTra.

a corpus de especialidades distintas (uno especializado y otro de carácter general o corpus de referencia), teniendo en cuenta los segmentos o palabras que no tienen en común. Por ejemplo, comparando la lista de palabras de un corpus de la cerámica y una lista de palabras de un corpus de referencia, se genera una lista de palabras específicas del ámbito de la cerámica.

Existen otros programas de análisis textual además de los ya comentados, tanto privativos como libres, que incorporan técnicas de búsqueda similares a las ya mencionadas, entre ellos *AntConc*, *LEXA*, *MicroConcord*, *MonoConc*, *ParaConc*, *Simple Concordance Program*, *TACT* y *TACTweb*, *TextSTAT* y *WordCruncher*.

### 3.3.3.3 Programas de traducción asistida por ordenador (TAO)

Los programas de traducción asistida por ordenador cuentan con una estructura modular que incluye principalmente un editor de traducción bilingüe, un gestor de memorias de traducción y un gestor de bases de datos terminológicas. Una memoria de traducción almacena el contenido traducido anteriormente de una forma organizada. Normalmente se procede a la segmentación de los textos originales y traducidos en unidades de traducción (párrafos, oraciones o frases) que se alinean para formar pares de unidades de traducción. Estas unidades antiguas pueden recuperarse cuando se detecta una coincidencia entre las unidades de traducción nuevas y las que hay almacenadas en la memoria.

Son muchos los programas de traducción asistida disponibles en el mercado, algunos privativos, entre ellos *SDL Trados*, *Déjà Vu*, *Transit*, *Wordfast*, etc., otros de software libre o código abierto<sup>48</sup>, como *OmegaT*, *Open Language Tools*, *ForeignDesk* o *Transolution*, entre otros.

Algunos cuentan con un editor de traducción independiente (*Déjà Vu*, *Transit*, *OmegaT*, *MemoQ*, *Across*, *SDL Trados Studio* versiones 2009 y 2011) y otros precisan de la interacción con editores externos, principalmente *Word* (*Wordfast*, *Metatexis*, *Trados* hasta la versión 2007).

---

<sup>48</sup> Véase Flórez y Alcina (2011) donde presentan un catálogo con herramientas de software libre y código abierto entre las que se incluyen programas de traducción asistida y programas de gestión de corpus, entre otras herramientas de utilidad para el traductor.

Según lo que hemos visto, las memorias de traducción<sup>49</sup> podrían considerarse corpus bilingües paralelos porque contienen textos equivalentes en más de una lengua, y además están alineados. Cuando el traductor se dispone a traducir un segmento de traducción, el programa busca en la memoria segmentos que se asemejen al que se está traduciendo. Los segmentos de la memoria pueden tener una correspondencia exacta con la frase que se traduce, o una correspondencia parcial, desde algunas palabras iguales, sintagmas completos iguales, etc. Al encontrar el programa correspondencias en la memoria, facilita al traductor de forma automática los segmentos equivalentes que podrían resultar útiles en su traducción. También existen funciones que insertan directamente los segmentos equivalentes en la traducción y el traductor solo debe decidir si utiliza o no las traducciones proporcionadas por el programa. En estas funciones que acabamos de describir, la tarea de búsqueda terminológica la realiza la propia herramienta, casi automáticamente.

Según Lagoudaki (2006: 3), existen diferentes técnicas de detección y recuperación de coincidencias en la memoria. La primera es la **coincidencia basada en secuencias de caracteres**. Esta es la técnica que emplean la mayoría de programas de traducción asistida. En este sentido, se comparan las secuencias de caracteres de los segmentos de traducción nuevos con los almacenados en la memoria. Con ello, es posible detectar coincidencias exactas o parciales. En el caso de que no existan segmentos completos exactos, el sistema busca coincidencias de caracteres dentro de los segmentos. Algunos programas incorporan técnicas de traducción automática basada en ejemplos gracias a las cuales son capaces de fusionar dos segmentos parcialmente coincidentes y formar un nuevo segmento.

La segunda técnica es la **coincidencia basada en análisis lingüísticos**. Las herramientas que la incorporan suelen considerarse la segunda generación de memorias de traducción. Estos sistemas cuentan con lexicones y programas de análisis gramatical. Detectan los patrones gramaticales dentro de los segmentos y tienen en cuenta los

---

<sup>49</sup> Existen memorias de traducción disponibles en línea para su descarga, proporcionadas por organismos internacionales y públicos, plataformas de software libre, etc. A modo de ejemplo, véanse las siguientes:

- Memoria de traducción de la DGT, proporcionada por la Comisión Europea: <http://langtech.jrc.it/DGT-TM.html>
- Memoria de traducción de la ONU: <http://www.uncorpora.org/>
- Memoria del Instituto Vasco de Administración Pública: [http://opendata.euskadi.net/w79-contdata/es/contenidos/ds\\_recursos\\_linguisticos/memorias\\_traduccion/es\\_izo/memorias\\_traduccion\\_izo.html](http://opendata.euskadi.net/w79-contdata/es/contenidos/ds_recursos_linguisticos/memorias_traduccion/es_izo/memorias_traduccion_izo.html)
- Memoria de Softcatalà: [http://www.softcatala.org/wiki/Rebost:Mem%C3%B2ria\\_de\\_traducci%C3%B3\\_de\\_Softcatal%C3%A0](http://www.softcatala.org/wiki/Rebost:Mem%C3%B2ria_de_traducci%C3%B3_de_Softcatal%C3%A0)

vocabularios internos y las reglas gramaticales para proporcionar sugerencias de traducción. Además son capaces de aprender nuevos patrones gramaticales y vocabulario utilizados por el traductor.

A modo de ejemplo, *Similis* (Planas 2005) cuenta con un vocabulario monolingüe de las lenguas alemán, inglés, español, francés, italiano, neerlandés y portugués, e incorpora un analizador gramatical con el que detecta las categorías gramaticales de las palabras de un segmento. Así, puede dividir las oraciones en unidades sintácticas que denomina *chunks* y que consisten principalmente en sintagmas nominales y verbales, con lo que es capaz de generar de manera automática glosarios técnicos. De esta forma, la herramienta puede detectar segmentos completos coincidentes, pero también unidades o *chunks* coincidentes.

Asimismo, el método denominado *Translation Intelligence* (Grönroos 2005) se basa en lo que su autor denomina la segmentación flexible, de manera que cada segmento se divide también en partes más pequeñas y se emplean tres métodos diferentes. El método de reciclado de traducciones funciona igual que los sistemas tradicionales y reutiliza segmentos anteriores de traducción. El método de generación de traducciones convierte los segmentos divididos en patrones gramaticales de traducción y sugiere traducciones a partir de segmentos que tienen la misma estructura gramatical, utilizando también un lexicón o vocabulario y un generador morfológico. Como última opción, si el sistema no tiene ninguna sugerencia de traducción, aplica una traducción automática. El programa es capaz de aprender terminología, unidades y patrones de traducción, formas de segmentación y reglas ortográficas. De momento este sistema se ha desarrollado para el par de lenguas inglés-finlandés.

En *Déjà Vu X2*, el sistema *Deepminer*, por medio de cálculos matemáticos, es capaz de convertir coincidencias parciales en coincidencias exactas a partir de la combinación de traducciones de más de un segmento de la memoria. Por ejemplo<sup>50</sup>, imaginemos que la memoria tiene almacenados los siguientes segmentos: *The cat is black and lazy>El gato es negro y perezoso; The fish is brown and slow>El pez es marrón y lento; The eyes of the cat are brown>Los ojos del gato son marrones*. Así, dado el siguiente segmento de traducción *The cat is brown and lazy*, la memoria propone la coincidencia exacta *El gato es marrón y perezoso* (aunque ninguno de los segmentos de la memoria es exacto al segmento para traducir). Esta coincidencia se obtiene a partir de cálculos realizados por la herramienta en

---

<sup>50</sup> Ejemplos extraídos del manual de usuario de *Déjà Vu X2 Professional* y adaptados al español.

función de traducciones de más de un segmento de la memoria. El segmento de la memoria que más se asemeja al nuevo segmento para traducir es una coincidencia parcial: *El gato es negro y perezoso* (solo hay una palabra diferente: *black* en la memoria y *brown* en el segmento para traducir). Sin embargo, el programa ha sido capaz de indagar que *brown* equivale a *marrón* porque hay otros dos segmentos en la memoria que contienen dicho equivalente. En versiones anteriores de *Déjà Vu*, estas asociaciones solo se realizaban si la herramienta encontraba equivalentes en la base de datos terminológica. Ahora es capaz de obtener dichos equivalentes de los segmentos de la memoria.

Según Planas (2005), el siguiente paso serán herramientas que permitan análisis sintácticos complejos, que resuelvan por ejemplo los problemas de anáforas, y también análisis semánticos que permitan a los programas distinguir correctamente entre distintos significados cuando exista más de una opción de traducción, lo que solo será posible si el sistema cuenta con lexicones semánticos u ontologías del campo que se esté traduciendo.

Lagoudaki (2006: 3) señala que aunque los sistemas basados en análisis lingüísticos en un principio están pensados para mejorar los resultados de detección de coincidencias en las memorias, tienen el inconveniente de que solo pueden trabajar con aquellas lenguas de las que cuentan con recursos lingüísticos (lexicones, analizadores gramaticales, etc.). Entre tanto, los sistemas basados en secuencias de caracteres pueden funcionar para cualquier par de lenguas.

Asimismo, es posible que al traductor le interese buscar manualmente en una memoria de traducción a modo de corpus bilingüe paralelo y alineado, ya que la recuperación de traducciones de la memoria depende mucho del porcentaje de coincidencia del texto origen del traductor con los segmentos de la memoria. Si la coincidencia es mínima, por ejemplo de una palabra, la memoria normalmente no ofrecerá un equivalente de traducción (dependiendo de nuestra configuración).

Por ejemplo, en *SDL Trados 2007*, se puede buscar en la memoria con la función de *Workbench, Tools > Concordance*. Esta función de búsqueda únicamente permite buscar una forma exacta o secuencia de formas en la memoria.

La versión *SDL Trados Studio 2009* presenta mejoras en la búsqueda respecto de la versión anterior. La opción *Enable character-based concordance search* recupera palabras por similitud ortográfica (se basa en la similitud de caracteres entre palabras) en el caso de que no se haya encontrado una palabra exacta en la memoria.

*Déjà Vu X2* ofrece la función *Scan* para buscar en la memoria y admite el uso de comodines. Admite los siguientes comodines: \* (cero o más caracteres), ? (un carácter), #

(un dígito), [a-m] (un carácter en el rango que se especifica), [!a-m] (un carácter que no se encuentre en el rango que se especifica).

El programa *OmegaT* permite realizar en la memoria una búsqueda exacta de una palabra o secuencia, o una búsqueda por palabras clave, lo que supone que se buscan segmentos que contengan todas las palabras introducidas por el traductor. Es posible utilizar expresiones regulares. El programa puede distinguir el uso de mayúsculas, y también es posible utilizar comodines (\* o ?).

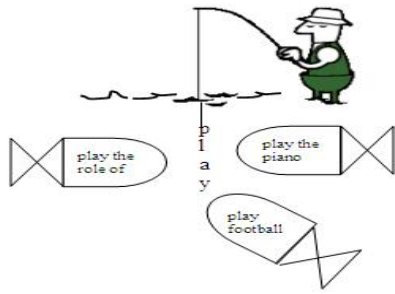
Normalmente las funciones de búsqueda manual de estos programas de traducción asistida son más limitadas que las de la mayoría de programas de gestión de corpus, aunque se están produciendo avances para que las búsquedas se asemejen a las de estos programas (Sánchez-Gijón 2009: 113). Así y todo, las memorias de traducción son los corpus paralelos alineados más utilizados por los traductores.

### 3.3.4 Clasificación de técnicas de búsqueda en corpus

Entendemos por técnica de búsqueda en corpus a las vías, herramientas u opciones de consulta que el corpus nos permite realizar para obtener un resultado. Hemos visto que en una técnica de búsqueda intervienen tres elementos: una consulta, un instrumento y un resultado.

La *consulta* es la palabra o expresión que introduce el usuario en el corpus. El *instrumento* es el corpus o la parte del corpus en el que se busca la palabra o expresión de búsqueda. El *resultado* de la búsqueda es el elemento al que se accede cuando se consulta el corpus.

En este apartado presentamos una clasificación de técnicas de búsqueda en corpus que se centra en la consulta, el instrumento y el resultado. Un ejemplo de técnica de búsqueda en corpus, teniendo en cuenta los tres elementos de nuestra clasificación, podría ser que el usuario introdujera como consulta una palabra exacta, por ejemplo *play*, en la interfaz de un corpus monolingüe inglés, que sería el instrumento, y obtuviera, como resultado, una lista de concordancias de la palabra *play*, entre las que se incluirán expresiones como *play the piano*, *play football* o *play the role of*.



**Ilustración 49. Representación de técnica de búsqueda en un corpus.**

A continuación explicamos con más detalle todas las técnicas de búsqueda que pueden emplearse en un corpus, clasificadas en función de la consulta, el instrumento y el resultado. Asimismo, aportamos ejemplos concretos de cómo se implementan estas técnicas de búsqueda en los distintos corpus que hemos analizado.

### 3.3.4.1 El primer elemento: la consulta

La consulta es la expresión que introduce el usuario en la interfaz de un corpus cuando realiza una búsqueda. Como veremos en este apartado, las consultas pueden ser 1) una expresión léxica, 2) una expresión gramatical, 3) un número, 4) una combinación de expresiones (léxicas, gramaticales, números o expresiones mixtas), con criterios de presencia de todas, alguna o ausencia de expresiones y con criterios de combinación continua o discontinua de las expresiones. También es posible utilizar filtros para restringir las consultas.

#### 3.3.4.1.1 Búsqueda mediante expresiones léxicas

Las expresiones léxicas pueden estar formadas por una forma o un lema, o por una secuencia de dos o más formas o lemas.

Un **lema** es una palabra en su forma canónica, es decir, sin morfemas flexivos. En el caso de un sustantivo, el lema es la palabra sin morfemas de género y número. En el caso de los verbos, el lema es el infinitivo.

Esta consulta resulta útil para acceder a todas las formas del corpus que están etiquetadas con el lema introducido. Por ejemplo, si introducimos en el *BYU-British National Corpus* el verbo *do* como lema, se recuperan todas las formas de dicho verbo que aparecen en el corpus: *do*, *did*, *does*, *done*, *doing*, etc.

Este tipo de consulta amplía los resultados de la búsqueda, de manera que se recuperan por ejemplo tanto las formas en singular como en plural de una palabra. En un corpus lematizado, como el corpus de la informática en español en *Bwananet*, podemos



introducir *memoria* como lema seguida de un adjetivo, y en las concordancias recuperamos tipos de memoria en español, como *memoria asociativa*, *memorias asociativas* (en singular y plural), *memorias auxiliares* (solo en plural), y *memoria principal* y *memoria masiva* (solo en singular). Si hubiéramos introducido la forma *memoria* en singular, no hubiéramos recuperado *memorias auxiliares*, que aparece en el corpus solo en plural.

Las **formas** pueden ser exactas o truncadas. Una **forma exacta** es una palabra completa. Puede ser útil si se quiere encontrar en el corpus una forma concreta. Todos los corpus que hemos analizado permiten introducir formas exactas. Por ejemplo, podemos introducir la forma exacta *Government* en el *BNC* para analizar sus contextos.

Una **forma truncada** es una palabra incompleta. La parte de la palabra que se omite, se sustituye con un carácter comodín. Los comodines más utilizados son el asterisco (\*) y el signo de interrogación (?). Puede ser útil si se quieren buscar todas las palabras que contengan, empiecen o terminen con una secuencia de caracteres. Por ejemplo, si introducimos en el corpus *COMPARA* la forma truncada “*hous.\**”, se recuperan formas completas: *house*, *housewife*, *housekeeper*, *house-doctor*, *houses*, *housing*, *household*, etc.

#### 3.3.4.1.2 Búsqueda mediante expresiones gramaticales

Las expresiones gramaticales son expresiones formadas por categorías gramaticales. Pueden ser útiles para localizar palabras o secuencias de palabras a partir de sus categorías gramaticales.

Este tipo de búsqueda puede realizarse en *Bwananet*, en los corpus de *BYU*, *COMPARA* y *HNC*. Por ejemplo, si introducimos en el corpus inglés de derecho de *Bwananet* la expresión gramatical *adjetivo nombre*, se recuperan expresiones como: *commercial legislation*, *fiscal protection*, *Social Fund*, etc. También en *Bwananet* pueden buscarse secuencias variables de categorías gramaticales. Por ejemplo, podemos buscar en el corpus de la informática en español de *Bwananet* secuencias formadas por 1 a 3 sustantivos seguidos de un adjetivo. Así recuperamos tanto términos simples como compuestos: *gestión documental*, *discos ópticos*, *protocolos gateway internos*, *máquina virtual*, etc.

#### 3.3.4.1.3 Búsqueda mediante un número

Los números pueden introducirse de forma exacta o truncada. En el caso de introducir un **número exacto**, este se busca en el corpus tal cual se ha introducido. Un **número truncado** es un número combinado con un comodín. En este caso, se busca en el

corpus todos los números que contienen la secuencia de números que acompaña al comodín. La búsqueda de números puede ser útil para localizar palabras que suelen aparecer en un contexto próximo a algún número significativo. Por ejemplo, si introducimos en el corpus de la informática en español de *Bwananet* el número 640, podemos localizar palabras como *píxel*, porque una medida que recupera el corpus es *640x480 píxels*; también se localiza *memoria RAM*, porque otra medida que se recupera es *640 Kb de memoria RAM*.

#### 3.3.4.1.4 Búsqueda mediante dos o más expresiones

La consulta mediante **dos o más expresiones** en un corpus puede estar formada por **lemas, formas exactas o truncadas, expresiones gramaticales, números y expresiones mixtas** (combinación de palabras, categorías gramaticales y números). Normalmente la interfaz de búsqueda del corpus busca aquellos contextos en los que estén presentes **todas las expresiones** introducidas, aunque también puede especificarse que solo estén presentes **algunas de las expresiones** o la **ausencia de expresiones**. Además, cuando se consultan dos o más expresiones es necesario indicar un criterio de combinación, que puede ser continua o discontinua.

##### 3.3.4.1.4.1 Búsqueda mediante una combinación continua de expresiones

Una **combinación continua de expresiones** es una consulta en la que todos sus componentes deben aparecer en el corpus en el orden que se han introducido. Puede servir cuando se quiere acceder al contexto de una expresión fija. En el siguiente ejemplo consultamos la combinación continua de formas *Cámara de los Comunes* en el corpus *CREA* para saber si el nombre de esta institución en español se encuentra extendido en los textos españoles. En los resultados, *Cámara de los Comunes* aparece 140 veces en 104 textos del corpus *CREA*. En la Ilustración 50 observamos algunos contextos del corpus en los que comprobamos que la denominación *Cámara de los Comunes* se encuentra extendida en español y no se añaden explicaciones o aclaraciones en los contextos recuperados.

Concordancias (RAE)

Consulta:	Cámara de los Comunes, en todos los medios, en CREA
Resultado:	140 casos en 104 documentos.

OBTENCIÓN DE EJEMPLOS

Concordancias. ▾ Normal. ▾
Clasificación: ▾

Agrupación: ▾
Marcas: ▾

Cómo citar el CORPUS

Concordancias.

Pantalla: 1 de 6. [Siguiente](#) 1 2 3 4 5 6 [Ver párrafos](#)

Nº	CONCORDANCIA	AÑO
1	te polémica suscitada en el país, anunció ante la Cámara de los Comunes que se va a crear una comisión	** 2001
2	responsabilidades penales. A continuación, en la Cámara de los Comunes, prometió a los padres que "aqu	** 2001
3	la sesión de preguntas y respuestas de ayer en la Cámara de los Comunes, respondiendo de manera aguda,	** 1995
4	vor y 314 en contra. LONDRES. (Corresponsal). La Cámara de los Comunes respaldó ayer por el más estrecho	** 1995
5	n individuo. Clarke, en una comparecencia ante la Cámara de los Comunes, donde la oposición laborista h	** 1995
6	bajo que proporcionaba Barings, pero aseguró a la Cámara de los Comunes que "no hay un número significa	** 1995
7	ación Major, sir Patrick Mayhew, compareció en la Cámara de los Comunes para reiterar que Londres no im	** 1995
8	l proceso de paz. La intervención de Mayhew en la Cámara de los Comunes no sirvió sin embargo para aliv	** 1995
9	rebelión de los euroescépticos y la derrota en la Cámara de los Comunes de la subida del impuesto sobre	** 1994
10	rían hasta el punto de llevarse 607 escaños en la Cámara de los Comunes y aniquilar por completo a los	** 1994
11	que se celebró en la oficina que éste posee en la Cámara de los Comunes. Por si no estaba claro, Klaus	** 1994
12	un manifiesto <i>ad hoc</i> ; se creó una comisión en la Cámara de los Comunes de carácter consultivo-, Franci	** 1994
13	jefe del Bloque Quebequés y de la oposición en la Cámara de los Comunes de Ottawa, Lucien Bouchard, de	** 1994
14	a última guerra, Winston Churchill afirmara en la Cámara de los Comunes que "América se halla en la oú	** 1994
15	jo. Clark, que tuvo que abandonar su escaño en la Cámara de los Comunes hace dos años como uno de los i	** 1994
16	ente diputado laborista David Blunket acudía a la Cámara de los Comunes con su lazarillo. Hace ya seis	** 1994
17	o británico del Interior, Michael Howard, ante la Cámara de los Comunes, quien indicó que no se habían	** 1994

Ilustración 50. Consulta de una combinación continua de formas, *Cámara de los Comunes*, en CREA.

Las **expresiones mixtas** que combinan expresiones léxicas, expresiones gramaticales y números pueden ser útiles para localizar una expresión de la que se conoce la forma o el lema de alguna de sus palabras y la categoría gramatical del resto. Por ejemplo, si buscamos en el corpus inglés de derecho de *Bwananet* la expresión *adjetivo law*, se recuperan expresiones como: *organic law*, *civil law*, *common law*, *Federal law*, *budgetary law*, etc.

En el siguiente ejemplo queremos saber qué verbos se utilizan con mayor frecuencia con la palabra *law*. Buscamos en el *Corpus of Contemporary American English* (de BYU) la expresión mixta *[v\*] a law*, para acceder a expresiones compuestas por un verbo seguido de *a law*. En los resultados encontramos: *passed a law*, *pass a law*, *signed a law*, *enacted a law*, *earned a law*, *need a law*, *break a law*, *write a law*, *become a law*, *upheld a law*, etc.

También podemos buscar en *Scientext* la expresión mixta de la categoría gramatical *adjetivo + ADN*, especificando una relación sintáctica: que la consulta 1 *adjetivo* será un adjetivo epíteto de la consulta 2 *ADN*. Como resultado obtenemos concordancias en las que aparecen *duplex ADN*, *ADN nucléaire*, *ADN chromosomique*, *ADN simple*, *ADN total*, etc. Además de ofrecer estos contextos a modo de concordancias KWiC, pinchando en cada una de las concordancias, se puede acceder a un contexto más amplio, donde también puede verse el análisis sintáctico efectuado.

The screenshot shows the '4. Recherche' (Search) section of the Scientext interface. The search criteria are: Mot 1: Catégorie Adjectif (A), Mot 2: Forme ADN. The search relation is 'adjectif épithète de (ADJ)'. The results table shows 18 occurrences, with the 6th occurrence highlighted, showing the phrase 'ADN chromosomique' in red. A detailed view of the 6th occurrence is shown below the table, with 'ADN chromosomique' circled in red. A checkbox 'Voir l'analyse syntaxique' is also visible at the bottom of the detailed view.

N°	Contexte gauche	Occurrence	Contexte droit	Ref. text
1	Les brins d'	ADN issus	de l' échantillon par rétro transcription sont appelés cible complexe	[hin-the-276-hnd
2	Les acides nucléiques sont présents sous forme d'	hétéro duplex ADN	ARN , il est donc nécessaire de dégrader les ARN	[bio-the-276-bod
3	Les acides nucléiques sont présents sous forme d' hétéro	duplex ADN	ARN , il est donc nécessaire de dégrader les ARN	[bio-the-276-bod
4	Environ 80 % de l'	ADN nucléaire	est engagé dans la formation des nucléosomes .	[bio-the-281-intrc
5	, cet ADN constitue les 20 % restants de l'	ADN nucléaire	.	[bio-the-281-intrc
6	sont l' analyse des fragments de restriction de l'	ADN chromosomique	après migration en champ pulsé ( pulsotype ) et l'	[bio-art-197-body
6	sont l' analyse des fragments de restriction de l'	ADN chromosomique	après migration en champ pulsé ( pulsotype ) et l'	[bio-art-197-bod

[bio-art-197-body] Philippe Glaser - Les puces à ADN vont-elles révolutionner l'identification des bactéries ?

L'identification bactérienne nécessite l'isolement de la bactérie sous forme d'une colonie. Cette première étape peut être difficile si l'échantillon biologique n'est pas normalement stérile, comme un prélèvement de gorge ou un prélèvement de selles. Traditionnellement, l'identification de l'espèce se faisait en combinant l'observation microscopique et l'analyse phénotypique, en étudiant la forme et la couleur des colonies et les propriétés métaboliques et enzymatiques. Cette méthode peut être remplacée par une analyse moléculaire soit en utilisant des tests immunologiques, soit sur la base de séquences d'ADN, par exemple après amplification par PCR. Les méthodes de PCR présentent l'avantage de pouvoir être réalisées directement à partir d'un échantillon biologique. Le gène codant pour l'ARN ribosomique 16S est un des marqueurs d'espèce les plus utilisés : il est en effet possible de définir des amorces universelles pour son amplification, et la comparaison de sa séquence avec des banques de données de référence permet de déterminer l'espèce [6]. De la même manière, après l'identification de l'espèce, les premières méthodes de typage bactérien étaient phénotypiques, par comparaison des antigènes de surface (sérotypie) ou analyse de la sensibilité à des bactériophages (lysotypie). Durant ces vingt dernières années, de nombreuses méthodes de typage moléculaire ont été développées [7]. La valeur de ces méthodes dépend de leur pouvoir de discrimination entre les isolats, de leurs facilité et rapidité d'utilisation et de la reproductibilité des résultats entre différents laboratoires. Cette reproductibilité est essentielle pour l'échange des données, leur organisation sous forme de bases de données internationales et la mise en place de systèmes de surveillance épidémiologique performants. Les méthodes considérées aujourd'hui comme les plus fiables sont l'analyse des fragments de restriction de l'ADN chromosomique après migration en champ pulsé (pulsotype) et l'analyse des séquences nucléotidiques de plusieurs gènes de ménage (exprimés dans toutes les cellules) (STML, séquençotypage multilocus, MLST en anglais). Le pulsotype dépend du polymorphisme des sites de reconnaissance par les enzymes de restriction et de l'organisation du génome. Le type MLST dépend uniquement de la vitesse d'évolution nucléotidique au sein de l'espèce et d'éventuels transferts génétiques horizontaux.

Voir l'analyse syntaxique

Ilustración 51. Consulta de una combinación continua de expresiones mixtas en *Scientext*.

Asimismo, existe la posibilidad de utilizar secuencias fijas que representan relaciones entre conceptos, o lo que es lo mismo, **expresiones semánticas**, que en otros trabajos han recibido distintas denominaciones<sup>51</sup>. Algunas de estas expresiones pueden utilizarse, por

<sup>51</sup> Distintos autores se han referido a *patrones*, *patterns* o *knowledge patterns*; *knowledge probes* y *sondas*; *marcadores*, *markers* y *marqueurs*, entre otros. Véanse a modo de ejemplo Marshman et al. (2002: 2) que hacen una recopilación de denominaciones que se han utilizado en la bibliografía: *formulae* (Lyons 1977), *diagnostic frames* o *test frames* (Cruse 1986), *frames* (Winston et al. 1987), *knowledge probes* (Ahmad y Fulford 1992), *definitional meta-language* y *defining expositives* (Pearson 1996b). Marshman et al. (2002: 2) se decantan por *knowledge patterns*, lo mismo que Halskov y Barrière (2008: 21) y Meyer (2001: 290) que distingue entre patrones léxicos, gramaticales y paralingüísticos. En la misma línea, López Rodríguez y Tercedor (2008: 10) se refieren a *linguistic patterns* y también *search structures*. Soler y Alcina (2008) también utilizan *patrones*, al igual que Faber et al. (2001: 181-182) que, como Ahmad y Fulford (1992) hablan de *sondas*. Asimismo, es frecuente el uso de *marcadores*, como por ejemplo *marcadores discursivos de reformulación* (MR) (Bach 2005), *marcadors lingüístics de relacions conceptuals* (MLRC) (Feliu 2004), *marqueurs* (Condamines y Rebeyrolle 2000; Marshman y L'Homme 2008) o *explicit relation*

ejemplo, para localizar contextos definitorios en un corpus (*es un\**, *conocid\* como*, *llamad\**, etc.). Por ejemplo, en un texto médico sobre genoma en inglés, aparece la abreviatura PCR. En el caso de que no encontremos su definición en el diccionario, podríamos buscar una explicación de este término o abreviatura en un corpus de especialidad, por ejemplo en los corpus de medicina y genoma en inglés de *Bwananet*. Introducimos la consulta *PCR is*. Como resultado, obtenemos entre las concordancias contextos definitorios de PCR; algunos los incluimos en la Tabla 15.

The <b>PCR is</b> a relatively straightforward laboratory technique, although because the technique is so versatile and the range of applications so wide, it is difficult to give a “typical” example.
This study also demonstrated that extended <b>PCR is</b> a reliable, rapid method to obtain genomic probes (such as the 12 kb used in this report) for FISH.
<b>PCR is</b> a rapid and versatile in vitro method for amplifying defined target DNA sequences present within a source of DNA.
<b>PCR is</b> an in vitro method for amplifying DNA sequences using defined oligonucleotide primers

**Tabla 15. Contextos definitorios de PCR extraídos del corpus de medicina e informática en inglés de *Bwananet*.**

A su vez, el programa *TerminoWeb* permite acceder concordancias a partir de la selección de relaciones semánticas de antonimia, causa, definición, función, hiperonimia, meronimia, similitud y sinonimia. Asimismo, es posible que el usuario defina nuevos tipos de relaciones. El programa busca por defecto una serie de patrones por cada tipo de relación y el usuario puede introducir patrones nuevos para las relaciones que existen y para las relaciones nuevas que haya incorporado<sup>52</sup>.

### 3.3.4.1.4.2 Búsqueda mediante una combinación discontinua de expresiones (con delimitación de proximidad)

Esta consulta consiste en introducir en el corpus un elemento determinado y establecer la distancia en la que debe aparecer también un segundo (o más) elemento(s). Tanto el primer como el segundo elemento pueden ser cualquiera de las consultas anteriormente explicadas: **lemas, formas exactas o truncadas, expresiones gramaticales,**

---

*markers* (Bowden et al. 1996). Otros autores han trabajado también sobre patrones definitorios y contextos definitorios (Alarcón 2009; Alarcón et al. 2008; Sierra et al. 2010). Además, estos últimos autores están desarrollando un prototipo de herramienta de búsqueda en línea, llamado *Describe*, que permite a los usuarios obtener la descripción de un término de forma resumida y organizada (Sierra et al. 2009).

<sup>52</sup> Véase la Ilustración 59.

**números** o **expresiones mixtas**. Puede ser útil si se quiere buscar una expresión en el corpus y solo se conoce alguna de las palabras que la componen.

En el siguiente ejemplo queremos conocer expresiones en español que se utilicen para expresar que llueve mucho. Introducimos en el *Corpus del español* (de BYU) una combinación discontinua con lema y forma: el verbo *llover* como lema seguido de la preposición *a* y un comodín, *[llover] a \**. El resultado son expresiones como, *llovía a cántaros, llueve a torrentes, lloviendo a mares, lloviendo a raudales, lloviendo a chuzos, llovía a baldes, etc.*

Otro ejemplo de búsqueda con formas truncadas lo vemos en la búsqueda en *WordSmith Tools 4*, donde buscamos en un corpus de la cerámica en español *mol\** en una posición cercana a *sec\** y recuperamos expresiones como *molienda por vía seca* o *molienda en seco*.

En otro ejemplo combinamos de forma discontinua en el *BNC* dos expresiones léxicas y delimitamos que la proximidad entre ellas será de un máximo de 5 posiciones de distancia. Introducimos la forma *Cytomegalovirus* y la expresión semántica *is a*. El resultado son algunos contextos definitorios de la palabra *Cytomegalovirus*.

<b>Cytomegalovirus</b> (CMV) <b>is a</b> virus with many similarities to the herpes virus.
<b>Cytomegalovirus is a</b> less well-known infection which affects considerably greater numbers of babies than rubella.

**Tabla 16. Resultados de la búsqueda de las expresiones *Cytomegalovirus* y *is a* hasta 5 posiciones de distancia en el *BNC*.**

A continuación mostramos un ejemplo en el que combinamos en el *Corpus del español* (de BYU) la forma *metros* a un máximo de 5 posiciones de distancia del número 100. Algunos de los resultados incluyen expresiones como *100 metros libre* o *100 metros cuadrados*.

[...] terminó ayer su participación en Phoenix, Arizona, con el quinto lugar en los <b>100 metros</b> libre.
[...] del lugar, con una área mínima de construcción de 200 <b>metros</b> cuadrados y <b>100 metros</b> cuadrados para parqueo.

**Tabla 17. Resultados de la búsqueda del número 100 hasta 5 posiciones de la forma *metros* en el *Corpus del español* (BYU).**

De los corpus analizados, algunos que permiten combinar discontinuamente más de dos elementos sin necesidad de conocer el lenguaje CQL son el *Hellenic National Corpus*, que permite combinar hasta tres formas, lemas, números o categorías gramaticales a una



distancia máxima de cinco posiciones; o *Scientext* que combina cualquier cantidad de formas, lemas o categorías gramaticales y se pueden especificar relaciones sintácticas entre ellos.

### 3.3.4.1.5 Búsqueda mediante el historial de consultas

Al igual que en los diccionarios, algunos corpus como por ejemplo *BNCweb* permiten acceder al historial de consultas (*query history*) y repetir en el corpus cualquier consulta que se haya efectuado anteriormente.

### 3.3.4.1.6 Filtros

Los filtros añaden una condición de búsqueda a la consulta introducida. Existen, por ejemplo, filtros de **categoría gramatical** como los de *Bwananet*, que permiten buscar palabras que sean o que no sean de una determinada categoría gramatical.

En la Ilustración 52 observamos una búsqueda en el corpus inglés de la informática en *Bwananet* en la que se introduce la palabra *e-mail* restringida a que no sea sustantivo (se activa la opción *negación* debajo de la casilla *categorías*). El resultado es una concordancia en la que *e-mail* es adjetivo: [...] *it can be obtained through CD-ROM, e-mail server, [...]*.

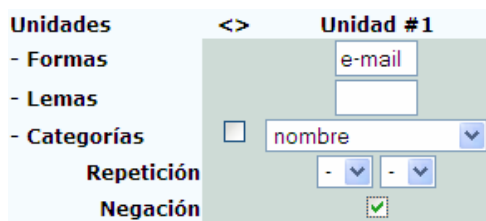


Ilustración 52. Búsqueda de una palabra con filtro de restricción de categoría gramatical en *Bwananet*.

Por otra parte, si buscamos la palabra *e-mail* con la restricción de que sea sustantivo (véase Ilustración 53), el resultado son concordancias de la palabra *e-mail* como sustantivo: *By sending e-mail to clinton-info@campaign92.org, I was able to request press releases on foreign policy; [...] suggestions on how to use everything from e-mail to remote databases, tutorials, lists of frequently asked questions, [...]*.

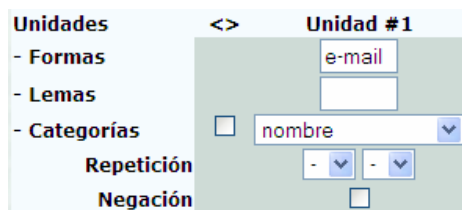


Ilustración 53. Búsqueda de una palabra con filtro de restricción de categoría gramatical en *Bwananet*.

Con los filtros también es posible restringir la consulta a una parte del corpus, por ejemplo con los filtros de área temática, tipo de texto, área geográfica, idioma, autor, fecha y lugar del texto. El filtro de **área temática** restringe la búsqueda a textos del área temática que indiquemos. Esta opción la tienen por ejemplo *Bwananet* (informática, medio ambiente, derecho, medicina, genoma, economía) y *CREA* (ciencias y tecnología, biología, veterinaria, informática, etc.).

El filtro de **tipo de texto** restringe la búsqueda a textos de un determinado canal (oral y escrito, por ejemplo en el *BNCweb*) o género textual (por ejemplo, ficción, prensa o académico, en el *Corpus del español*; académico, administrativo, institucional, etc., en *BNCweb*; libros periódicos, revistas, en *CREA*).

El filtro de **área geográfica** restringe la búsqueda a textos de una determinada zona. Por ejemplo, en el corpus *COMPARA* se puede restringir al portugués de Angola, Portugal, Brasil y Mozambique, o al inglés de Sudáfrica, Reino Unido y Estados Unidos. El *BNCweb* también ofrece este filtro para restringir a textos escritos de autores del Reino Unido o Irlanda, Commonwealth, Europa Continental y Estados Unidos, o textos orales de hablantes cuya lengua materna es el inglés británico, inglés americano, lengua india desconocida, alemán, francés, o desconocido. En corpus multilingües también es posible restringir la búsqueda al corpus de un **idioma** determinado, como en los corpus *COMPARA* (inglés y portugués) o *Bwananet* (inglés, español y catalán).

El filtro de **autor** restringe la búsqueda a textos publicados por uno o varios autores. Este tipo de restricción puede realizarse en el corpus *COMPARA* y en *CREA*. El filtro de **fecha** restringe la búsqueda a textos publicados en un período o fecha. Este tipo de restricción es propia del corpus *COMPARA*, *BNCweb*, *CREA* y los corpus de BYU. El filtro de **lugar del texto** restringe la búsqueda a títulos de los textos, las listas, las tablas, etc. *Bwananet* permite realizar este tipo de restricción (títulos, listas, tablas, resto del texto, cualquiera). El *BNCweb* también permite restringir el corpus a títulos y palabras clave.

#### 3.3.4.2 El segundo elemento: el recurso o instrumento

El instrumento en un corpus siempre es el conjunto de textos que componen dicho corpus. No obstante, los corpus pueden ser de distinto tipo en función de la consulta y el resultado de la búsqueda. En este apartado tratamos los corpus monolingües, los corpus bilingües o multilingües (corpus comparable, corpus paralelo y corpus alineado) y los corpus etiquetados.



### 3.3.4.2.1 Búsqueda en un corpus monolingüe

Un corpus **monolingüe** está compuesto por textos en una sola lengua. En este tipo de corpus la consulta y el resultado serán siempre en una sola lengua. Por ejemplo, en el *Corpus of Contemporary American English* (BYU), un corpus monolingüe en inglés, podrán introducirse consultas en inglés y los resultados que se obtengan serán también en inglés, ya que esta es la única lengua disponible del corpus. Podemos mencionar como ejemplo de corpus monolingües analizados *CREA*, *BNC*, *HNC* y todos los corpus de BYU.

### 3.3.4.2.2 Búsqueda en un corpus bilingüe o multilingüe: corpus comparable, corpus paralelo y corpus alineado

Un corpus también puede ser bilingüe o multilingüe, lo que significa que está compuesto por textos de más de una lengua. Además, como hemos visto un corpus bilingüe puede ser **comparable** (compuesto por textos similares en más de una lengua, por ejemplo del mismo género o la misma temática) o **paralelo** (formado por textos en lengua origen y sus traducciones). Finalmente, los corpus paralelos pueden estar **alineados**. Un corpus alineado es un corpus paralelo en el que cada frase o segmento del texto origen está vinculado o alineado con su traducción en la lengua meta. Por ejemplo, de los corpus paralelos analizados los corpus *Bwananet* y *COMPARA* se encuentran alineados, mientras que *EUR-Lex* no lo está. En los programas de traducción asistida, como *SDL Trados*, *Déjà Vu*, etc., las memorias de traducción también se encuentran alineadas. A continuación vemos ejemplos de búsqueda en los corpus alineados analizados.

En un corpus alineado la consulta se introduce, normalmente, en una sola lengua, al igual que en un corpus monolingüe. Sin embargo, el resultado de la búsqueda incluye los segmentos del idioma en el que aparece la consulta introducida y, además, los segmentos equivalentes en otro idioma o idiomas del corpus. Por ejemplo, en el corpus *COMPARA*, si introducimos como consulta la forma *run* en inglés, el resultado son concordancias en las que la palabra *run* aparece destacada. Junto a las concordancias de la palabra *run* aparecen los segmentos equivalentes en portugués. Las palabras portuguesas equivalentes a *run* (*montar*, *correr* y *comprar*) no aparecen destacadas (véase la Tabla 18).

I said to Nizar, «You could probably <b>run</b> a little rental business [...]»)	Em resposta, sugeri ao Nizar: «Talvez pudesse montar um negócio de aluguer [...]»
We had to <b>run</b> for a train once at Euston: [...]	Uma vez em Euston tivemos de correr para apanharmos um comboio: [...]
Neither of our families could afford to <b>run</b> a car in those far-off days.	Naquele tempo nem a minha família nem a dela tinham dinheiro para comprar carro.

**Tabla 18. Resultados de búsqueda de la consulta *run* en el corpus paralelo alineado *COMPARA*.**

Algunos corpus alineados permiten utilizar consultas al mismo tiempo en los dos idiomas del corpus. En el corpus *COMPARA* existe una opción de restricción de alineamiento con la que se puede buscar una consulta en un idioma y, al mismo tiempo, una consulta en otro idioma. En este ejemplo se introduce como consulta en inglés la forma *run* y en portugués el lema *correr*. El resultado son concordancias en inglés de la palabra *run* y en portugués el lema *correr*. El resultado son concordancias en inglés de la palabra *run* cuyos segmentos equivalentes en portugués contienen el lema *correr*. En las concordancias aparecen destacadas *run* y todas las formas de *correr* (véase la Tabla 19).

We had to <b>run</b> for a train [...]	[...] tivemos de <b>correr</b> para apanharmos um comboio [...]
If they had broken into a <b>run</b> , [...].	Se tivessem desatado a <b>correr</b> , [...].
But I feel we <b>run</b> a grave risk by doing so.	Mas eu acho que <b>corremos</b> um risco grave se o fizermos.

**Tabla 19. Resultados de búsqueda de las consultas de la forma *run* en inglés y el lema *correr* en portugués en el corpus paralelo alineado *COMPARA*.**

Las memorias de traducción también son corpus alineados que, gracias a la ayuda de los programas de traducción asistida, permiten buscar a partir de una de las lenguas de la memoria, para obtener la traducción en otra lengua. Asimismo, los programas de traducción asistida, cuando detectan un porcentaje de similitud determinado (el establecido por el usuario) entre el segmento que se está traduciendo y un segmento almacenado en la memoria, el programa propone automáticamente la traducción almacenada, como posible equivalencia. De esta forma, el proceso de búsqueda en el corpus paralelo alineado se automatiza.

### 3.3.4.2.3 Búsqueda en un corpus etiquetado

Los corpus también pueden estar o no etiquetados a distintos niveles. Los corpus pueden estar anotados a nivel gramatical, sintáctico, semántico y también con información extralingüística. Los corpus **etiquetados gramaticalmente** indican la categoría gramatical de todas sus formas. Las consultas formadas por expresiones gramaticales y los filtros de categoría gramatical solo pueden emplearse en corpus etiquetados gramaticalmente. Algunos corpus etiquetados gramaticalmente son *HNC*, *COMPARA*, *BNC*, los corpus de BYU y los corpus de *Bwananet*.

Algunos corpus etiquetados gramaticalmente también permiten realizar búsquedas **sintácticas**, generando esquemas del comportamiento gramatical y colocacional de las palabras, como en los corpus disponibles en *Sketch Engine* (como veremos a continuación en el apartado de resultados, es posible obtener una lista de palabras relacionadas sintácticamente). Otros corpus como *Sciencetext* además ofrecen un análisis sintáctico de las oraciones del corpus.

Otro tipo de etiquetado es el **semántico**, que permite la realización de búsquedas de relaciones semánticas en el corpus, por ejemplo a través de la sinonimia (como veremos en el apartado de resultados, es posible acceder a sinónimos o palabras relacionadas semánticamente).

Los corpus **lematizados** son aquellos en los que se indica el lema de todas sus formas. Solo pueden introducirse lemas en corpus lematizados. Algunos corpus lematizados que hemos analizado son *HNC*, *COMPARA*, *BNC*, los corpus de BYU, los corpus de *Bwananet* y los corpus disponibles en *Sketch Engine*.

Asimismo, existe un tipo de **etiquetado extralingüístico** que proporciona información sobre los textos que componen el corpus: área temática, tipo de texto, área geográfica, idioma, autor, fecha y lugar del texto. Como hemos visto en el apartado de filtros, este tipo de anotación extralingüística facilita la restricción de la búsqueda.

### 3.3.4.3 El tercer elemento: el resultado

El resultado es el elemento al que se accede en un corpus con una determinada búsqueda. Las interfaces de los corpus electrónicos incluyen opciones con las que se puede escoger el resultado que se desea obtener. En nuestro análisis, hemos detectado los siguientes resultados: 1) lista de concordancias monolingües o bilingües, 2) lista de palabras, 3) lista de agrupaciones de palabras, 4) lista de colocaciones, 5) lista de sinónimos o palabras relacionadas semánticamente, 6) lista de palabras relacionadas

sintácticamente, 7) lista de categorías gramaticales, 8) representación gráfica de las relaciones de una palabra y 9) información estadística.

### 3.3.4.3.1 Lista de concordancias

Las concordancias son contextos en los que aparece la consulta que hemos introducido. La mayoría de los corpus muestra las concordancias en un formato fácil de leer denominado KWIC (key words in context), lo que significa que la consulta aparece en el centro rodeada de su contexto y normalmente destacada. Aunque también hay posibilidad de recuperar las concordancias en oraciones, párrafos o documentos completos, dependiendo del corpus. En función del corpus utilizado en una técnica de búsqueda, las concordancias pueden ser monolingües o bilingües.

#### 3.3.4.3.1.1 Lista de concordancias monolingües

Los corpus monolingües generan listas de concordancias monolingües, es decir, listas de contextos en una lengua. Las concordancias monolingües pueden ser útiles por varias razones. Su utilidad principal es observar una palabra en su contexto. Por ejemplo, podemos generar una lista de concordancias monolingües en el *BNC* de la expresión léxica *Prime Minister* para observar los contextos en los que se utiliza esta expresión.

En otro ejemplo, se introduce en el *BYU-OED Oxford English Dictionary* el lema *arrow* a máximo 9 posiciones de distancia de la categoría gramatical *sustantivo*. El corpus nos propone concordancias de sustantivos que aparecen cercanos a la palabra *arrow*, entre los que se encuentra *quiver*, que es el recipiente en el que un arquero guarda las flechas (véase la Tabla 20).

A gaily-painted <b>quiver</b> , full of <b>arrows</b>
He could draw an <b>arrow</b> from his <b>quiver</b> [...]

**Tabla 20. Resultado de concordancias del lema *arrow* a máximo 9 posiciones de distancia de un sustantivo.**

En la mayoría de los corpus, es posible ordenar los resultados de las concordancias por las palabras que se encuentran a izquierda y derecha de la palabra de búsqueda, por ejemplo, por la primera palabra a la derecha. En el corpus *BNCweb* también es posible especificar una categoría gramatical, además de una posición, para ordenar las concordancias. Por ejemplo, supongamos que estamos buscando un adjetivo para describir en inglés a la palabra *rain*. Como vemos en la Ilustración 54, en el *BNCweb*, podemos buscar concordancias de la palabra *rain* y ordenarlas por los adjetivos que aparecen en la

primera posición a la izquierda de *rain*. De esta forma visualizamos fácilmente combinaciones como *abrupt rain*, *abundant rain*, *acid rain*, *African rain*, *Amazonian rain*, *American rain*, *chilling rain*, *constant rain*, *ceaseless rain*, etc.

Your query "rain" returned 6235 hits in 1316 different texts (98,313,429 words [4,048 texts], frequency: 63.42 instances per million words), thinned with method *random selection* hits, sorted on *position -1* with tag-restriction *any adjective* (989 hits) [\[data removed from table\]](#)

<	<<	>>	>	Show Page:   1	Show Sentence View	Show in random order	New Query	Go!
Sort parameters:		Position: 1 Left	Tag restriction: any adjective	<input type="checkbox"/> exclude	Starting with letter: all	Submit		
No	Filename	Hits 1 to 50		Page 1 / 20				
1	<a href="#">K4L 479</a>	to pursue a victory target of 174, and even an 18-minute_	AJ0	rain	stoppage just prior to the start of the final hour failed to			
2	<a href="#">GV2 1392</a>	course, that you take care to stay alive.' 23_	CRD	Rain	and Oliver heard Foucard and Denis rattle downstairs and then cobble			
3	<a href="#">G1V 921</a>	I thought that they were good For any neighbourhood_	4_	CRD	Rain	punishes the city, like raw wind that batters flesh, ever		
4	<a href="#">K3H 117</a>	Tour match WARWICK v AUSTRALIA at Edgbaston Australia first inns: 59-2_	CRD	Rain	stopped play Rowing challenge MEMBERS of the public have the cha			
5	<a href="#">BPO 1551</a>	The fourth day dawned, and wore through its hours of abrupt_	AJ0	rain	and mild, blustering wind 'They all worked in scuffed brigandines			
6	<a href="#">J28 74</a>	areas, and the need to protect against strong winds and abundant_	AJ0	rain	, are a few of the many practical considerations which have had			
7	<a href="#">B14 424</a>	vast tracts of forest (Begley, 1988). Increasingly acid_	AJ0-NN1	rain	is a problem in Europe too. Not only the air we			
8	<a href="#">GU5 17</a>	(sometimes 5.6) is termed acid precipitation, or simply acid_	AJ0-NN1	rain	. The precipitation that fell before the Industrial Revolution and was pre			
9	<a href="#">B7M 268</a>	what to do to control the emissions of sulphur that cause acidified_	AJ0-VVD	rain	. Canada wants action. The US does not. Canada's			
10	<a href="#">AM2 327</a>	coloured of the African forest monkeys known as guenons. The African_	AJ0	rain	forests are inhabited by several closely related guenons. Because each			
11	<a href="#">CJ3 1289</a>	in the desert. VENOMOUS ANIMALS In the heart of the African_	AJ0	rain	forest a chameleon crouches on a branch, stalking an insect.			
12	<a href="#">ATR 1158</a>	the atmosphere with moisture to form sulphuric and nitric acids. All_	DT0	rain	is naturally acidic. Acidity is reckoned on the pH scale of			
13	<a href="#">J18 1474</a>	, it has been noted that by 1981 only 20% of all_	DT0	rain	forests including tropical ones in Australia remained when compared w			
14	<a href="#">ARS 1302</a>	tripod and forms a natural 'thatch', which sheds all_	DT0	rain	except in a severe gale. The outside thatch is bleached by			
15	<a href="#">BLX 876</a>	on the scene, the local Indians living deep in the Amazonian_	AJ0	rain	forests caught and killed these animals to provide themselves with the p			
16	<a href="#">K5C 315</a>	tribes still to be found living in remote areas of the Amazonian_	AJ0	rain	forests. Most often known to us as witch doctors or medicine			
17	<a href="#">CFH 765</a>	ozone layer, and how to stop the deforestation of the Amazonian_	AJ0	rain	forests. The latter explains why the Summit is being held in			
18	<a href="#">K1D 2786</a>	been presented with an international award for his research on the Amazonian_	AJ0	rain	forest. Working with trained helpers, he's compiling data on			
19	<a href="#">HL2 743</a>	expressed his support for the government's efforts to protect the Amazonian_	AJ0	rain	forest. He declined to postpone his visit to Argentina, expressing			
20	<a href="#">K1C 1349</a>	he needs. Michael Croggon's studies could help save the Amazonian_	AJ0	rain	forests ... but only if he can find the money to fund			
21	<a href="#">J18 856</a>	endemism have been identified, 15 in the great block of Amazonian-Orinoco-Guayanan_	AJ0	rain	forest. Today, the forest is at a maximum, so			
22	<a href="#">E9F 560</a>	, tiny monkeys that live in the canopy of the South American_	AJ0	rain	forest, have a very hard time raising their babies. They			

**Ilustración 54. Resultado de concordancias ordenadas por el primer adjetivo a la izquierda de *rain* en *BNCweb*.**

### 3.3.4.3.1.2 Lista de concordancias bilingües

Los corpus bilingües paralelos generan listas de concordancias bilingües. Las concordancias bilingües pueden ser útiles para acceder y seleccionar equivalentes de traducción, ya que tanto la consulta en lengua origen como su equivalente en lengua meta aparecen rodeados de un contexto que se puede comparar con el contexto de la traducción para comprobar si es el mismo.

A continuación incluimos un ejemplo de técnica de búsqueda en la que el resultado es una concordancia que puede servir para elegir un equivalente de traducción en función del contexto. Introducimos la palabra *play* en el corpus paralelo *COMPARA* y accedemos a las concordancias de dicha palabra. Como vemos en los resultados de la Tabla 21, dependiendo del contexto, *play* se traduce en portugués por *tocar* (cuando se refiere a un instrumento), *jogar* (cuando se refiere a un deporte) o *fazer a* (cuando se refiere a adoptar un rol).

«[...] and not being able to <b>play</b> the piano.»	«[...] e à incapacidade de <b>tocar</b> piano.»
Joe wanted to switch partners and <b>play</b> the best of three sets, [...]	Joe queria trocar de parceiros e <b>jogar</b> de novo, uma melhor de três, [...]
([...] he likes to <b>play</b> the father in our relationship.)	([...] gosta de <b>fazer a</b> figura paterna no nosso relacionamento.)

**Tabla 21. Lista de concordancias de la palabra *play* con sus equivalentes en portugués (énfasis añadido).**

Todos los programas de traducción asistida que hemos analizado generan listas de concordancias bilingües. Por ejemplo, en *SDL Trados*, cuando se detecta en la memoria un segmento igual o similar al que estamos traduciendo, nos muestra dicho segmento junto con su traducción o traducciones de la memoria.

### 3.3.4.3.2 Lista de palabras

Las listas de palabras pueden ser de dos tipos. Por un lado, están las listas de palabras más frecuentes del corpus. Por otro lado, existen las listas de palabras clave, que se extraen comparando las listas de palabras más frecuentes de dos corpus; el resultado es una lista de palabras propias de uno de los corpus, las que no se encuentran en el otro corpus.

Las listas de palabras pueden ser útiles para tener una visión rápida de la terminología propia de un ámbito. *Bwananet*, con la opción *unidades fuera de contexto*, genera listas de formas, lemas o categorías. La lista de lemas permite restringir la longitud de la lista, sin por ello disminuir el número de términos diferentes que se encuentran en la lista, ya que no aparecerán términos repetidos (por ejemplo, el mismo sustantivo en singular y en plural, o diferentes conjugaciones de un verbo).

A continuación presentamos un ejemplo de generación de listas de palabras frecuentes, en el corpus en inglés de economía en *Bwananet*. Si buscamos una lista de lemas, las primeras palabras que aparecen son del lenguaje general, preposiciones y determinantes principalmente, lo cual es lógico ya que son las palabras más frecuentes en todos los corpus. Sin embargo, en primera posición de la lista aparece el signo =, propio de los textos sobre economía y también encontramos otras palabras propias del ámbito como *rate, market, price, good, capital, investment, etc.*

Los corpus de BYU también generan listas de palabras (pero solo si se especifica una categoría gramatical, extrae una lista de palabras con dicha categoría). El *BNCweb* también genera listas de formas o lemas frecuentes en un corpus y permite restringir la búsqueda utilizando patrones lingüísticos o gramaticales. Por ejemplo puede generarse una lista de

los verbos más frecuentes del corpus que empiezan por *pre*: *present, prepare, prevent, press, prefer, predict, etc.*

Este corpus genera, asimismo, listas de palabras clave comparando las listas de palabras frecuentes de todo el *BNC*, el *BNC* de textos escritos y el *BNC* de textos orales. Por ejemplo, si comparamos en *BNCweb* las listas de palabras de todo el *BNC* y el *BNC* de textos orales. La lista resultante de palabras clave incluye *er, yeah, erm, it, oh, yes, cos, gon, na, etc.*, que son palabras propias del lenguaje oral.

En general la mayoría de programas de gestión de corpus, como *WordSmith Tools 4* también proporciona listas de palabras clave, siempre que el usuario facilite al programa al menos dos listas de palabras de corpus de naturaleza distinta (uno especializado y uno de referencia preferiblemente) para ser comparados por la herramienta.

*TerminoWeb* también permite extraer una lista de terminología de los corpus que recupera de Internet. Por ejemplo, como vemos en la Ilustración 55, a partir de un corpus de cerámica, se puede generar una lista de términos de dicho ámbito, como *residual stresses, strain, neutron diffraction, particle, alumina, etc.*

TERM	FREQUENCY	SOURCE	STATUS	Select	Accept	Reject
<a href="#">New Term</a>				Select All	Accept All	Reject All
residual stresses	22	Extraction	Accepted	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
residual	37	Extraction	Undefined	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
stresses	33	Extraction	Undefined	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
strain	38	Extraction	Accepted	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
neutron diffraction	11	Extraction	Accepted	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
particle	54	Extraction	Accepted	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
alumina	28	Extraction	Accepted	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Ilustración 55. Generación y validación de una lista de términos en *TerminoWeb*.

### 3.3.4.3.3 Lista de agrupaciones de palabras

Las agrupaciones de palabras (opción *n-gramas, clusters*, dependiendo del recurso) son secuencias de dos o más palabras que aparecen con frecuencia en el corpus. Las agrupaciones de palabras pueden ayudar a tener una visión rápida de cómo se combina la terminología propia de un ámbito.

Cuando se buscan agrupaciones de palabras, la consulta puede variar. Puede que no se especifique consulta y únicamente se indique el número de palabras que debe aparecer

en la agrupación (dos o más). Puede que se indique una palabra que debe aparecer en la agrupación, por ejemplo *mesita*. En otras ocasiones se indica la secuencia gramatical de las palabras de la agrupación, por ejemplo agrupaciones de *nombre+adjetivo+adjetivo*. También es posible que se indiquen palabras y categorías gramaticales que deben aparecer (o no) en la agrupación, por ejemplo *mesita+preposición+adjetivo*.

De los corpus que hemos analizado solo algunos tienen la opción de agrupaciones de palabras. *Bwananet* genera listas de agrupaciones de máximo dos palabras (opción *unidades fuera de contexto, incluir bigramas*) sin especificar ninguna consulta. Los corpus de *BYU* generan agrupaciones de palabras, especificando palabras o categorías gramaticales. *CREA* genera *agrupaciones* de palabras especificando una o más palabras que deben aparecer en la agrupación. Por ejemplo, pueden buscarse agrupaciones de tres palabras en las que se incluye la palabra *mesita*. La lista de agrupaciones de palabras incluye: *mesita de noche, mesita de madera, mesita de luz, mesita del teléfono*, etc.

*WordSmith Tools 4* también genera listas de agrupaciones de palabras que contienen una determinada palabra o palabras de búsqueda. El usuario puede indicar a la herramienta un número mínimo y máximo de n-gramas para recuperar en las agrupaciones, por ejemplo agrupaciones de 3 a 5 palabras. También puede configurarse el mínimo de repeticiones que debe tener la agrupación en el corpus para que aparezca en la lista, así como el número de palabras a derecha e izquierda que debe tener en cuenta el programa para generar las agrupaciones. En los resultados, el programa especifica la longitud de las agrupaciones y el número de repeticiones de cada agrupación en el corpus. Por ejemplo, podemos generar una lista de agrupaciones (*clusters*) de 2 a 4 palabras que contengan *seco* en un corpus de la cerámica; como resultado obtendremos los términos compuestos *molienda en seco, molturación en seco, prensado en seco*, etc.

Finalmente, la interfaz *Phrases in English* permite acceder a agrupaciones de hasta 8 palabras (*n-grams*), que pueden seguir un patrón (añadiendo o excluyendo de la búsqueda palabras o categorías gramaticales en determinada posición). Asimismo, la opción *chargrams* busca agrupaciones de hasta 8 caracteres que se repiten con frecuencia dentro de las palabras del corpus. La función *phrase-frames* presenta grupos de agrupaciones de palabras que son iguales excepto por una palabra, lo que ayuda a identificar patrones fraseológicos. Por ejemplo, indicando a la herramienta que el número de *frames* es 3 y que una de las palabras que debe aparecer en el patrón es *play*, recupera entre otros el patrón *play \* the* (actuando como verbo \* artículo) y las variantes a las que puede accederse que siguen dicho patrón son agrupaciones de palabras donde el asterisco se sustituye en la



mayoría de los casos por una preposición (*play in the, play for the, play with the, play down the, etc.*). La función “*Drill down*” *RegEx search* accede a expresiones o frases más largas que incluyen las palabras de búsqueda e indica el porcentaje de aparición de determinadas frases dentro de las frases más largas. Por ejemplo, con la consulta de *door* se accede a *the door*, más allá *opened the door and, she opened the door and*, y otras expansiones como *knocked on the door of, knocked on the door and, etc.*

#### 3.3.4.3.4 Lista de colocaciones

Una colocación es una palabra que aparece frecuentemente cercana a otra palabra en el corpus. La utilidad de las colocaciones reside en que se puede acceder a palabras que aparecen en un contexto inmediato a otra palabra sin tener que leer todas las concordancias de un corpus en las que aparece dicha palabra. En definitiva, puede ayudar a agilizar el proceso de búsqueda.

A continuación mostramos ejemplos de resultados de listas de colocaciones en algunos de los corpus analizados: *Corpus of Contemporary American English* (de BYU), *BNCweb* y *WordSmith Tools 4*. En el caso de los corpus de BYU, es necesario especificar la categoría gramatical que deben tener dichas colocaciones. Por ejemplo, si buscamos los *sustantivos* que aparecen en el *Corpus of Contemporary American English* (de BYU) como colocación de la palabra *television*, se genera una lista de colocaciones entre las que se incluye: *radio, news, show, cable, network, station, series, etc.*

El *BNCweb* genera listas de colocaciones de la palabra de búsqueda, y es posible especificar que las colocaciones pertenezcan a una determinada categoría gramatical. Por ejemplo las colocaciones nominales más frecuentes a cinco posiciones a izquierda o derecha de la palabra *house* en este corpus son *commons, lords, white, representatives, opera, manor, prices, boarding, etc.*

*WordSmith Tools 4* genera listas de colocaciones de una palabra y proporciona información acerca de la situación en la que aparece cada colocación respecto a la palabra de búsqueda (posiciones a izquierda y derecha), así como la cantidad de veces que se repite dicha colocación en el corpus en el lugar señalado. Por ejemplo, generamos una lista de colocaciones del término *software* de un corpus de la informática en español<sup>53</sup>. En la siguiente imagen solo incluimos una muestra de los resultados obtenidos. Aunque si

---

<sup>53</sup> Corpus de la informática en español desarrollado por el grupo OLST, de la Universidad de Montreal.

observamos la lista completa de colocaciones proporcionada por el programa, vemos que *hardware* se sitúa en 22 ocasiones dos posiciones delante de *software* porque una expresión frecuente en el corpus es *hardware y software*. *Libre* se repite 18 veces una posición a la derecha de *software* porque otra expresión habitual es *software libre*.

N	Word	With	elation	Total	tal	Left	al	Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3
1	SOFTWARE	software	0.000	1.052	20	31	10	8	5	5	1	992	2	5	5		
2	DE	software	0.000	993	597	396	80	88	96	62	272	0	140	34	85		
3	EL	software	0.000	385	292	93	25	23	32	28	184	0	9	19	25		
4	Y	software	0.000	288	136	153	23	21	32	26	33	0	50	28	19		
5	EN	software	0.000	197	91	106	13	19	19	13	27	0	26	20	21		
6	LA	software	0.000	194	102	92	22	28	42	10	0	0	3	26	27		
7	QUE	software	0.000	185	73	112	25	18	14	13	3	0	45	9	20		
8	PARA	software	0.000	173	46	127	15	19	8	4	0	0	48	28	25		
9	DEL	software	0.000	150	132	18	13	6	11	9	93	0	4	3	2		
10	LOS	software	0.000	117	67	50	13	20	32	2	0	0	3	16	9		
11	UN	software	0.000	117	73	44	8	13	13	3	36	0	1	11	12		
12	A	software	0.000	110	38	72	14	14	5	5	0	0	12	11	21		
13	CON	software	0.000	107	64	43	10	15	10	19	10	0	7	9	8		
14	LAS	software	0.000	80	37	43	13	9	15	0	0	0	1	15	4		
15	ES	software	0.000	70	26	44	6	7	4	9	0	0	12	6	16		
16	COMO	software	0.000	66	26	40	10	2	3	8	3	0	13	7	6		
17	SE	software	0.000	66	22	44	5	6	11	0	0	0	11	14	7		
18	HARDWARE	software	0.000	57	42	15	0	1	11	22	8	0	1	8	2		

**Ilustración 56. Resultado de colocaciones de *software* en *WordSmith Tools 4*.**

Otro resultado que puede proporcionar una búsqueda en corpus, en relación con las colocaciones, son patrones lingüísticos. En el programa *WordSmith Tools 4*, es posible observar en una tabla las colocaciones más comunes a izquierda y derecha de la palabra pivote o palabra de búsqueda en el lugar exacto en el que con más frecuencia aparecen (opción *patterns*). Vemos en la Ilustración 57 que algunos de los términos que aparecen con más frecuencia en este corpus dos posiciones a la derecha de *software* son *desarrollo*, *hardware*, *desarrolladores*, *aplicaciones*, *distribución*, *productos*, *herramientas*, etc. Un verbo que suele situarse inmediatamente anterior a la palabra *software* en este corpus es *adquirir*. Algunos términos que se sitúan una posición a la derecha de *software* son *libre*, *antivirus*, *propietario*, etc.

N	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	DE	DE	DE	DE	DE	SOFTWARE	DE	DE	DE	DE	DE
2	EL	LA	LA	EL	EL		Y	LA	LA	Y	EL
3	Y	EL	Y	Y	DEL		PARA	Y	EL	EN	Y
4	LA	EN	LOS	DESARROLLO	PANDA		QUE	PARA	EN	QUE	LA
5	QUE	Y	EL	HARDWARE	UN		EN	EN	PARA	EL	QUE
6	EN	PARA	EN	CON	Y		ASSURANCE	EL	A	LA	EN
7	PARA	LOS	QUE	QUE	EN		LIBRE	LOS	QUE	A	LOS
8	A	QUE	LAS	EN	ESTE		NO	LAS	Y	LAS	CON
9	DEL	CON	UN	RROLLADORES	NUEVO		3DS	SE	ES	LOS	A
10	LAS	UN	DEL	POR	SU		ES	MAX	UN	PARA	PARA
11	LOS	A	UNA	LA	BMC		COMO	CON	LOS	SOFTWARE	UN
12	CON	SOFTWARE	SE	APLICACIONES	CON		SE	UN	SE	UN	SOFTWARE
13	SOFTWARE	LAS	ARDWARE	ES	AL		A	A	CON	CON	UNA
14	MÁS	SU	PARA	DEL	POR		EL	UNA	SUS	COMO	ES
15	COMO	DEL	CON	DISTRIBUCIÓN	HARDWARE		ANTIVIRUS	ISTRACIÓN	7	SE	LAS
16	UN	UNA	AL	PRODUCTOS	EXPRESIÓN		PROPIETARIO	QUE	COMO	DEL	SISTEMA
17	SISTEMA	ES	SOFTWARE	COMO	HEYNE		ADOBE	POR	SOFTWARE	ES	NO
18	VERSIÓN	NO	SU	PAQUETE	ADQUIRIR		CON	O	POR	LIBRE	COMO
19	ES	SIN	VERSION	HERRAMIENTAS			MICROSOFT	LIBRE	ESTE	DATOS	O

**Ilustración 57. Resultados de la búsqueda de patrones de colocaciones de la palabra *software* en *WordSmith Tools 4*.**

#### 3.3.4.3.5 Lista de palabras relacionadas semánticamente

Algunos corpus han incorporado la función de búsquedas semánticas. Esta opción consiste en que se puede acceder a los sinónimos de la palabra introducida y realizar distintos tipos de búsqueda en relación con dichos sinónimos. Esta técnica de búsqueda puede ser útil para acceder a una palabra a partir de otras palabras que están relacionadas semánticamente. De los corpus analizados, los corpus de BYU y *TerminoWeb* permiten la realización de búsquedas semánticas.

Por ejemplo en el *Corpus of contemporary American English* (de BYU) pueden buscarse los sinónimos de *beautiful*, introduciendo [=beautiful]. El resultado son sinónimos de esta palabra como *wonderful, attractive, striking, lovely, handsome*, etc. De todos estos sinónimos se indica su frecuencia de aparición en el corpus y se puede acceder a sus concordancias o contextos. También es posible comparar su frecuencia y distribución por tipos de texto o fechas.

Asimismo, podrá accederse a colocaciones de los sinónimos de una palabra, por ejemplo en el *Corpus del español* (de BYU) buscamos todos los sinónimos de *lluvia* que aparecen cerca de sinónimos del adjetivo *fuerte*, entre los que se recuperan *tormenta, tempestad, aguacero, borrasca, precipitación, granizada, diluvio, chubasco y chaparrón*, entre otros. Como vemos en la Ilustración 58, *aguacero* aparece junto al adjetivo *fuerte* y también junto a sinónimos de dicho adjetivo como *violento, recio o penetrante*.

**CORPUS DEL ESPAÑOL**  
Mark Davies / Brigham Young University

MOSTRAR  
 LISTA  GRÁFICO  PCEC  COMPARAR

BUSCAR  
 PALABRA(S) [-fuerte]  
 CONTEXTO [=lluvia] 4 4  
 CAT GRAM [-seleccionar-]  
 AI A7AR BUSCAR BORRAR

SECCIONES  MOSTRAR

ORDENAR Y LIMITAR  
 ORDENAR FRECUENCIA  
 MÍNIMO FRECUENCIA  0

VER CONTEXTO: HACER CLIC EN LA PALABRA (TODAS LAS SECCIONES), NÚMERO (UNA SECCIÓN), O [CONTEXTO] (VARIAS)

	<input type="checkbox"/>	CONTEXTO	TOT <input type="checkbox"/>
1	<input type="checkbox"/>	LLUVIA [S]	52
2	<input type="checkbox"/>	TORRENTA [S]	31
3	<input type="checkbox"/>	TEMPESTAD [S]	21
4	<input type="checkbox"/>	HUMEDAD [S]	19
5	<input type="checkbox"/>	AGUACERO [S]	14

PALABRA CLAVE EN CONTEXTO (PCEC) Más información...

illas. Antuco abrió de par en par la ventana del patio. El **aguacero violento** había cesado al hacerse la noche. Ahora llovía misteriosamente en l...  
 rganizadas Tarde de lluvias Apagones en 35 sectores de San José y Heredia Con un **fuerte aguacero**, ayer en la tarde, la estación lluviosa dijo...  
 alo Hincado, del Ejército Nacional, junto a otros dos presidiarios que aprovecharon un **fuerte aguacero** y un prolongado apagón para fugarse...  
 alo Hincado, del Ejército Nacional, junto a otros dos presidiarios que aprovecharon un **fuerte aguacero** y un prolongado apagón para fugarse...  
 partido entre olímpicos y jaijos se suspendió en su fecha original, debido a una **fuerte aguacero** que cayó sobre esta ciudad minutos antes de...  
 del ermitaño. Los relámpagos y truenos se sucedían con mucha rapidez, y un **fuerte aguacero** debía desprenderse muy en breve sobre las casc...  
 resistiesen de su propósito los que lo armaban, vino en auxilio del orden un **recio aguacero**, al que habría resistido un entusiasmo popular arr...  
 evándose el Sacristán las llaves, como las de una casa deshabitada. Y un **fuerte aguacero** que principió á caer en seguida, dejó completamente...  
 regra y medrosa soledad no se escuchaban sino el sordo rumor del **aguacero**, el **penetrante** silbido del viento, el ronco y vago eco de las ondas

**Ilustración 58.** Lista de sinónimos de *lluvia* en el contexto de sinónimos de *fuerte*; concordancias de *aguacero* en el *Corpus del español* (de BYU).

*TerminoWeb* contiene una función de búsqueda de patrones semánticos en los contextos del corpus. Los patrones disponibles en *TerminoWeb* son de antonimia, causa, definición, función, hiperonimia, meronimia, similitud y sinonimia. Por ejemplo, los patrones de antonimia proporcionados por *TerminoWeb* son *be antonym*, *be opposed to*, *be the opposite of* y *contrast with* (véase Ilustración 59).

En el ejemplo de la Ilustración 60 hemos descargado desde Internet con *TerminoWeb* un corpus sobre cerámica en inglés, y hemos buscado dentro del corpus los patrones que indican relaciones de sinonimia y similitud en el corpus. Se han recuperado contextos en los que aparecen los patrones *also called*, *known as*, *distinguished from*, *analogous to*. Podemos seleccionar cada concordancia para ampliar el contexto. Hemos accedido a un contexto en el que se explica qué es la frita (*frit*). También podemos acceder a la página de Internet de la que se ha extraído el contexto.



Ilustración 59. Patrones de relaciones semánticas en *TerminoWeb*.

TerminoWeb 2.0

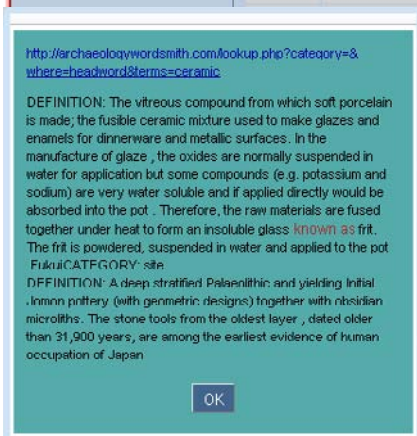
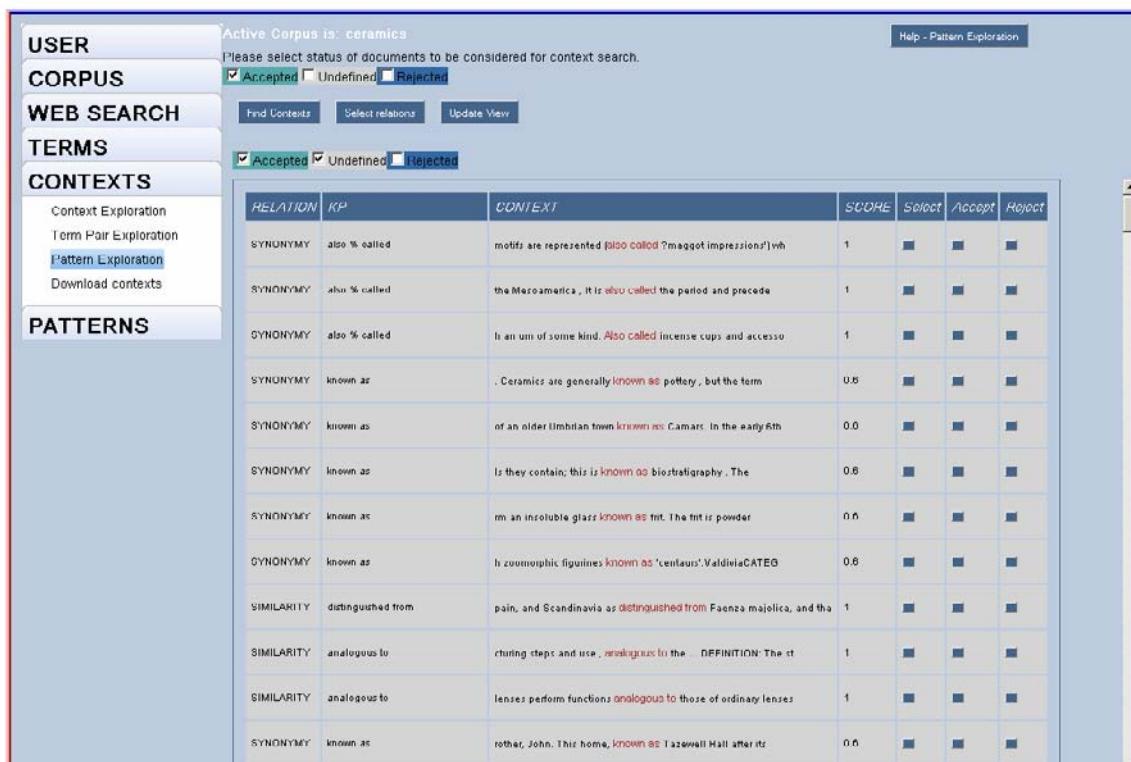


Ilustración 60. Resultado de la búsqueda de patrones de similitud y sinonimia en un corpus virtual sobre cerámica en *TerminoWeb*.

### 3.3.4.3.6 Lista de palabras relacionadas sintácticamente

Algunos corpus también han incorporado la función de búsqueda de relaciones sintácticas. Hemos encontrado este resultado en *Sketch Engine*. Como hemos visto en la descripción de esta herramienta, la búsqueda de palabras relacionadas sintácticamente está basada en la generación de *word sketches*, que son resúmenes del comportamiento sintáctico y colocacional de una palabra, extraídos de un corpus (Kilgarriff et al. 2004: 105). Por medio de esta función podemos observar con qué palabras está relacionada sintácticamente una palabra. Por ejemplo, en el caso de un sustantivo, con qué verbos coaparece funcionando como complemento directo, qué adjetivos lo modifican, con qué otros sustantivos aparece combinado por medio de las conjunciones y/o, etc.

También es posible buscar palabras que presentan un mismo comportamiento sintáctico en el corpus (con la función *Thesaurus* de *Sketch Engine*). Por ejemplo, algunas de las palabras que muestran el mismo comportamiento sintáctico que *university* en el *BNC* (disponible para la consulta en *Sketch Engine*) son *college*, *school*, *council*, *department*, *authority*, *institute*, etc.

Finalmente, con *Sketch Engine*, podemos comparar el comportamiento sintáctico de dos palabras relacionadas, por ejemplo, *university* y *college*. Si utilizamos la función *Sketch-Diff* con el corpus *BNC* en *Sketch Engine*, observamos que ambas palabras muestran un comportamiento sintáctico muy similar, aparecen combinadas con los mismos verbos, modificadas por los mismos adjetivos, etc. En la Ilustración 61 observamos que *college* y *university*, en los contextos del *BNC* funcionan en ocasiones a modo de adjetivo modificando algunos sustantivos, como en *university lecturer*, *college lecturer*, *university campus*, *college campus*, etc. Asimismo, observamos que los sustantivos *degree* y *education* aparecen con mayor frecuencia modificados por *university* que por *college*. En el corpus se han encontrado 70 contextos en los que *university* modifica a *degree* (*university degree*), mientras que solo se han encontrado 7 contextos en los que *college* modifica a *degree* (*college degree*).

En *Scientext* también se puede especificar que en los resultados se recuperen relaciones sintácticas entre términos del corpus. Por ejemplo de una preposición que introduce al lema *effect*. Como vemos en la Ilustración 62, podemos acceder a un contexto ampliado de una concordancia de *on long-term effects*, en la que se puede observar el análisis sintáctico efectuado en este corpus, por ejemplo, el de la frase *on long-term effects of stable type 2 diabetic subjects or poorly controlled diabetic subjets*.



modifies	5105	2031	1.1	0.6
lecturer	77	46	8.4	8.3
campus	53	14	8.1	7.2
library	180	65	8.1	6.8
graduate	60	13	7.9	6.3
funding	67	13	7.5	5.5
student	210	91	7.4	6.3
professor	42	7	7.3	5.3
entrance	28	6	6.5	4.8
degree	70	7	6.4	3.3
tutor	20	8	6.3	5.6
hospital	83	17	6.3	5.6
filmmaker	6	10	4.7	6.3
London	104	86	6.1	5.9
teacher	76	31	6.1	4.9
education	90	19	5.9	3.7
archive	8	2	5.1	5.7
lecture	7	11	4.5	5.7
course	78	54	3.6	5.1
staff	44	17	5.0	5.2
accommodation	6	10	3.9	5.1
council	68	21	4.9	3.2
examination	13	6	4.7	3.9
school	39	20	4.5	3.5
building	25	21	4.2	4.1
authority	33	23	4.0	3.3

Ilustración 61. Lista de palabras modificadas por *college* y *university* en el BNC con la función *Sketch-Diff* de *Sketch Engine*.

The screenshot shows the Sketch Engine search interface. At the top, the search criteria are set to 'effect' (Word 1) and 'Preposition (0)' (Word 2). The syntactic relation is 'introduced by the preposition (NOMPREP)'. The search results show 18 occurrences. A red arrow points to a detailed syntactic analysis window for the word 'effects' in the context 'on long-term effects of stable type 2 diabetic subjects or poorly controlled diabetic'. The analysis shows the following structure:

- PREP= on
- NOMPREP= effects
  - NN= long-term
  - PREP= of
    - NOMPREP= or
      - CC= subjects
        - NN= stable
        - NN= type
        - NNSP= 2
        - NN= diabetic
      - CC= subjects
        - NN= controlled
        - ADV= poorly
        - NN= diabetic

Ilustración 62. Resultado de relaciones sintácticas en el corpus *Scientext*.

### 3.3.4.3.7 Lista de categorías gramaticales

En algunos corpus etiquetados gramaticalmente es posible acceder a una lista de las categorías gramaticales más frecuentes del corpus. Por ejemplo, en *Phrases in English*, la opción *PoS-Grams* permite generar una lista de categorías gramaticales frecuentes del *BNC*. Es posible especificar la frecuencia mínima y máxima, el número de resultados por búsqueda, el orden de la lista (frecuencia ascendente, descendente, orden alfabético). También se puede recuperar una lista de agrupaciones de categorías gramaticales que sigan un determinado patrón (incluyan o excluyan categorías gramaticales).

En el siguiente ejemplo buscamos agrupaciones de dos categorías gramaticales que incluyan verbos con terminación en *ing* en segunda posición. La primera agrupación de la lista indica que el mayor número de agrupaciones es de preposición seguida de verbo en *ing*, a continuación sustantivo singular seguido de verbo en *ing* y también hay un alto porcentaje de adverbio seguido de verbo en *ing*, entre otras agrupaciones.

The screenshot shows the 'Explore PoS-Grams from the British National Corpus' interface. On the left, search filters are set: Minimum frequency 20, Maximum frequency (empty), Start with item 1, Data chunk size 1000, Order: Descending frequency, and PoS filters: 'verbs: -ing form' with 'V7G' selected. On the right, the results are displayed as a table titled 'Displaying 2-POS-grams by Types'.

POS-GRAMS	Types	>=3 Tokens	>=3
PRP VVG	11,183	204,433	
NN1 VVG	6,790	45,402	
AV0 VVG	5,528	53,072	
NN2 VVG	4,248	28,831	
CJC VVG	3,318	69,320	
VBD VVG	2,489	94,895	
PRF VVG	2,298	76,622	
VBB VVG	1,920	73,666	
CJS VVG	1,612	19,155	
VBZ VVG	1,541	56,854	
PNP VVG	1,223	19,729	
VV1 VVG	1,153	8,964	
VVD VVG	1,137	9,951	
NN1 VBG	924	7,815	
VBI VVG	904	24,972	
VBN VVG	776	19,869	
VVB VVG	748	5,796	
XXD VVG	737	21,858	
DT0 VVG	644	6,589	
VVN VVG	577	4,208	
AVP VVG	563	5,406	
NN2 VBG	538	4,209	
NND VVG	410	5,302	
VVZ VVG	374	2,241	
PN1 VVG	368	2,945	
PNX VVG	236	1,533	
CJT VVG	208	1,776	
ATO VVG	200	1,959	
AJD VVG	196	1,233	
CRD VVG	196	1,373	
NPO VVG	185	833	
NN1 VHG	171	947	
DPS VVG	144	1,010	
AV0 VBG	137	5,256	

Ilustración 63. Resultado de lista de categorías gramaticales en *Phrases in English*.



### 3.3.4.3.8 Representación gráfica de las relaciones de una palabra

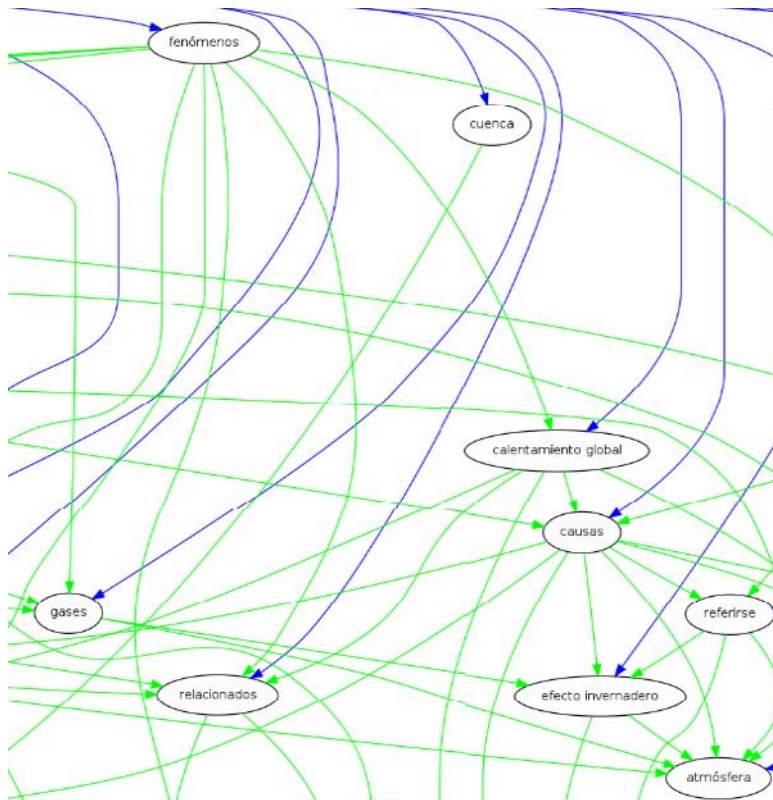
Hemos observado que se pueden representar gráficamente las relaciones de una palabra en el corpus, por ejemplo a partir de un grafo<sup>54</sup>. En nuestro análisis hemos observado que la herramienta de gestión y análisis de corpus *Jaguar* genera grafos para representar la densidad terminológica de un corpus que se ha descargado desde Internet. Este grafo permite al usuario detectar con un golpe de vista la terminología del corpus (similar a en una lista de palabras claves), pero además se observan las relaciones que existen entre los distintos términos.

Asimismo, pinchando en cada uno de los nodos del grafo (correspondiente a una unidad terminológica) el usuario accede a los contextos de dicho término en el corpus con lo que puede decidir si el término pertenece o no al ámbito temático del corpus. Con los términos que el usuario ha validado, la herramienta puede realizar búsquedas en Internet para restringir el corpus de especialidad, con lo que se generará un nuevo grafo, en el que el usuario podrá incluir o descartar más términos. Este proceso puede repetirse tantas veces como el usuario considere necesario para delimitar todo lo posible los textos del corpus. Esta función puede ayudar al traductor a compilar un corpus especializado *ad hoc* que se descarga desde Internet para una traducción.

Por ejemplo, descargamos un corpus de Internet con *Jaguar* que debe contener los términos *cambio climático* y *calentamiento global*. En el grafo de este corpus aparecen numerosos términos relacionados con el cambio climático: *sequías*, *carbono*, *desastres*, *ecosistemas*, *recursos hídricos*, *calentamiento global*, *efecto invernadero*, *atmósfera*, etc. El grafo es demasiado grande para incluirlo al completo en una captura de pantalla, pero en la Ilustración 64 vemos una parte del grafo, en el que el término *cambio climático* aparece relacionado con *fenómenos*, que a su vez está vinculado a *calentamiento global*, de ahí se conecta con el nodo *causas*, que está vinculado a *efecto invernadero*, y *efecto invernadero* está unido a *atmósfera*, con lo que podemos observar con un vistazo rápido las relaciones que se establecen entre algunos términos clave de este corpus especializado.

---

<sup>54</sup> Definición de grafo en el apartado 3.2.4.3.2



**Ilustración 64.** Resultado de grafo generado por *Jaguar* a partir de un corpus sobre *cambio climático*.

### 3.3.4.3.9 Información estadística

Los resultados de la búsqueda en un corpus incluyen información estadística sobre el comportamiento de la expresión de consulta en el corpus. En este sentido, la cantidad de palabras de un corpus se puede medir por *tokens* o *casos*, la cantidad total de palabras que contiene el corpus, o por *types* o *tipos*, la cantidad de palabras diferentes. El *type-token ratio* es la relación entre la cantidad de palabras diferentes y el total de palabras de un corpus. Es una manera de calcular la variedad léxica de un texto. Así, en la tabla de resultados de la Ilustración 63, por cada agrupación de categorías gramaticales de la lista, podemos conocer el número de types y de tokens que se corresponden con dicho patrón gramatical en el corpus. Por ejemplo, para la primera agrupación de la lista, preposición seguida de verbo en *ing*, el número de types es 11.183 y de tokens 204.433.

Por otro lado, la *frecuencia absoluta* es el número de veces que un elemento aparece en un corpus, mientras que la *frecuencia relativa* refleja el porcentaje entre la cantidad de apariciones de un elemento respecto del tamaño del corpus. Por ejemplo, en la Ilustración 65, por la frecuencia absoluta y relativa observamos que la palabra *ordenador* es más frecuente en textos de España que en cualquier otra variedad del español y que también aparece en mayor proporción en textos del ámbito temático de la ciencia y la tecnología.



Resultados con estadísticas (RAE)

Consulta: *ordenador*, en todos los medios, en CREA  
 Resultado: 3919 casos en 1299 documentos.

Filtros: Documentos  
 Ratio: 10  
 Mantener documentos (Sólo para filtro sobre casos).  
 Filtrar

OBTENCIÓN DE EJEMPLOS

Recuperar Documentos Normal.  
 Clasificación:  
 Agrupación:  
 Marcas:

Cómo citar el CORPUS

Estadísticas

Año	%	Casos	País	%	Casos	Tema	%	Casos
1997	15.99	593	ESPAÑA	94.26	3581	1.- Ciencia y Tecnología.	47.31	1807
2004	11.29	419	ARGENTINA	1.60	61	4.- Artes.	10.10	386
2000	8.87	329	MÉXICO	0.84	32	2.- Ciencias sociales, creencias y pensamiento.	9.06	309
2003	7.20	267	CHILE	0.73	28	3.- Política, economía, comercio y finanzas.	8.30	317
1994	7.17	266	EE. UU.	0.47	18	7.- Ficción.	8.06	308
1996	6.41	238	PERÚ	0.44	17	5.- Ocio, vida cotidiana.	5.39	206
1995	6.09	226	VENEZUELA	0.42	16	8.- Miscelánea.	4.76	182
2002	5.47	203	CUBA	0.31	12	6.- Salud.	3.48	133
1993	4.96	184	BOLIVIA	0.26	10	9.- Oral.	2.90	111
Otros	26.51	983	Otros	0.63	24			

Nueva consulta: [CREA](#) [CORDE](#) [Nómina de autores y obras](#) [Ayuda](#).

Ilustración 65. Resultados de información estadística de la palabra *ordenador* en CREA.

### 3.3.5 Recapitulación de nuestra propuesta de clasificación de técnicas de búsqueda en corpus

En este capítulo hemos visto que las técnicas de búsqueda pueden variar de unos corpus a otros y que algunas técnicas se encuentran más desarrolladas y extendidas que otras. Por eso, resulta necesario establecer una clasificación y nomenclatura homogénea que ayude a evaluar los corpus y sus funciones, enseñarlas a sus usuarios y diseñar nuevos corpus.

Nuestra clasificación, que resumimos en la Tabla 22, puede utilizarse, por un lado, como una guía para enseñar a los traductores técnicas de búsqueda en corpus electrónicos. Por otro lado, nos ayudará a reflexionar, en otros apartados, sobre qué técnicas son más útiles para los traductores y vislumbrar así posibles mejoras en los corpus electrónicos para adaptarlos a las necesidades de los traductores.

CONSULTA	RECURSO O INSTRUMENTO	RESULTADO
<ul style="list-style-type: none"> <li>•Una expresión léxica</li> <li>•Un lema</li> <li>•Una forma (exacta o truncada)</li> <li>•Una expresión gramatical</li> <li>•Un número (exacto o truncado)</li> <li>•Dos o más expresiones (lemas, formas, expresiones gramaticales, números, expresiones mixtas)</li> <li>•Con presencia total, parcial o ausencia</li> <li>•En combinación continua o discontinua</li> <li>•Historial de consultas</li> <li>•Filtros                             <ul style="list-style-type: none"> <li>-Categoría gramatical</li> <li>-Área temática</li> <li>-Idioma</li> <li>-Tipo de texto</li> <li>-Área geográfica</li> <li>-Autor</li> <li>-Fecha</li> <li>-Lugar del texto</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>•Corpus monolingüe</li> <li>•Corpus bilingüe o multilingüe                             <ul style="list-style-type: none"> <li>-comparable</li> <li>-paralelo</li> <li>oalineado</li> </ul> </li> <li>•Corpus etiquetado</li> </ul>	<ul style="list-style-type: none"> <li>•Lista de concordancias monolingües o bilingües</li> <li>•Lista de palabras</li> <li>•Lista de agrupaciones de palabras</li> <li>•Lista de colocaciones</li> <li>•Lista de de palabras relacionadas semánticamente</li> <li>•Lista de palabras relacionadas sintácticamente</li> <li>•Lista de categorías gramaticales</li> <li>•Representación gráfica de las relaciones de una palabra</li> <li>•Información estadística</li> </ul>

TÉCNICAS DE BÚSQUEDA

Tabla 22. Resumen de la clasificación de técnicas de búsqueda en corpus.

### 3.4 Internet

Internet es una red mundial de ordenadores que están conectados entre sí por medio de protocolos de comunicación TCP/IP. De entre todos los servicios que ofrece Internet, en este apartado vamos a centrarnos en la World Wide Web, o sencillamente la Web<sup>55</sup>. La Web son un conjunto de archivos (principalmente en formato HTML que se caracterizan por estar conectados entre sí por medio de un sistema de enlaces hipertextuales) que se

<sup>55</sup> Este es el servicio que más importancia ha tenido y, por ello, se utilizan a menudo los términos *Internet* y *Web* como sinónimos.

encuentran en servidores de Internet disponibles para la consulta a través de un protocolo de transferencia de hipertexto, Hypertext Transfer Protocol o HTTP.

En este apartado abordamos las técnicas de búsqueda en la Web o Internet. Comenzamos en el apartado 3.4.1 resumiendo las principales utilidades que ofrecen algunos recursos de Internet frente a otros recursos para la búsqueda terminológica. En el apartado 3.4.2 presentamos la evolución de la Web y su influencia en las posibilidades de búsqueda. En el punto 3.4.3 realizamos un análisis de las funciones de búsqueda de una serie de recursos que facilitan la búsqueda en páginas de Internet. En el apartado 3.4.4 aplicamos a los recursos de Internet una clasificación de técnicas de búsqueda que ya hemos utilizado para recopilar las técnicas de búsqueda que posibilitan otros recursos, los diccionarios electrónicos y los corpus. Para terminar, en el apartado 3.4.5 resumimos la clasificación de técnicas de búsqueda en Internet.

### **3.4.1 Utilidad de Internet como recurso de búsqueda terminológica, según la bibliografía**

En la actualidad, la Web se considera la principal fuente de información disponible y, además, ofrece gran cantidad de información léxica (Maia 2000: 57). Por un lado, es una gran fuente de recursos lingüísticos: diccionarios, glosarios, bases de datos, enciclopedias, tesauros, corpus en línea, etc. (Abaitua 2002; Austermühl 2001: 38; Castrillón Cardona et al. 2005: 347; Maia 2000: 56). Por otro lado, igualmente se reconoce a Internet como un recurso lingüístico en sí mismo. Es lo que algunos autores llaman la Web como corpus (Morales 2006: 2; Zanettin 2002b: 241), tanto monolingüe como, cada vez en mayor medida, también multilingüe (Abaitua 2002; Maniez 2007: 163; Sánchez-Gijón 2003a: 21-28)<sup>56</sup>.

De ahí, la creciente demanda de herramientas que permitan procesar los textos multilingües de Internet (Abaitua 2002: 82). No en vano, este autor señala que el futuro de la lingüística de corpus y las técnicas de alineación de textos está en Internet, en el desarrollo de buscadores inteligentes que permitan indizar, catalogar, extraer información y gestionar el conocimiento de los textos multilingües que se encuentran en línea.

---

<sup>56</sup> Según Maniez (2007: 163), aunque todavía el número de páginas en Internet en inglés es superior al de otras lenguas, la tendencia está cambiando progresivamente. Se estima que en 1997 el 84% de las páginas de la Web eran en inglés, en 1999 ese número descendió a 64%, en 2001 a 60%, y en 2004 a 35%.

En esta línea, según Roberts y Bossé-Andrieu (2006: 201), actualmente hay que ampliar la definición de corpus a todo el material disponible en versión electrónica, incluido el que se encuentra en Internet. Para estas autoras, la diferencia entre lo que denominan un corpus lingüístico (LC) según la definición tradicional de corpus (textos seleccionados y compilados en formato electrónico) para un fin específico, según veíamos en el apartado anterior, y lo que llaman un corpus general (GC), que puede encontrarse en Internet, es que el corpus lingüístico tiene un objetivo más definido ya que se ha compilado para satisfacer unas necesidades concretas, y además puede consultarse con un programa de análisis de corpus, mientras que los textos de Internet tienen que consultarse bien con un programa de análisis disponible para su consulta, en el caso de que lo haya, que normalmente tendrá funciones más limitadas, o con un buscador de Internet, que no cuenta con funciones de análisis de corpus. Aun así, estas autoras señalan que Internet es el tipo de corpus más utilizado por los traductores (Roberts y Bossé-Andrieu 2006: 203).

Internet presenta una serie de ventajas, pero también de desventajas frente a otras herramientas de búsqueda terminológica que cabe mencionar.

Entre las ventajas de Internet, encontramos la disponibilidad y rapidez en las búsquedas; es más versátil y está más actualizado que cualquier recurso (Friedbichler y Friedbichler 2000: 112; Rundell 2000; Varantola 2006: 221). Internet contiene más terminología y más terminología actualizada y en contexto que cualquier otro recurso (Harman 2002: 2; Varantola 2006: 222).

En comparación con los diccionarios en papel, la ventaja de Internet es evidente en cuanto a actualización, rapidez en las búsquedas, presentación de los términos en contexto, y facilidades de búsqueda. Los diccionarios electrónicos y en línea han mejorado algunos de estos aspectos y, así, han disminuido sus desventajas, pero la realidad es que las mejoras en los diccionarios, al menos hasta ahora, han sido mínimas por regla general, ya que muchos diccionarios electrónicos disponibles todavía son meras copias de diccionarios en papel (Harman 2002: 5-6; Sharpe 1995: 48). Por ejemplo, las búsquedas en Internet son útiles en traducción para acceder a terminología sobre todo en lenguas en las que los diccionarios son escasos, caros y se desfasan rápidamente, como por ejemplo en el caso del chino (Harman 2002).

Con respecto a los corpus, Internet, al ser más amplio que un corpus, ofrece más contextos que un corpus (Maniez 2007: 163; Robb 2003: 20). También en lo que respecta al lenguaje hablado: «The index sites include blogs and discussions which come very close

to spoken language whereas much of the data in formal corpora are from more formal written styles» (Robb 2003: 21).

Asimismo, la tarea de compilación de corpus hemos visto que requiere una gran inversión de tiempo y esfuerzo (Ferraresi 2009) y generalmente los traductores no pueden perder el tiempo en compilar sus propios corpus (Roberts y Bossé-Andrieu 2006: 203). Esto sucede sobre todo en el caso de ámbitos especializados, donde resulta cuando menos complicado encontrar un buen corpus de la temática que se requiere y crearlo requiere una gran inversión de tiempo (Barrière 2009: 1). En este sentido, Internet es más accesible que un corpus (Robb 2003: 21).

Sin embargo, el principal problema que identifican los autores actualmente de utilizar Internet como un corpus es que los buscadores están pensados para recuperar contenidos (páginas web), no formas lingüísticas. Por eso, los resultados de una búsqueda no atienden a motivos relevantes para el traductor (validez lingüística) sino a la adecuación entre la consulta introducida y los supuestos contenidos de una página web.

En este sentido, las funciones de búsqueda que implementan los buscadores, en lo que respecta al acceso a la terminología, son más limitadas que las de los programas de gestión de corpus. Como señala Robb (2003: 21), los buscadores como *Google* solo permiten la búsqueda de palabras exactas, no es posible especificar categorías gramaticales ni acceder a las distintas formas de una palabra, además los resultados no se presentan en un formato fácil de leer. Los buscadores proporcionan, a lo sumo, párrafos con los términos de búsqueda destacados. La manera en que los buscadores muestran los resultados de sus búsquedas no facilita la tarea de detección de terminología como lo hacen los programas de gestión de corpus (Roberts y Bossé-Andrieu 2006: 203; Sánchez-Gijón 2009: 114).

Los programas de gestión de corpus, como hemos visto en apartados anteriores, muestran sus resultados en formato KWIC, con la terminología destacada en el centro de la concordancia, con posibilidad de ordenar las concordancias en función de las palabras que rodean al término de búsqueda a derecha e izquierda, por frecuencia y alfabéticamente. Además, estos programas también ofrecen acceso a colocaciones, listas de palabras, agrupaciones de palabras o n-gramas, etc.

En este sentido, surgen iniciativas para mejorar la búsqueda eficaz de terminología en Internet. Por ejemplo, se diseñan nuevas herramientas que pretenden conjugar la capacidad de búsqueda de los buscadores con las funciones de búsqueda de un corpus con

el fin de optimizar la búsqueda en Internet como un gran corpus, lo que algunos autores han denominado *Web as Corpus*.

En respuesta a esta necesidad se están desarrollando herramientas, que examinaremos a continuación, como *KWiCFinder* (Fletcher 2007a), *Web Concordancer* (Fletcher 2007b), *WebCONC* (Hüning 2001-2009) y *WebCorp* (Renouf et al. 2007), que ofrecen funciones de análisis lingüístico que no se han incorporado todavía en los buscadores tradicionales. También hemos visto en los apartados anteriores que los programas de análisis de corpus apuestan por incorporar aplicaciones que permiten recuperar textos especializados de Internet y posteriormente realizar búsquedas terminológicas en ellos a modo de corpus. Así, observamos que los expertos tratan de mejorar constantemente las búsquedas en Internet.

Aunque la Web es el recurso por excelencia en cantidad e inmediatez de la información, el problema se plantea más bien en cómo gestionar y acceder a dicha información. En el caso de la terminología, los recursos de la Web as Corpus presentan mejoras (ofrecen concordancias en formato KWiC, e incluso algunos disponen de otras funciones como la generación de listas de palabras y colocaciones).

No obstante, los autores también coinciden en que estas herramientas todavía no han alcanzado el nivel que se esperaría para poder analizar la Web como un verdadero corpus. Estos programas siguen utilizando los buscadores tradicionales de Internet (*Google*, *Yahoo!*, etc.) para la recuperación de los textos que analizan y, por tanto, como reconocen sus diseñadores, todavía adolecen de muchos inconvenientes de estos buscadores, como por ejemplo los problemas de fiabilidad del contenido de los textos analizados.

Entre las desventajas de Internet frente a otros recursos también encontramos que la Web es demasiado grande para poder procesarse en su plenitud como se haría con un corpus, por ejemplo con una herramienta de gestión de corpus (Varantola 2006: 221). Además, como la información disponible es casi infinita, el usuario puede verse desbordado y tardar más tiempo del que desearía en encontrar lo que necesita (Friedbichler y Friedbichler 2000: 116). Esta situación ha recibido el nombre de *information overload* o *infoxication*, y es cada vez más frecuente sobre todo desde la proliferación de la Web 2.0, en la que el volumen de información crece exponencialmente (Benito-Ruiz 2009).

Asimismo, los autores desmienten la idea de algunos usuarios de que cualquier información que necesiten la encontrarán en Internet. Codina (1995) se refiere a la *ciber-ingenuidad*. Por ejemplo muchos de los resultados encontrados por un buscador no siempre dan respuesta a la consulta efectuada (García Figuerola 1999: 129).



También existe consenso en que los criterios de selección de las páginas por parte de los buscadores no atienden a factores de calidad, fiabilidad o representatividad, sino más bien al número de visitas que reciben, al número de enlaces a una página, o incluso a criterios económicos.

Por tanto, otra desventaja de Internet reside en que la información no está validada, depende del propio usuario discernir qué información es válida y qué información habría que descartar. Internet presenta problemas de fiabilidad (Friedbichler y Friedbichler 2000: 112). Por tanto, siempre es necesaria una labor de evaluación (Cabré 1999: 35; Gonzalo García 2004: 276; Gonzalo García y Fraile Vicente 2004: 338; Larsen 1997: 89; Palomares Perraut 1999: 181; Pinto y Sales 2007: 534) que tenga en cuenta criterios como el autor de la página, que la lengua en la que está escrita la página sea original, la actualidad de la página, la fiabilidad del servidor que aloja la página, etc. (Alcina et al. 2005: 232).

La escasa fiabilidad de la información que circula en Internet afecta tanto a la información que encontramos en páginas web en general como a diccionarios, glosarios o bases de datos en línea.

Comenzaremos haciendo referencia a los equivalentes electrónicos de los glosarios y diccionarios tradicionales. Hay que empezar diciendo que existen incontables recursos de este tipo en Internet, aunque, en mi experiencia, para el tipo de consultas que pueden interesar a un traductor, y sobre todo a un traductor técnico, difícilmente sustituyen aún al diccionario clásico en papel. Su valor, por otra parte, es muy desigual, porque cualquier particular o empresa puede poner en la Red (que de alguna manera sitúa en pie de igualdad cuanto contiene) lo que le parezca, desde un glosarillo sin contrastar hasta una base de datos refinada a lo largo de muchos años, y se diría que lo primero abunda más que lo segundo. Aparece de nuevo el problema de la fiabilidad al que ya se ha hecho alusión. El traductor, que normalmente no puede permitirse tener sobre su mesa un número elevado de diccionarios especializados en papel, debe resistir la tentación de fiarse, sin más averiguaciones y por el simple hecho de que le resulte fácil acceder a ello, de lo que encuentra en Internet.

(Gómez 1999)

Asimismo, se añade el inconveniente de que el conocimiento más especializado se encuentra escondido o no está accesible directamente desde los buscadores de Internet, «[...] el conocimiento visible es como la punta del iceberg, que oculta una cantidad muy superior de conocimiento no visible» (Bonet Heras 2004: 39), lo que muchos autores denominan la *Web invisible*.

Conviene tener en cuenta que las estimaciones más optimistas calculan que la consulta a través de un buscador permite acceder a menos del 50% del contenido que se encuentra en Internet. Es lo que se conoce como la Web superficial, aquellas páginas que puede indizar un buscador. La Web superficial son páginas estáticas accesibles a los buscadores gracias al sistema de enlaces con otras páginas.

Sin embargo, se estima que la Web profunda o invisible, aquella a la que los buscadores no tienen acceso, es entre 400 y 550 veces más grande que la Web superficial, además crece más rápido y presenta el contenido más valioso y novedoso disponible en línea (Bergman 2001, 2004: 54).

La Web profunda se compone de contenido dinámico, presente en bases de datos, donde normalmente el usuario debe registrarse para su consulta aunque sea de forma gratuita. Además, los buscadores tienen limitaciones técnicas para indizar contenidos en determinados formatos. La información en línea que forma parte de la Web profunda se corresponde, entre otros, con catálogos de bibliotecas, bases de datos bibliográficas y no bibliográficas, obras de referencia como diccionarios, enciclopedias, revistas electrónicas u otros documentos (prensa, legislación, informes, etc.). La consulta de estos recursos debe efectuarse directamente en su base de datos, no a partir de un buscador. En este punto cabe señalar que los problemas de acceso a la Web profunda son más técnicos que económicos, ya que según Bergman el 95% de la Web profunda es de uso público y gratuito.

A todo lo anterior se añade que Internet todavía presenta en ocasiones problemas de conexión (Friedbichler y Friedbichler 2000: 112), lo que puede ralentizar las búsquedas o impedir las.

Asimismo, la inestabilidad de los sistemas de búsqueda supone que, con una misma búsqueda, se recuperan páginas diferentes dependiendo del momento del día o del ordenador desde el que se realice la búsqueda. Además, la virtualidad de Internet provoca una cierta relatividad en las búsquedas, ya que una página recuperada en un momento dado, puede no estar disponible minutos o incluso segundos después.

Finalmente, excepto algunos tutoriales, Internet, como instrumento de búsqueda terminológica, no viene con un manual específico que podamos leer para aprender a utilizar todas las funciones que ofrece, a diferencia de otras herramientas que sí incluyen manuales de uso. Esto se debe también a la variedad de tecnologías que encontramos en los recursos de Internet.

Methods for term mining are grounded in the efficient use of *search engines*. These Web-based tools reflect different technologies, which is one reason why we obtain different results when we use the same search statement with different search engines.

(Budin y Wright 1997: 845)

Lo cierto es que a pesar de todas las ventajas y las desventajas que hemos mencionado, el principal inconveniente de Internet es que se trata de un recurso demasiado grande y poco controlado y los autores coinciden en que, para sacarle provecho y no caer en una búsqueda infructuosa permanente, el usuario debe conocer los lenguajes de

interrogación de los recursos, saber utilizar una serie de técnicas de búsqueda (Corpas Pastor 2004: 224; Gonzalo García 2004: 279).

Por eso, al igual que con los diccionarios electrónicos y los corpus, se requiere incorporar en los programas de formación de futuros traductores y estudiantes de lenguas la enseñanza de técnicas de búsqueda en Internet (Alcina 2003: 80). La mejor forma de potenciar las ventajas y disminuir las desventajas de Internet como recurso terminológico es enseñar a los traductores a utilizar Internet eficazmente (Alcina et al. 2005: 222; Clark et al. 2002: 63; Kübler 2003: 41; Maniez 2007: 162). Por tanto, los traductores deberán aprender el funcionamiento de los buscadores de Internet con el fin de obtener el máximo provecho de las búsquedas.

### **3.4.2 La evolución de la Web y su influencia en las posibilidades de búsqueda**

Desde la proliferación de la Web a principios de los años 90 impulsada por el ingeniero Berners-Lee, a la que él mismo llamó la World Wide Web, la actividad en Internet no ha dejado de crecer.

Tal evolución se está reflejando en las transformaciones de las denominaciones que recibe Internet. Se habla por ejemplo de la Web 1.0, Web 2.0 y Web 3.0, como si fueran las versiones mejoradas de un programa<sup>57</sup>. Por ello, antes de analizar las técnicas de búsqueda que ofrece Internet resulta conveniente explicar qué se entiende actualmente por la Web.

Independientemente de las distintas denominaciones empleadas, se están produciendo desarrollos sobre todo desde la perspectiva de los usuarios, donde se ha pasado de una Web en la que pocas personas determinaban de manera centralizada lo que los usuarios podían utilizar en Internet, a una Web en la que cualquiera puede participar y de manera colaborativa crear y publicar contenido en Internet. La proliferación de los blogs, las wikis y las redes sociales han permitido crear comunidades, redes sociales y enciclopedias en línea, donde los usuarios pueden crear conocimiento e inteligencia colectivos, reunirse y compartir experiencias. Los miembros de estas comunidades y estos recursos compartidos han pasado de emplear estos servicios de manera lúdica a, cada vez

---

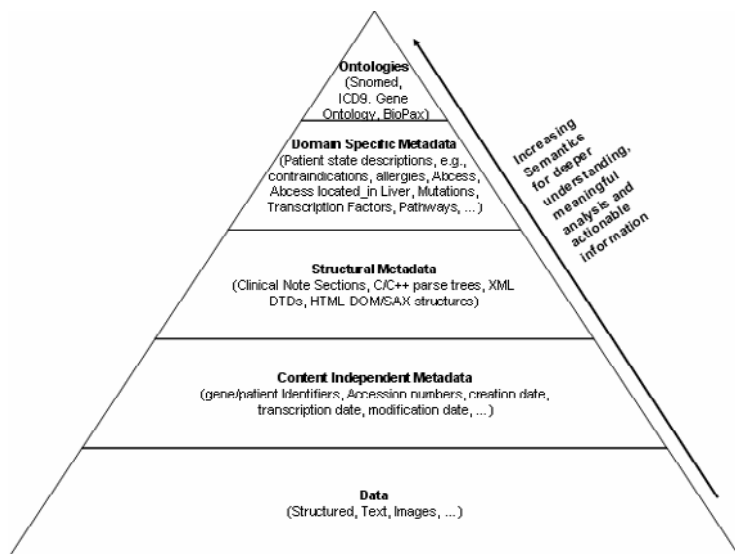
<sup>57</sup> Aunque como reconocen algunos autores estos nombres hay que tratarlos con cuidado: «‘Web 3.1’ is an arbitrary and silly label like ‘Web 2.0’, but even sillier, and 55% cooler! And everyone knows nothing works right until the 3.1 version. This should let us handle the term with care, and should prevent us from overestimating the novelty behind it» (Vossen y Hagemann 2007: 68).

más, aplicarlos a la vida profesional (Vossen y Hagemann 2007: 60). En palabras de Vossen y Hagemann (2007: 66) «[...] while initially content was mostly *read from* the Web, content is meanwhile more and more *written to* the Web. This is why some people speak of Web 2.0 as the “read/write Web”».

Más allá de la Web 2.0 o Web social, se habla de la Web 3.0 o Web semántica, cuya idea es añadir significado a Internet y sus aplicaciones. Estas tecnologías pretenden favorecer la capacidad de los buscadores mejorando las posibilidades de consulta, exploración y navegación (véase la Ilustración 66).

La idea de la Web semántica es que se describa el contenido de las páginas web para que los ordenadores puedan *entender y expresar* el significado. El XML permite al usuario crear sus propias etiquetas y añadir estructura a un documento, pero no explica lo que significa la estructura. El RDF permite expresar el significado de términos y conceptos de forma que el ordenador puede procesarlo. Una de las aplicaciones del RDF es la creación de estándares de metadatos interoperables, según la iniciativa del *Dublin Core Metadata Initiative* (DCMI)<sup>58</sup>, que pueden utilizarse en la descripción de los recursos en línea para facilitar la búsqueda inteligente de información (Vossen y Hagemann 2007: 301).

Las relaciones entre los conceptos y las reglas de razonamiento lógico se definen mediante ontologías (Berners-Lee et al. 2001: 37) con lenguajes como OWL, extensión semántica del RDF (Kashyap et al. 2008: 46).



**Ilustración 66. Semantics-based characterization of metadata, según Kashyap et al. (2008: 30).**

<sup>58</sup> Véase <http://dublincore.org/> para una descripción más detallada de los metadatos.

Los autores sugieren un control terminológico estricto en la indexación de las páginas web. Asimismo, los sistemas de búsqueda avanzan en la línea de ampliar las búsquedas tanto basándose en las formas, lo que Abadal (2004: 416) denomina *búsquedas por patrones*, como en los conceptos, también llamada *búsqueda semántica*<sup>59</sup>. En la búsqueda por patrones el sistema compara el código binario o la apariencia física del término de la consulta con aquellos que tiene almacenados. En el ejemplo que aporta el autor, si el usuario introduce *Eltsin*, el sistema recuperará *Yeltsin*, *Elsin* y *Ieltsin*. En la búsqueda semántica la consulta se amplía a términos relacionados, «[...] derivaciones morfológicas<sup>60</sup>, equivalencias lingüísticas, sinonimia, antonimia, términos generales, específicos, etc.». Para ello, el sistema utiliza «[...] tesauros, diccionarios multilingües o diccionarios especializados previamente elaborados que indiquen cuáles son las relaciones entre los términos» (Abadal 2004: 413). También algunos sitios web tratan de fusionar las tecnologías semánticas y las arquitecturas de participación de los usuarios (Vossen y Hagemann 2007: 68).

Estos avances de Internet (lo que hemos visto que los expertos llaman la Web 2.0 y la Web 3.0) afectan al funcionamiento de los sistemas de búsqueda. Por ejemplo, en lo que respecta a las etiquetas, los buscadores como *Google* en un principio se basaban principalmente en las etiquetas meta del HTML para clasificar las páginas web. Sin embargo, como los diseñadores de páginas web manipulaban las etiquetas meta para mejorar sus posiciones en los resultados de búsqueda, los buscadores empezaron a plantearse nuevos métodos de obtener la información meta, por ejemplo se fijaron en la estructura de enlaces de las páginas web.

En la actualidad, una novedad en el etiquetado son las etiquetas meta creadas manualmente. Los usuarios pueden adjuntar palabras clave a un objeto o recurso web, ya sea un sitio web completo, una foto, un enlace, de manera que se favorece el almacenamiento y la búsqueda (Vossen y Hagemann 2007: 183).

Según Kashyap et al. (2008: 26-29), las etiquetas meta facilitan la manipulación y recuperación de contenido digital, por ejemplo, en la navegación y recuperación de colecciones de imágenes basándose en el contenido semántico. Así, en la imagen de un

---

<sup>59</sup> Otros autores como Woods (2004) se refieren a las relaciones morfológicas y semánticas.

<sup>60</sup> Según Woods (2004) la técnica de *stemming* permite indexar los documentos por las raíces de las palabras que contienen y, para la recuperación de los mismos, el sistema solo tiene en cuenta las raíces de los términos de la consulta. De este modo, en el caso de que existieran dos textos en línea y en uno apareciera *acid glass* y en el otro *acidic glass*, el buscador los relacionaría, indexaría y recuperaría conjuntamente.

tumor cerebral los conocimientos del contenido espacial de la imagen y la forma en que cambia con el paso del tiempo pueden representarse con anotaciones de metadatos. También puede emplearse la información meta para buscar videos, archivos de audio o discursos, para la gestión de documentos estructurados, por ejemplo para la recuperación a partir de información contenida en el título de un documento, en sistemas de información geográfica y ambiental, en bibliotecas digitales y en sistemas de gestión del conocimiento. Las etiquetas meta pueden ser *Content-Independent Metadata* no dependientes del contenido del documento, por ejemplo el número de registro de una muestra de laboratorio, identificadores de localización como los URI, *Content-Based Metadata* que pueden ser *Structural Metadata*, por ejemplo que identifique las secciones de un informe médico, y *Domain-Specific Metadata* donde el vocabulario es muy importante, ya que los metadatos deben corresponderse con términos de un dominio específico. Dentro de este último tipo de metadatos, encontramos *Intra-domain-specific Metadata* que engloban relaciones y asociaciones entre datos de un mismo dominio e *Inter-domain-specific Metadata* que incluyen las relaciones y asociaciones entre diferentes dominios.

Además, otro avance es que estas etiquetas pueden compartirse en comunidad, se ponen a disposición de otros usuarios para que cuando estos busquen a través de dichas etiquetas puedan adoptarlas para su propio uso. Esto es lo que los autores denominan *folksonomies*, «[...] collaboratively created open-ended labeling systems» (Breslin et al. 2009: 38; Vossen y Hagemann 2007: 184).

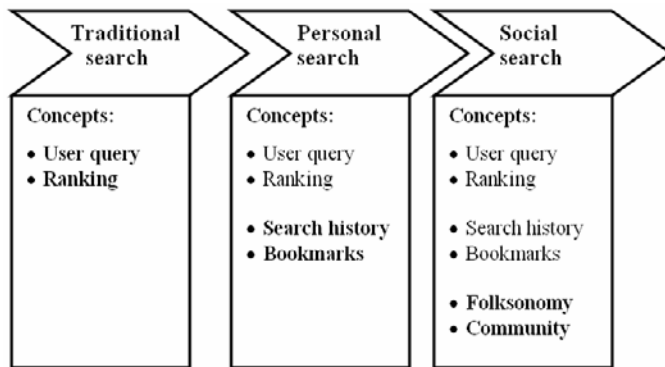
Asimismo el *social bookmarking* o los *marcadores sociales* pueden influir en la búsqueda. La página que otros usuarios han considerado útil o relevante en relación con un tema no tiene por qué coincidir con las páginas que un algoritmo de búsqueda ha situado en las primeras posiciones de los resultados (Vossen y Hagemann 2007: 195). Las ventajas de los marcadores sociales se vislumbran tanto en la traducción profesional como en la enseñanza de la traducción<sup>61</sup>, donde traductores y estudiantes de traducción pueden compartir enlaces lo que facilita sobremanera la tarea de documentación (Enríquez Raído 2011: 305; Muñoz Sánchez 2007: 170).

---

<sup>61</sup> En una época donde prima el trabajo cooperativo y en equipo, la Web 2.0 ha supuesto un enorme avance social, donde los traductores se han visto involucrados en lo que podría conocerse como la traducción social, participando activamente por ejemplo en proyectos de traducción de software libre, traducción activista como la del proyecto Babels, plataformas colaborativas de ayuda a traductores e intérpretes como foros y recursos creados colaborativamente (enciclopedias, diccionarios, glosarios), etc. Aunque también se advierte a los traductores de los peligros de perderse en esta nueva era social, donde en ocasiones prevalece la inmediatez y la gratuidad sobre la calidad (Estella Clota 2010).

En este sentido, la búsqueda social se basa en el dilema de que, por un lado, los sistemas de búsqueda no son inteligentes para decidir de modo automático y evaluar la información que contiene un sitio web. Por otro lado, un usuario sí que es capaz de decidir qué información es útil en cada situación pero no es capaz de procesar los millones de resultados que proporciona un buscador en cada búsqueda.

Por ello, como señalan Vossen y Hagemann (2007: 266), cada vez son más los intentos de personalizar las búsquedas (véase la Ilustración 67). Por ejemplo, a un buscador el historial de búsquedas o los marcadores de un usuario pueden ayudarle en la recuperación de páginas web útiles para el usuario, según las búsquedas que ha efectuado anteriormente y los sitios web que ha visitado o considera útiles. Un paso más es que los buscadores tengan en cuenta la información etiquetada con palabras clave, valoraciones personales, comentarios, clasificaciones, etc. Además, cuando el etiquetado se produce en comunidad, los expertos de un campo comparten sus conocimientos con el resto de usuarios.



**Ilustración 67. Pasos de evolución en la búsqueda (Vossen y Hagemann 2007: 268).**

Sin embargo, la explotación de estos metadatos proporcionados por ejemplo por las *folksonomies* generan también problemas para su procesamiento automático por cuestiones semánticas como la homonimia y la polisemia y aquí es donde las técnicas de la Web semántica entran en juego para aportar soluciones, poniendo en conjunción las características disponibles de la Web 2.0 o Web social (marcadores sociales, etiquetado social, *folksonomies*) con los desarrollos de la Web semántica (taxonomías, ontologías) (Vossen y Hagemann 2007: 283).

[...] the Social Web and social networking sites can contribute to the Semantic Web effort. Users of these sites often provide metadata in the form of annotations and tags on photos, ratings, blogroll links, etc. In this way, social networks and semantics can complement each other. Already within online communities, common vocabularies or folksonomies for tagging are emerging through a consensus of community members.

(Breslin et al. 2009: 19)

Finalmente, como señalan Breslin et al. (2009: 283) «[...] we are only now able to realise the vision of a Web that is both socially aware and machine understandable: in fact, this was essentially the vision of Berners-Lee when he designed the original Web. Hence, it should not be considered as either Web 3.0 or Web n.0 but simply as the Web, the only one».

### 3.4.3 Análisis empírico de técnicas de búsqueda en Internet

En este apartado efectuamos un análisis de algunas herramientas que permiten efectuar búsquedas de terminología en Internet<sup>62</sup>. En este sentido, nos centraremos en las consultas y no en la exploración y navegación, según lo entienden la mayoría de autores (Baeza-Yates y Ribeiro-Neto 1999: 265; Vossen y Hagemann 2007: 10-14).

En un principio lo que dio popularidad a Internet fueron las páginas web, en HTML, que se caracterizan por los hiperenlaces, que permiten que unas páginas estén conectadas con otras y que el usuario pueda realizar una búsqueda por navegación (*navigating*) a través de ellas con la ayuda de un navegador.

Pronto, debido al ingente número de páginas disponibles en línea se reconoce la necesidad de organizarlas, a través de *directorios*, catálogos que organizan los sitios web jerárquicamente (Vossen y Hagemann 2007: 19) y que facilitan al usuario la exploración (*browsing*) dentro de estas estructuras de información para acceder a páginas web que se corresponden con una misma temática.

Sin embargo, si el usuario busca información específica, los directorios no resultan demasiado satisfactorios, ya que no está claro en qué categoría encontraremos un término concreto. La manera más directa y rápida de acceder a la información disponible en Internet ha sido hasta el momento la realización de consultas (*querying*) con *buscadores* que han tenido tanto éxito que: «“To search” the Web is nowadays often identified with “to Google” the Web [...]»<sup>63</sup> (Vossen y Hagemann 2007: 13).

Según lo que acabamos de explicar, empezaremos nuestro análisis con la descripción de los modos de búsqueda de algunos buscadores, como *Google* y *Yahoo!*. Los buscadores

---

<sup>62</sup> No incluimos aquí la búsqueda directa en recursos lingüísticos disponibles en línea, como diccionarios, glosarios, bases de datos o corpus, aunque también son recursos en los que el traductor puede efectuar búsquedas terminológicas en línea. Ya hemos examinado recursos de este tipo en los apartados anteriores 3.2.3 y 3.3.3.

<sup>63</sup> En la práctica la distinción entre *directorio* y *buscador* es difusa ya que la mayoría de portales de Internet ofrecen ambos servicios, aunque el que prevalece es el de buscador, mientras que el de directorio se encuentra más oculto (dir.*Yahoo.com*, dir.*Google.com*).



se sirven de las palabras clave introducidas por el usuario en el buscador, buscan dichas palabras clave en las páginas que tiene indexadas en su base de datos, y ofrecen al usuario un listado de páginas que contienen las palabras introducidas.

Asimismo, analizaremos un nuevo tipo de buscadores de Internet que funcionan a modo de herramientas de gestión de corpus (Web as Corpus), como *WebCorp*, *WebCONC*, *Web concordancer* y *KWiCFinder*. La diferencia principal de estas herramientas y los buscadores radica principalmente en el modo en que se presentan los resultados de la búsqueda.

Los buscadores recuperan páginas web, en las que se encuentra el término o expresión de búsqueda. Normalmente, en los resultados de la búsqueda, debajo de cada una de las páginas recuperadas aparece un resumen de la página con las palabras de búsqueda resaltadas. Para acceder al contenido de la página y observar el uso de las palabras de búsqueda en contexto, el usuario tiene que acceder a cada una de las páginas ofrecidas por el buscador. En resumen, los buscadores no son herramientas específicamente diseñadas para realizar búsquedas lingüísticas.

En cambio, las herramientas de búsqueda en Internet como corpus están pensadas para facilitar las búsquedas lingüísticas en páginas de Internet. Para la recuperación de las páginas analizadas se sirven de los buscadores de Internet, pero los resultados de la búsqueda se visualizan a modo de concordancias KWiC, y algunos de estos recursos ofrecen otras opciones de los programas de gestión de corpus, como generación de colocaciones o de listas de palabras. Estas herramientas permiten buscar terminología en Internet como si fuera un gran corpus.

Finalmente, en apartados anteriores hemos visto que cada vez son más los programas de gestión y análisis de corpus que facilitan la rápida descarga y compilación de textos de Internet para la creación de corpus virtuales y su posterior análisis<sup>64</sup>.

#### 3.4.3.1 Buscadores

En este apartado analizamos<sup>65</sup> en primer lugar el buscador *Google* por ser actualmente el más utilizado y después examinamos *Yahoo!* que ha implantado su interfaz de búsqueda en otros buscadores como *Altavista* y *AlltheWeb*<sup>66</sup>.

---

<sup>64</sup> Véase el análisis de los recursos *WordSmith Tools 4*, *Corpógrafo*, *TerminoWeb*, *Jaguar* o *WebBootCat* y *Sketch Engine* en el apartado 3.3.3.2

<sup>65</sup> Última revisión del análisis 20 de abril de 2012. Es posible que desde esta fecha algunas funciones

Para el análisis de *Google* (www.google.com) nos basamos en su interfaz de búsqueda en español. Este buscador permite realizar dos tipos de búsqueda: la simple y la avanzada. En la página inicial de *Google* se puede realizar una búsqueda por palabras clave. En este tipo de búsqueda, el usuario puede introducir palabras clave relacionadas con el término o tema que busca. Con el conocimiento de una serie de operadores es posible combinar las palabras de búsqueda para que se tengan en cuenta todas (con el uso del signo + delante de cada una), que se descarte alguna (colocando el signo – delante), o que se busque una expresión exacta (utilizando *comillas*). También es posible emplear comodines dentro de las comillas para sustituir palabras en la búsqueda de una secuencia de términos (pueden utilizarse tantos asteriscos seguidos como palabras quieran sustituirse). También es posible emplear el símbolo | en búsquedas como “tomar|adoptar una decisión” que recupera tanto *tomar una decisión* como *adoptar una decisión*. Otros operadores funcionan a modo de filtro para las páginas recuperadas. *Google* utiliza los siguientes operadores filtro.

- cache: es un operador que recupera la página que introduzcamos a continuación de este operador, si se encuentra almacenada en la caché de *Google*. Asimismo, con el uso de este operador si, además de una dirección URL, introducimos otras palabras de búsqueda, *Google* las resalta con colores en la página que recupera<sup>67</sup>. Por ejemplo, si introducimos *cache:www.tecnociencia.es/entidades/itc/tcn\_itc4.html prensado ceramica*, se recupera esta página del Instituto de Tecnología Cerámica, que se encuentra en la caché de *Google*, y se destacan las palabras *prensado* y *cerámica* cada vez que aparecen en el documento.

- link: recupera páginas vinculadas a la dirección URL que indiquemos.

- related: recupera páginas relacionadas con la dirección URL que introduzcamos.

- info: ofrece información sobre una página.

- define: funciona a modo de diccionario monolingüe para definir la palabra que introduzcamos a continuación del operador.

- site: restringe los resultados de la búsqueda a la página o dominio que indiquemos.

---

de búsqueda hayan cambiado.

<sup>66</sup> En análisis anteriores estos buscadores contaban con una interfaz de búsqueda propia (Pastor 2008b; Pastor y Alcina 2011).

<sup>67</sup> Función de *Google* citada por Vargas Sierra (2002) para destacar las palabras clave de la búsqueda en contexto, como en la búsqueda de concordancias en un corpus.

- *allintitle*: restringe los resultados de la búsqueda a los títulos de los documentos. Por ejemplo *allintitle:informática Google* recuperará páginas que contengan tanto *informática* como *Google* en el título.

- *intitle*: restringe la búsqueda a los títulos solo para la primera palabra de la búsqueda. Por ejemplo, *intitle:informática Google* recuperará páginas que contengan *informática* en el título y *Google* en cualquier lugar del documento.

- *allinurl*: restringe los resultados de la búsqueda a direcciones URL en las que aparezcan las palabras clave que indiquemos. Por ejemplo, *allinurl:Google search* buscará direcciones URL que contengan ambas palabras.

- *inurl*: es una variante de *intitle*:, pero aplicada a las URL. Por ejemplo *inurl:Google search* recuperará direcciones en las que aparezca la palabra *Google* solamente y *search* en el resto del documento.

Normalmente, no puede haber espacio entre los dos puntos del operador filtro y la primera palabra de la búsqueda.

Asimismo, es posible buscar intervalos de números, añadiendo dos puntos entre las dos cifras del intervalo, por ejemplo, *125..512 Mb*. Además, con la introducción de una vírgula ~ delante de una palabra, *Google* busca también sinónimos de dicha palabra, aunque esta función está solo parcialmente desarrollada en las búsquedas en inglés y no está implementada en otras lenguas.

La búsqueda avanzada de *Google* permite al usuario restringir la búsqueda en función de una serie de criterios sin necesidad de tener conocimientos de lógica booleana y operadores. Este tipo de búsqueda posibilita, en primer lugar, la limitación de los términos introducidos en el buscador para la búsqueda *con todas las palabras*, *con la frase exacta*, *con alguna de las palabras* y *sin las palabras* que introduzcamos; también un intervalo de números.

Además, se puede indicar a la herramienta que busque páginas escritas en el idioma que deseemos, páginas de una región en concreto, con un formato de archivo determinado, por ejemplo en pdf, o publicadas en un intervalo de tiempo. Otra de las opciones que ofrece es que las palabras de la búsqueda aparezcan en páginas de un dominio, y también en el lugar de la página que especifiquemos (en cualquier lugar de la página, en el título, en el contenido, en la dirección o en los enlaces).

*Google* también presenta otras herramientas como la búsqueda de *Imágenes* y *Sitios con imágenes*, con la que en los resultados únicamente aparecen páginas web que contengan imágenes.

Con la opción *Búsquedas relacionadas* el buscador sugiere palabras de búsquedas relacionadas en forma de lista. *Intervalo personalizado* ordena las páginas recuperadas en el intervalo de tiempo que indiquemos, *Páginas visitadas* únicamente busca en aquellas páginas que figuren en el historial de búsqueda y *Páginas no visitas* aquellas que no. La opción *Diccionario* busca definiciones de la palabra de búsqueda introducida. Finalmente, ofrece la posibilidad de filtrar los resultados recuperados a imágenes, videos, noticias (páginas de periódicos principalmente), libros (búsqueda en *Google Books* o *Google Libros*), lugares (búsqueda en *Google Maps*), blogs y foros de debate.

Asimismo, *Google Scholar* o *Google Académico* busca bibliografía especializada en diferentes campos y fuentes disponibles en línea, principalmente estudios revisados por especialistas, tesis, libros, resúmenes y artículos de editoriales académicas, sociedades profesionales, depósitos de impresiones preliminares, universidades y otras organizaciones académicas.

*Google Trends*<sup>68</sup> permite observar la distribución por dominio geográfico y por fecha de una determinada expresión.

Anteriormente, con la *Rueda de búsquedas* el buscador a partir de una palabra generaba un grafo con palabras de búsqueda relacionadas, como vemos en la Ilustración 68, al introducir la palabra *sequía*, se proponían las siguientes palabras de búsqueda relacionadas *inundaciones*, *aridez*, *agua*, *sequía diccionario*, *sequía definición*, *fotos sequía*, *sequía en España* y *letra sequía*.



Ilustración 68. Resultados de la *Rueda de búsquedas* a partir de la palabra *sequía* en *Google*.

<sup>68</sup> Véase <http://www.google.es/trends>

Antes también, **Google Sets** buscaba términos relacionados por grupos de palabras (Bergman 2004: 37). Por ejemplo, con la introducción de las palabras *sandía* y *melón*, se obtenían los nombres de otras frutas, como *manzana*, *piña*, *fresa*, *naranja*, *limón*, etc.

**Yahoo!** (www.yahoo.com) ha modificado las interfaces de los buscadores **Altavista** (www.altavista.es) y **AlltheWeb** (www.alltheweb.com) y, actualmente, los tres buscadores presentan una misma interfaz de búsqueda. Dicha interfaz ha adoptado algunas funciones de los buscadores *Altavista* y *AlltheWeb*, como los operadores booleanos: AND, OR, NEAR, + y -. NEAR busca documentos en los que las palabras introducidas aparezcan en una posición cercana entre sí. Por tanto, las expresiones de búsqueda *agua NEAR presión* recuperarán *agua a presión*, *presión de agua*, *presión de vapor del agua*, etc. Asimismo, acepta el empleo de comillas para recuperar citas literales, y el uso de asteriscos dentro de las comillas para sustituir palabras en la búsqueda de una secuencia de términos (pueden utilizarse tantos asteriscos seguidos como palabras quieran sustituirse). Es posible restringir la búsqueda a páginas de uno o varios idiomas y páginas del país que especifiquemos. También permite restringir la búsqueda a imágenes, videos y noticias.

A su vez, utiliza los siguientes operadores filtro:

- site: recupera las páginas que se corresponden con el dominio indicado. Por ejemplo, *site:uk*, *site:es*, *site:rae.es*, etc.

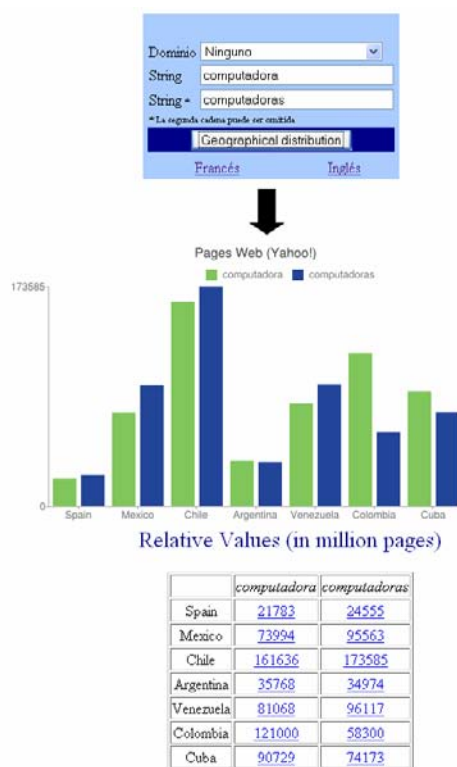
- url: funciona de manera similar al operador anterior, restringe los resultados a páginas de la dirección URL que especifiquemos. Por ejemplo, *url:europa.eu.int* limitará los resultados a páginas de la Unión Europea.

- intitle: encuentra las páginas que contienen la expresión de búsqueda en el título. Por ejemplo, *intitle:industria cerámica* encontrará páginas que contengan en el título la expresión *industria cerámica*, y también solo *industria* o solo *cerámica*.

Las funciones de búsqueda avanzada de *Yahoo!* permiten buscar con todas las palabras que se indiquen, con una secuencia exacta, con cualquiera de las palabras que se introduzcan o con ninguna de ellas. Contiene la opción de restricción por países o por lenguas. Asimismo, es posible restringir la búsqueda a un formato de archivo y a una URL específica o dominio de Internet.

Además, se ha incorporado una función de sugerencias, o también *búsquedas relacionadas*.

*Diatopix*<sup>69</sup> es una herramienta que permite observar la distribución geográfica de una determinada expresión dentro de un idioma. Los resultados se basan en la búsqueda con *Yahoo!*. Es recomendable introducir la expresión tanto en singular como en plural y se puede restringir por área temática. Funciona con el idioma francés (Canadá, Francia, Bélgica, Suiza y Luxemburgo), español (España, Méjico, Chile, Venezuela, Colombia y Cuba) e inglés (Canadá, Estados Unidos, Reino Unido, Irlanda, Australia y Nueva Zelanda). Por ejemplo, es posible comprobar la frecuencia de la palabra *computadora(s)* en páginas de países de habla hispana. Observamos que esta palabra aparece con mayor asiduidad en páginas de Chile y con menor en dominios de España.



**Ilustración 69. Distribución geográfica de *computadora(s)* en *Diatopix*.**

En Internet encontramos otros buscadores como *About* ([www.about.com](http://www.about.com)), *Ask Jeeves* ([www.ask.com](http://www.ask.com)), *Bing* ([www.bing.com](http://www.bing.com)), *Lycos* ([www.lycos.es](http://www.lycos.es)) *Teoma* ([www.teoma.com](http://www.teoma.com)), *Terra* ([www.terra.es](http://www.terra.es)), entre otros.

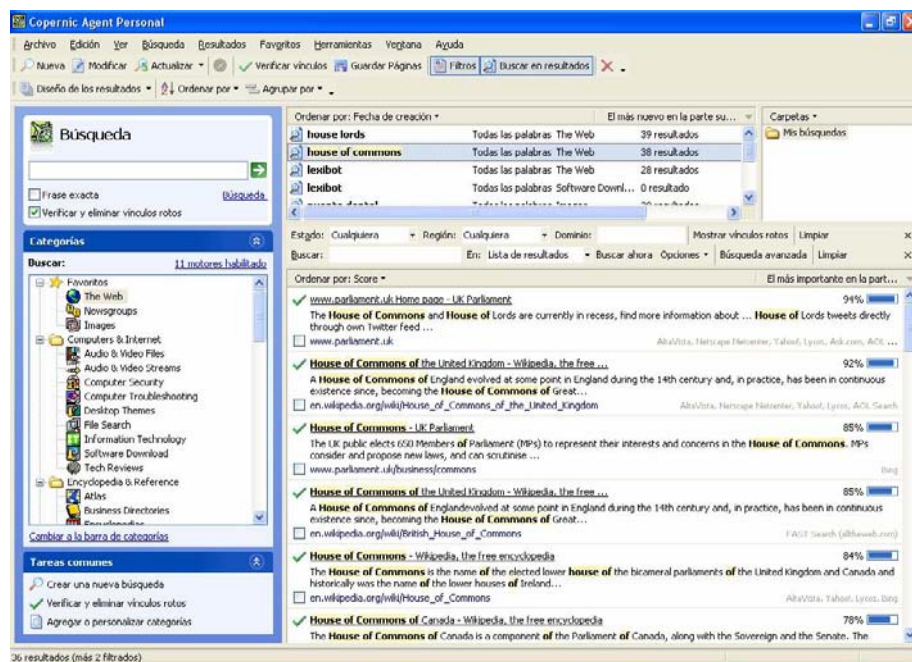
También existen metabuscadores, que permiten buscar en varios motores de búsqueda al mismo tiempo, como *Beaucoup* ([www.beaucoup.com](http://www.beaucoup.com)), *Buscopio* ([www.buscopio.net](http://www.buscopio.net)), *Dogpile* ([www.dogpile.com](http://www.dogpile.com)), *FinderSeeker* ([www.finderseeker.com](http://www.finderseeker.com)),

<sup>69</sup> Desarrollada por Drouin, grupo OLST de la Universidad de Montreal.

*Metacrawler* (www.metacrawler.com), *Metasearch* (www.metasearch.com), *Oneseek* (www.oneseek.com), *Search Engine Guide* (www.searchengineguide.com) y *Webcrawler* (www.webcrawler.com).

Asimismo, algunos metabuscadores denominados *agentes inteligentes de búsqueda*, como *Copernic Agent*, *LexiBot* o *Wisenuit*, pueden buscar información en red de forma autónoma de acuerdo con unos parámetros o necesidades especificadas por el usuario, además aprenden y mejoran los resultados de sus búsquedas en función del comportamiento del usuario.

A modo de ejemplo, *Copernic Agent Personal*<sup>70</sup> permite la consulta simultánea en varios motores de búsqueda clasificados por categorías, en función del tipo de información que interese buscar al usuario. Tiene una opción con la que se descartan de la lista de resultados aquellas páginas que ya no se encuentran disponibles en línea (vínculos rotos). Almacena los resultados de cada búsqueda por carpetas de manera que es posible acceder al historial de las búsquedas y a sus resultados en cualquier momento. También es posible realizar nuevas búsquedas dentro de las páginas recuperadas.



**Ilustración 70.** Interfaz de búsqueda de *Copernic Agent Personal*.

La versión *Copernic Agent Professional* incluye otras mejoras como la realización de tareas automáticas: programación de búsquedas periódicas, análisis automático de

<sup>70</sup> Este programa es de descarga gratuita en <http://www.copernic.com/>

resultados y almacenamiento en el disco duro de las páginas recuperadas, así como informar al usuario de las búsquedas. El análisis de resultados permite la detección del idioma de las páginas, de páginas duplicadas, de la última modificación de una página y la extracción de conceptos clave de las páginas.

### 3.4.3.2 Herramientas de análisis de la Web como corpus

Aunque hemos visto que Internet puede considerarse una de las herramientas de búsqueda más utilizadas actualmente por los traductores, todavía presenta inconvenientes en cuanto a las búsquedas de terminología, ya que no ordena ni clasifica las palabras de búsqueda y los contextos. En este sentido, por ejemplo Friedbichler y Friedbichler (2000: 116) abogan por el desarrollo de buscadores que muestren los resultados en modo KWIC pero con la misma velocidad de los buscadores actuales.

*WebCorp*<sup>71</sup> (Renouf et al. 2007) es una herramienta de búsqueda disponible en línea que permite realizar búsquedas en Internet a modo de corpus. Utiliza como fuente para realizar sus búsquedas las páginas indexadas en los buscadores *Bing* y *Google*. El usuario puede seleccionar con qué buscador desea que se recuperen las páginas para el análisis.

*WebCorp* incluye dos pestañas de búsqueda, *Search* y *Wordlist Tool*. La búsqueda simple de la página de inicio permite la introducción de una palabra exacta o una secuencia de palabras (pueden utilizarse asteriscos para sustituir términos en una secuencia de palabras, y corchetes junto con el operador “|” para incluir alternativas de palabras o de caracteres). Por ejemplo, [el|la] niñ[o|a] llor[ó|a] debería recuperar las combinaciones *el niño llora, el niño lloró, la niña llora y la niña lloró*. Asimismo, se selecciona el número de caracteres o términos que queremos que se muestren en las concordancias a cada lado de la palabra pivote (de 10 caracteres a 10 palabras). Podemos elegir que no se tenga en cuenta el uso de mayúsculas y minúsculas.

Además, si se accede a las opciones avanzadas (véase la Ilustración 71), se pueden especificar palabras de búsqueda que estarán ausentes en los resultados, se selecciona el número de páginas que se recuperarán, se especifica si se visualizarán o no las URL, si solo se recuperará una concordancia por página y también pueden indicarse URL o dominios de países para restringir la búsqueda. El modo de visualización de los resultados es de concordancias KWIC, con la palabra de búsqueda destacada en el centro de los

---

<sup>71</sup> *WebCorp* está disponible en la página <http://www.webcorp.org.uk/>



contextos. También es posible acceder al texto completo del que se ha extraído la concordancia haciendo clic en la opción *Text* (se visualiza solo el texto de la página con las palabras clave de búsqueda destacadas), o haciendo clic en cada una de las palabras pivote de las concordancias se accede a la página web de las que se han extraído los contextos y las palabras de búsqueda también aparecen destacadas.

The screenshot shows the WebCorp Live search interface. At the top, there is a navigation bar with links for Search, Wordlist Tool, User Guide, WebCorp LSE, Publications, and Feedback. Below this is a search form with the following options:

- Search:** A text input field.
- Case insensitive:** A checked checkbox.
- Span:** A dropdown menu set to "10 characters".
- Search API:** A dropdown menu set to "Bing".
- Language:** A dropdown menu set to "Not specified".
- Show URLs:** A checked checkbox.
- Pages:** A dropdown menu set to "50".
- One concordance line per web page:** An unchecked checkbox.
- Site:** A text input field with a list of popular sites below it:
 

UK Broadsheet Newspapers	US Academic	France
UK Tabloid Newspapers	UK Academic	Germany
French Newspapers		Italy
Greek Newspapers	Argentina	Japan
US Newspapers	Australia	Netherlands
BBC News	Brazil	New Zealand
Wikipedia	Canada	Spain
	China	UK
- Word Filter:** A text input field.

At the bottom of the form, there are two buttons: "Restablecer" and "Search". Below the form, there is a footer with the text: "By using the WebCorp tools you are agreeing to be bound by the Terms of Use and Privacy Policy. Copyright © 1999-2012 Research and Development Unit for English Studies, Birmingham City University."

### Ilustración 71. Opciones de búsqueda avanzada de WebCorp.

En la Ilustración 72 mostramos los resultados de una búsqueda en la que hemos introducido en *WebCorp* la palabra *hormigón* con la selección de idiomas español y la restricción a páginas con dominio (.es). Aparece una muestra de las concordancias generadas. Al lado de cada página con las concordancias, se da la posibilidad de generar una lista de palabras de dicha página (*Word List*).

## Results for query "hormigón"

case insensitive,  
using the Google API,  
from sites: .es

1) <http://www.upv.es/entidades/ICITECH/index.html>

Text, Wordlist, text/html, ISO8859\_1 (HTML source), date unknown

1: Información Instituto de Ciencia y Tecnología del **Hormigón** Información de contacto Servidor WWW <http://www.1>  
2: .es Dirección postal Inst. de Ciencia y Tecn. del **Hormigón** Camino de Vera, s/n 46022 Valencia Teléfono +34 9

2) [http://www.todo propiedades.com.es/informacion/historia\\_hormigon.htm](http://www.todo propiedades.com.es/informacion/historia_hormigon.htm)

Text, Wordlist, text/html, ISO8859\_1 (HTML source), 2011-10-19 (Server header)

3: Historia del **Hormigón** Mezclas para Construcción INICIOCASAPISOFICINA  
4: res españoles de materia prima para construcción. **Hormigón** Todo sobre el material más resistente para edific  
5: piedras de calidad normada. Primeros en Utilizar **Hormigón** No se tiene certeza quien descubrió o utilizó por  
6: teza quien descubrió o utilizó por primera vez el **Hormigón**. Es probable que al mismo tiempo que el hombre de  
7: dominó el fuego también descubrió el concepto de **hormigón**. Uno puede imaginar al hombre primitivo junto a s  
8: y agua. Esta mezcla puede ser considerada como un **hormigón** rudimentario. Los egipcios por su parte, utilizar  
9: alcinada que mezclaron con arcilla cocida y agua. **Hormigón** Romano El pueblo romano también usó **hormigón** en e  
10: gua. **Hormigón** Romano El pueblo romano también usó **hormigón** en sus construcciones, para lo cual utilizaron ca

3) <http://www.agenciasinc.es/Noticias/La-historia-del-hormigon-armado-en-el-siglo-XX-recopilada-en-una-tesis>

Text, Wordlist, text/html, UTF8 (Content-type), 2011-10-25 (Server header)

11: por qué se utilizó este material La historia del **hormigón** armado en el siglo XX, recopilada en una tesis No  
12: oy poseemos un extenso patrimonio de elementos de **hormigón** armado, pero apenas hace un siglo no existía ning  
13: a los rasgos de las construcciones levantadas con **hormigón** armado en cada década, y ha explicado cómo y por  
14: emarcado la autora, en Guipúzcoa la expansión del **hormigón** armado estuvo muy ligada a la industrialización.  
15: icado en dos grupos las primeras edificaciones de **hormigón** armado. Por una parte, las construcciones totalme  
16: te, las construcciones totalmente funcionales: el **hormigón** se encuentra a la vista, ya que no resulta necesa  
17: ser bellas y funcionales, por lo que se cubrió el **hormigón**. El puente de María Cristina, la sede del Banco G  
18: y teatro calcinadas por los incendios, al ser el **hormigón** un material incombustible. Libre de patentes El h  
19: n un material incombustible. Libre de patentes El **hormigón** armado quedó pronto libre de patentes, por lo que  
20: tradicional local, por lo que se construyeron con **hormigón** armado edificios que emulaban caseríos y palacios  
21: ientras, en Europa, a medida que se sabía más del **hormigón** armado surgieron nuevas expresiones arquitectónic  
22: s ingenieros llevaron hasta el extremo el uso del **hormigón** armado. Las construcciones resultaban resistentes  
23: partir de 1960 se retrocedió en la evolución del **hormigón** armado, con lo que hubo una pausa en su desarroll  
24: arquitectura e ingeniería siguen haciendo uso del **hormigón**. Se ha recurrido a la nanotecnología para obtener  
25: llas propiedades que se le atribuyen en su día al **hormigón**: hormigones translúcidos, aligerados, flexibles.  
26: que el reto ahora es conseguir una producción de **hormigón** sostenible. Se apunta al reciclaje y a la eficien  
27: oducción, lo cual puede suponer la renovación del **hormigón** armado y, en consecuencia, el surgimiento de nuev  
28: rquitectónica guipuzcoana unida al desarrollo del **hormigón** armado). Localización: País Vasco Comentarios (0)

## Ilustración 72. Concordancias generadas con WebCorp.

Una vez efectuada la búsqueda, *WebCorp* permite utilizar un filtro de fechas, en el que se pueden limitar los resultados a páginas publicadas en un intervalo de tiempo o un intervalo de fechas. También permite ordenar las concordancias por fecha, o alfabéticamente por la expresión de búsqueda o por palabras en posición 1 a 9 a derecha o izquierda de la palabra pivote. Asimismo, es posible recuperar colocaciones. La función *exclude stopwords* elimina de la lista de colocaciones las palabras vacías de significado, como determinantes o preposiciones.

La pestaña *Wordlist Tool* de *WebCorp* permite generar listas de palabras de la página web que indiquemos, y ordena las palabras de la lista alfabéticamente o por frecuencia. Es posible utilizar la opción *Filter Out Stopwords* (para excluir palabras vacías de significado) y especificar un número de n-gramas que deben contener los resultados de la lista (de 1 a 5 palabras).

*WebCONC*<sup>72</sup> es una herramienta desarrollada por Hüning (Freie Universität Berlin). Ofrece dos tipos de búsqueda. En el primer tipo de búsqueda, *WebCONC* utiliza *Google* para la recuperación de las páginas de las que genera las concordancias. En este caso, el usuario puede introducir una palabra o secuencia exacta. Además, permite especificar el

<sup>72</sup> *WebCONC* estaba disponible en la página <http://www.niederlandistik.fu-berlin.de/cgi-bin/web-conc.cgi?sprache=en>

número máximo de páginas que se desea recuperar (10, 20 o 50), el idioma de la búsqueda y el número de caracteres máximos que deberán aparecer en la concordancia a cada lado de la expresión de búsqueda (25, 50; 75 o 100). Véase la Ilustración 73.

### Ilustración 73. Interfaz principal de búsqueda de *WebCONC*.

En el segundo tipo de búsqueda, el usuario puede restringir el corpus de búsqueda indicando las URL exactas en las que desea que se busquen las concordancias. Esta opción permite también la utilización de expresiones regulares en la búsqueda. Por ejemplo, podemos comparar cómo se escribe Irak en diferentes periódicos del mundo para comprobar cómo es la forma de escritura correcta en cada lengua. En el ejemplo de la Ilustración 74 introducimos la expresión regular `Iraq|Irak` y restringimos la búsqueda a las siguientes páginas: `http://www.elpais.es`, `http://www.elmundo.es`, `http://www.guardian.co.uk`, `http://www.timesonline.co.uk`, `http://www.lemonde.fr`, `http://www.lefigaro.fr`.

### Ilustración 74. Restricción del corpus de búsqueda en *WebCONC*.

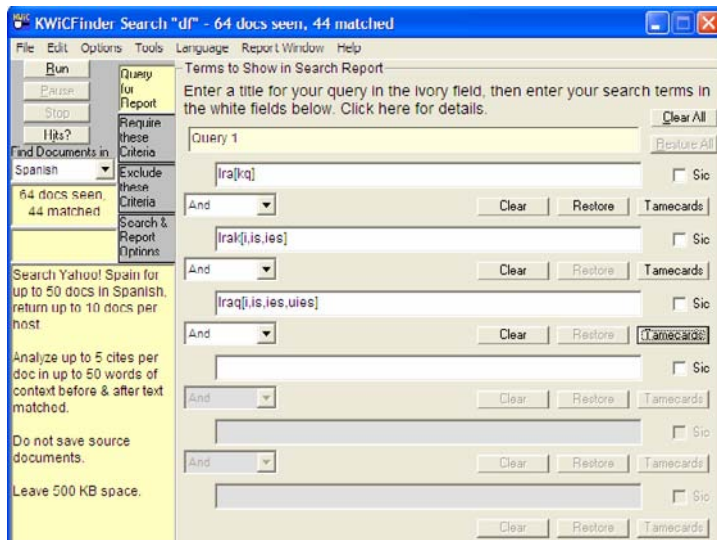
*KWiCFinder* (Fletcher 2007a) también es un programa de búsqueda de concordancias en Internet. A diferencia de *WebCorp* y *WebCONC*, no está disponible en línea, es necesaria su instalación en el ordenador, aunque puede descargarse gratuitamente

desde Internet, registrándose para obtener una clave que permite activar el producto una vez instalado<sup>73</sup>.

Esta herramienta utiliza como fuente para la extracción de concordancias las páginas indexadas por el buscador de *Yahoo!*. Tiene cuatro pestañas para configurar la búsqueda. En la primera pestaña (*query for report*) se introducen los términos de búsqueda (hasta un máximo de 6) que se pueden combinar con el uso de los operadores AND y OR. También es posible utilizar corchetes y llaves para añadir variaciones lingüísticas en las búsquedas. Si las variaciones son de un único carácter; se introducen todos los caracteres seguidos entre corchetes. Por ejemplo, *s[iau]nk* recuperará *sink*, *sank* y *sunk*. Si las variantes son de 0 o más de un carácter, se separan con comas dentro de los corchetes. Por ejemplo, *s[iau]nk[,s,ing]* podría recuperar: *sink*, *sank*, *sunk*, *sinks*, *sanks*, *sunks*, *sinking*, *sanking*, *sunking*. Para incluir variaciones de palabras en una expresión, se utilizan las llaves, {I,we} *behave \*sel{f,ves}* recupera *I behave myself* y *we behave ourselves*. La herramienta ignora los signos de puntuación, los acentos, las mayúsculas y los espacios dobles. Por ejemplo, al introducir la palabra *jugó*, se recuperaría el sustantivo *jugó* en minúscula, pero también se incluiría la conjugación verbal *jugó* o variaciones en mayúsculas como *Jugó* o *Jugó*. Sin embargo, marcando la casilla *Sic*, las minúsculas, mayúsculas y acentos se respetan. También es posible restringir la búsqueda a páginas de un determinado idioma. Por ejemplo, en la Ilustración 75, mostramos la búsqueda de la palabra *Irak*, utilizando corchetes para introducir variantes: *Ira[kq]*, *irak[i,is,ies]*, *iraq[i,is,ies,uies]*. Restringimos la búsqueda a páginas en español.

---

<sup>73</sup> *KWiCFinder* está disponible para su descarga en la página <http://www.kwicfinder.com/KWiCFinder.html>



**Ilustración 75.** Introducción de términos de búsqueda con *KWiCFinder*.

La segunda pestaña de búsqueda (*require these criteria*), permite especificar criterios que deben cumplirse en las páginas recuperadas. Se pueden restringir las páginas de la búsqueda con el uso de los operadores filtro domain:, host:, link:, title: y url:. También es posible limitar las páginas por país. Por último, puede delimitarse la búsqueda a páginas actualizadas en un período de tiempo. En la tercera pestaña (*exclude these criteria*) se pueden utilizar las mismas opciones que en la segunda pestaña, pero para excluir páginas de búsqueda.

En la cuarta pestaña (*search and report options*) se modifican opciones de la búsqueda y de los resultados, como el máximo de documentos para recuperar (10 a 1000), máximo de documentos para recuperar por dominio (todos, o 1 a 100), máximo de concordancias por documento (de 1 a 200), máximo de contexto para mostrar por concordancia (de 50 a 500 caracteres), y otras opciones como el tamaño máximo de los documentos.

Los resultados obtenidos en la búsqueda pueden visualizarse en diferentes formatos, entre ellos, el formato KWIC. También es posible observar los contextos de cada una de las páginas recuperadas en formato de tabla (*table*), con tres columnas: las palabras de búsqueda destacadas en la columna central, el contexto izquierdo en la columna izquierda y el derecho en la columna derecha; por párrafos (*paragraph*) con la palabras de búsqueda destacadas; y por documento (*KWiC by Doc*), se visualiza el texto completo recuperado y se destacan las palabras de búsqueda. En el modo de visualización KWIC (*KWiC Sort & Analysis*) todas las concordancias aparecen seguidas, sin especificar de qué página se han

recuperado, y pueden ordenarse alfabéticamente por las palabras a derecha o izquierda de la palabra de búsqueda.

Además, es posible buscar nuevas palabras dentro de las concordancias disponibles y aquí se permite utilizar comodines, el asterisco \* (sustituye 0 o más caracteres), la interrogación ? (sustituye 1 carácter), el porcentaje % (sustituye 0 o 1 carácter). Asimismo, pueden generarse listas de colocaciones de la palabra o palabras de búsqueda, puede especificarse el número mínimo de repeticiones que deben tenerse en cuenta para considerar a un término colocación, así como las posiciones que deben ocupar las colocaciones a derecha e izquierda de la palabra de búsqueda (de 1 a 5, por ejemplo). También puede accederse a los contextos de las colocaciones generadas.

Asimismo, desde la página de *KWiCFinder* se tiene acceso a la dirección [Webascorpus.org](http://Webascorpus.org) (Fletcher 2007b). Esta página permite la consulta en diferentes recursos en línea para la búsqueda en la Web como corpus, la mayoría todavía en estado de desarrollo. Uno de estos recursos es la interfaz *Web Concordancer*<sup>74</sup>, que genera concordancias de páginas de Internet recuperadas con el buscador *Bing*. Tiene dos pestañas de búsqueda, una simple y otra avanzada.

Con la búsqueda simple, la herramienta busca un término o secuencia exactos, tiene en cuenta el uso de apóstrofes, pero no de guiones. No admite el uso de comodines. Es posible restringir la búsqueda por idioma. Recupera concordancias de páginas en distintos formatos, por ejemplo html y pdf. Debajo de cada página recuperada, el término o secuencia de búsqueda aparece destacada. Asimismo, las páginas recuperadas pueden descargarse fácilmente en un archivo zip, que contiene las páginas en html o txt. De esta manera, se puede obtener un pequeño corpus para continuar el análisis con un programa de gestión de corpus.

La búsqueda avanzada de *Web Concordancer* posibilita introducir más de una palabra de búsqueda, y especificar si queremos que se recuperen páginas con cualquiera (*any*) o con todas (*all*) las palabras introducidas. Asimismo permite incluir o excluir una o más palabras. También es posible restringir la búsqueda a páginas de una determinada lengua. Por ejemplo, si introducimos como palabra de búsqueda el término *pasta* e incluimos como palabras de búsqueda *construcción* y *edificios*, pero excluimos *comida*, *cocina*, *macarrones*, *pizza* y *espagueti*, el resultado son páginas sobre construcción en las

---

<sup>74</sup> Disponible en la página <http://webascorpus.org/searchwac.html>



que aparece la palabra *pasta* pero no aparecen páginas relacionadas con la cocina. En la Ilustración 76 aparece una muestra de las concordancias generadas.

The screenshot shows the WebAsCorpus.org Web Concordancer beta interface. The search query is "web concordances in Spanish" for the word "pasta". The results show a list of words: construcción, edificios, comida, cocina, macarrones. The interface includes tabs for Simple Query, Advanced Query, Details, and Options. Below the search results, there is a list of search options like Language (Spanish), Max. webpage (50), and Max. matches each (50).

**Ilustración 76. Búsqueda avanzada en Web Concordancer.**

La pestaña *Source options* permite modificar opciones de búsqueda y resultados. El programa puede mostrar un contexto a cada lado de la palabra pivote de 50 a 1000 caracteres. El número de páginas analizadas puede ser de 10 a 500. El número de concordancias por página recuperada puede ser de 10 a 100. Asimismo, es posible restringir los países a los que pertenecen las páginas recuperadas, se puede seleccionar uno o varios al mismo tiempo.

Finalmente, haremos una breve referencia a la aparición de una serie de herramientas que facilitan la recuperación de corpus paralelos en Internet, bien facilitando la producción de expresiones de búsqueda para interrogar a buscadores existentes, sobre todo *Google*, o bien incorporando su propia base de datos independiente de páginas web bilingües. Por ejemplo *WWW Search Interfaces for Translators* (alojada en [www.multilingual.ch](http://www.multilingual.ch)) (Harvey Ciampi 2004) asiste al usuario en la selección de palabras o expresiones para acceder en *Google* a glosarios, textos paralelos en diferentes idiomas y abreviaturas.

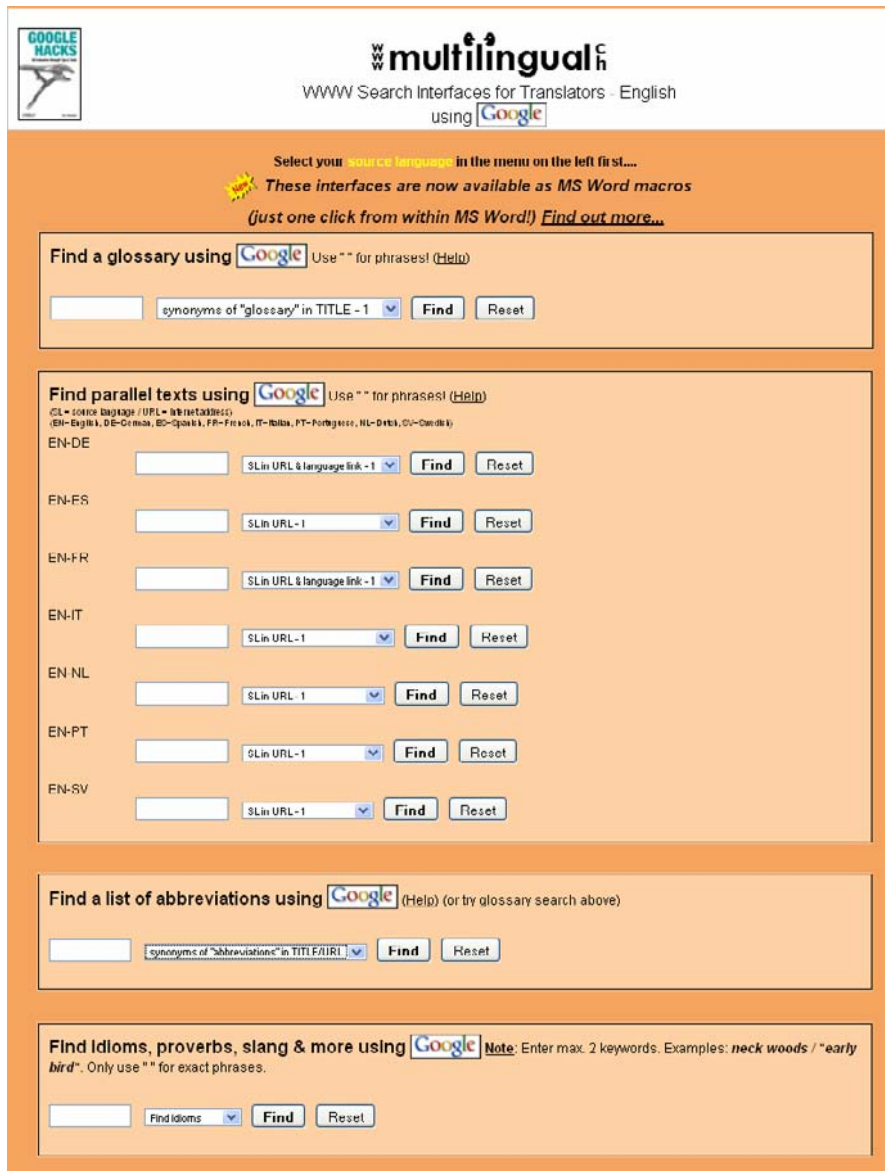


Ilustración 77. Interfaz de búsqueda de *WWW Search Interfaces for Translators*.

Una herramienta similar a la que acabamos de describir es *Term-minator* (Riediger 2009b) (<http://www.term-minator.eu/es.html>) que también facilita la introducción de consultas eficaces en buscadores (definiciones, traducciones en sitios multilingües, abreviaturas, etc.) y otros recursos de Internet como diccionarios, enciclopedias, libros, recursos institucionales como los de la Unión Europea, etc.



## Term-minator



**Ilustración 78.** Interfaz de búsqueda de *Term-minator*.

*Linguee* funciona a modo de buscador y recupera contextos de páginas web bilingües que se han recuperado de Internet por medio de un programa propio (araña web) que contiene algoritmos de aprendizaje automático para determinar la calidad de los textos. Asimismo, se incorporan funciones en la interfaz de resultados para que los usuarios puedan evaluar la calidad de los mismos y de esta manera el programa también aprende de las valoraciones de sus usuarios. Véase por ejemplo en la Ilustración 79 la búsqueda de *efecto invernadero* en los textos bilingües español-inglés recuperados por *Linguee*. Al usuario se le permite sugerir una traducción a partir de los contextos recuperados así como valorar positiva o negativamente cada uno de los contextos por medio de los iconos disponibles en la parte derecha de la interfaz.

*Keybot* (www.keybot.com) funciona de manera parecida al recurso anterior. Busca páginas web bilingües y las divide e indexa en función del formato de las páginas de manera que, a partir de la búsqueda de un término en una lengua, propone contextos supuestamente equivalentes en otra lengua. Las equivalencias propuestas pueden evaluarse accediendo al contexto completo de las páginas que es posible visualizar alineadas a modo de concordancias paralelas, como vemos en la Ilustración 80.

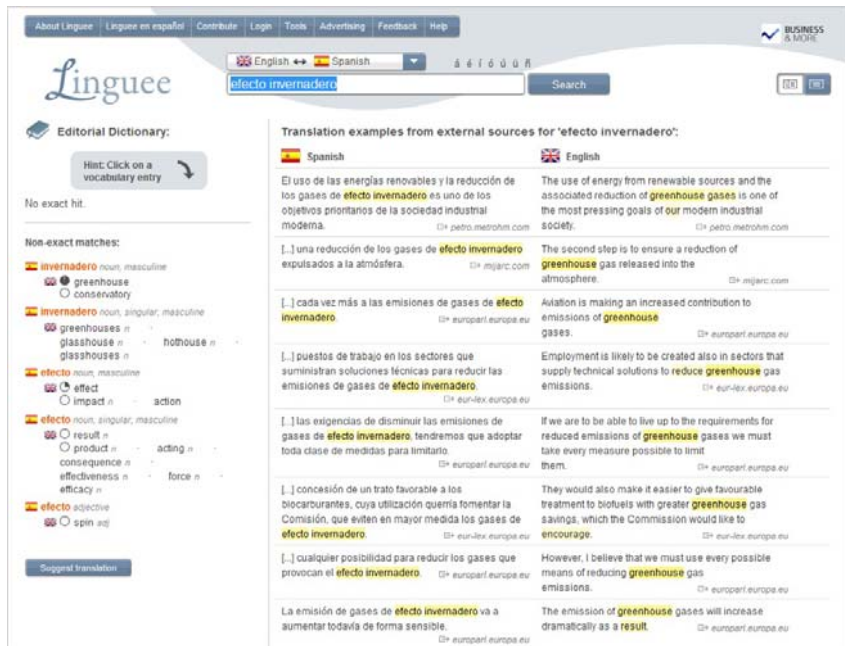


Ilustración 79. Búsqueda de *efecto invernadero* en la herramienta *Linguee*.



Ilustración 80. Búsqueda de *greenhouse effect* en *Keybot*.

**WeBiText** ([www.webitext.com](http://www.webitext.com)) permite la consulta de determinados sitios web a modo de corpus bilingüe. El usuario selecciona la lengua origen y meta para los contextos bilingües. También es necesario seleccionar un corpus de consulta (que puede estar compuesto por sitios web de distinto tipo, por ejemplo del Gobierno de Canadá), algunos de estos sitios se han procesado previamente, por lo que la consulta se agiliza. Otros sitios disponibles para la consulta no están procesados. La herramienta permite también al usuario añadir sus propios sitios para la búsqueda. Por ejemplo, la Ilustración 81 muestra la consulta de *divisa* para el par de lenguas español-inglés en el sitio web del Banco Central Europeo. Accedemos a contextos bilingües español-inglés donde el término *divisa* aparece en las páginas en español. Asimismo, es posible acceder al documento completo del cual se ha extraído cada contexto, tanto en formato texto como web, haciendo clic en los iconos centrales de los contextos bilingües recuperados.

The screenshot shows the WeBiText interface. At the top, there are logos for 'WEBITEXT' and 'TERMINOTIX', and navigation links for 'English', 'Français', 'Español', and 'Português'. A disclaimer states: 'Disclaimer: this prototype is provided to the public free of charge and for a limited time only.' Below this, there are sections for 'Select the languages:' (Spanish to English), 'And a predefined corpus from our list:' (European Central Bank), and 'Add your own sites to search (optional):'. A search bar contains the word 'divisa' and a 'Search' button. Below the search bar, it says 'Click on link to search in: Terminum' and 'Search completed. Results for "divisa": 6'. The results are presented in a table with two columns: 'Spanish' and 'English'. Each row contains a numbered entry with a small icon between the two columns. The first entry (1) discusses the history of the peseta and the introduction of the 'divisa'. The second entry (2) compares exchange rates of the LML and the peseta. The third entry (3) discusses the precarious state of the Treasury and the impact on monetary stability.

#	Spanish	English
1	<p>À tal fin, partiendo de los hechos tal como los han contado los historiadores, subrayaré algunos rasgos de la peseta que me han llamado la atención: las propicias circunstancias que acompañaron a su creación en 1868; su vinculación con otras monedas en su principio y en su final; el orden que desde el comienzo impuso en nuestro sistema monetario; su relativa estabilidad secular, a pesar de las "malas compañías", es decir, del déficit crónico de la Hacienda Pública; el hecho de que, pese a ello y también desde su origen hasta su destino final en el euro, la peseta ha sido bien administrada por gente solvente y, finalmente, trataré de hacerles ver que nuestra <b>divisa</b> ha tendido a ir contracorriente en la elección del patrón monetario en la primera mitad de su existencia, con cierta tendencia a la sobrevaloración en determinadas fases que el mercado inevitablemente acabó corrigiendo con drásticas depreciaciones.</p> <p><a href="http://www.ecb.int/press/key/date/2002/html/sp1212/">http://www.ecb.int/press/key/date/2002/html/sp1212/</a></p>	<p>Based on the facts as told by the historians, I would like to mention some of the peculiarities of the peseta which have caught my eye: the favourable circumstances surrounding its creation in 1868; its relationships with other currencies at its beginnings and end; the order it established in our monetary system from the onset; its relative stability over the centuries, despite falling into the "bad company" of chronic Treasury deficit; the fact that, notwithstanding these difficulties, right from its beginnings until its final destiny as part of the euro, the peseta has been well managed by reliable people. Finally, I will attempt to show you that our currency has tended to go against the tide in terms of the selection of the monetary standard during the first half of its existence, with a certain tendency towards overvaluation during specific periods, which the market inevitably corrected by drastic depreciation.</p> <p><a href="http://www.ecb.int/press/key/date/2002/html/sp1212/">http://www.ecb.int/press/key/date/2002/html/sp1212/</a></p>
2	<p>Si partimos de la equivalencia existente en la UML entre las monedas y la peseta y la comparamos con los cambios fijos e irrevocables de integración en el euro observamos que al cabo de 130 años una peseta se cambiaría por 3,94 francos viejos, 11,64 liras y 2,05 dracmas, lo que implica una relativa apreciación secular de nuestra <b>divisa</b>.</p> <p><a href="http://www.ecb.int/press/key/date/2002/html/sp0202...">http://www.ecb.int/press/key/date/2002/html/sp0202...</a></p>	<p>If we compare the old exchange rates of the currencies of the LML and the peseta with the fixed, irrevocable euro conversion rates of today, we see that after 130 years, one peseta would have an exchange rate value of 3.94 old francs, 11.64 lire and 2.05 drachmas, indicating a relative historical appreciation of our currency.</p> <p><a href="http://www.ecb.int/press/key/date/2002/html/sp0202...">http://www.ecb.int/press/key/date/2002/html/sp0202...</a></p>
3	<p>Pese a la precariedad de la Hacienda Pública y las negativas consecuencias que ello ha tenido para la estabilidad monetaria, yo sostengo que nuestra <b>divisa</b> ha sido, en general, bien administrada, que la intención de ortodoxia y de rigor técnico ha estado siempre presente y que con la peseta no se han cometido los blates que han sufrido otras monedas.</p> <p><a href="http://www.ecb.int/press/key/date/2002/html/sp0202...">http://www.ecb.int/press/key/date/2002/html/sp0202...</a></p>	<p>Despite the Treasury's precarious state and the resultant negative effects on monetary stability, generally speaking, I believe that our currency has been well managed, with a commitment at all times to orthodoxy and technical rigor, and that the peseta has not been played with in the way other currencies have.</p> <p><a href="http://www.ecb.int/press/key/date/2002/html/sp0202...">http://www.ecb.int/press/key/date/2002/html/sp0202...</a></p>

Ilustración 81. Búsqueda de *divisa* en el sitio web del Banco Central Europeo con *WeBiText*.

### 3.4.4 Clasificación de técnicas de búsqueda en Internet

En apartados anteriores hemos visto que las herramientas para analizar la Web son muchas y distintas, e implementan técnicas de búsqueda muy variadas difíciles de recopilar en un solo manual válido para cualquier búsqueda terminológica que se realice en Internet.

Asimismo, las clasificaciones de técnicas de búsqueda nos han permitido en este trabajo organizar las técnicas de búsqueda en diccionarios y corpus utilizando los mismos criterios de clasificación (consulta, instrumento y resultado).

Para el análisis de las técnicas de búsqueda en Internet, tomaremos como punto de partida la clasificación de técnicas de búsqueda que ya hemos creado para los diccionarios electrónicos y los corpus y analizaremos cómo podría aplicarse esta clasificación a la búsqueda de terminología en Internet.

Nuestra clasificación de técnicas de búsqueda está dividida en función de los tres elementos que hemos visto que intervienen en una búsqueda: la consulta, el instrumento y el resultado. La distinción entre consulta, instrumento y resultado nos permitirá reflejar de una forma más clara y ordenada todas las posibilidades de búsqueda de los recursos de Internet y compararlas con las de los otros recursos analizados.

#### 3.4.4.1 El primer elemento: la consulta

La consulta es la expresión que introduce el usuario en la interfaz de un recurso cuando realiza una búsqueda. En Internet, las consultas pueden ser: 1) una expresión léxica, 2) un número, 3) dos o más expresiones, y también es posible utilizar filtros para restringir las consultas. A diferencia de los corpus, en Internet no pueden consultarse expresiones gramaticales porque los textos no están etiquetados gramaticalmente. Asimismo, las expresiones léxicas no podrán ser lemas ya que los textos de las páginas web no se encuentran lematizados.

##### 3.4.4.1.1 Búsqueda mediante una expresión léxica

En Internet es posible la consulta de **formas exactas** y en algunos recursos también **truncadas**. Una forma exacta es una palabra completa. Las formas exactas pueden introducirse en cualquiera de las herramientas de búsqueda en Internet analizadas. Por ejemplo, podemos buscar en *Google* la forma exacta *diputados*.

Una **forma truncada** es una palabra incompleta. La parte de la palabra que se omite, se sustituye con un carácter comodín. Puede ser útil si se quieren buscar todas las palabras que contengan, empiecen o terminen con una secuencia de caracteres. Los corpus y los diccionarios electrónicos permiten el uso de caracteres comodín para truncar palabras en las consultas, mientras que su uso no está tan extendido en los recursos de Internet. Los buscadores, en general, no admiten el uso de comodines para truncar palabras. Los

comodines en los buscadores suelen sustituir a palabras completas dentro de la búsqueda de una secuencia de palabras entre comillas.

Algunas herramientas de análisis de la Web como corpus tampoco admiten el uso de formas truncadas porque sus motores de búsqueda son los buscadores tradicionales, algunas excepciones son *WebCONC* que permite el uso de expresiones regulares cuando el usuario especifica su propio corpus de búsqueda; *KWiCFinder* no admite el uso de comodines al introducir las palabras de búsqueda para recuperar el corpus de Internet pero, una vez que se han recuperado las páginas de Internet, es posible buscar nuevas palabras con el uso de comodines dentro de los contextos.

Finalmente, en algunos buscadores es posible la búsqueda mediante una **expresión aproximada**. Al igual que los diccionarios, las interfaces de búsqueda de algunos buscadores tienen una opción de sugerencias de consulta<sup>75</sup>, cuando la búsqueda no ha generado resultados. Por ejemplo, *Google* tiene una función en la que asiste al usuario proponiéndole expresiones a medida que este introduce la consulta en la casilla de búsqueda. Además, en el caso de que el usuario haya cometido un posible error al introducir la consulta, por ejemplo ortográfico, el buscador lo corrige y sugiere una consulta correcta. Así, si se consulta *palca base*, se muestran los resultados de *placa base*.

#### 3.4.4.1.2 Búsqueda mediante un número

En los corpus veíamos que los números pueden introducirse de forma exacta o truncada. En el caso de introducir un **número exacto**, este se busca en el corpus tal cual se ha introducido. Un **número truncado** es un número combinado con un comodín. En este caso, se busca en el corpus todos los números que contienen la secuencia de números que acompaña al comodín. La búsqueda de números puede ser útil para localizar palabras que suelen aparecer en un contexto próximo a algún número significativo. Por ejemplo, en los corpus poníamos el ejemplo de introducir en la interfaz de *Bwananet* el número 640 para buscarlo en el corpus de la informática en español, de esta forma localizábamos palabras como *píxel*, porque una medida que recupera el corpus es *640x480 píxels*; también se localiza *memoria RAM*, porque otra medida que se recupera es *640 Kb de memoria RAM*. Las expresiones que recupera el corpus pertenecen al ámbito de la informática.

---

<sup>75</sup> Véase por ejemplo Church (2008: 334)

En los buscadores de Internet pueden introducirse números en las búsquedas, pero los resultados no serán tan significativos como con los corpus, ya que la temática del conjunto de textos que componen la red es muy amplia y un mismo número puede utilizarse en textos de diferente temática en contextos diferentes. Si buscamos en *Google* el número 640, recuperamos expresiones tan variadas como *Real Decreto 640/2009*, *Teka TR-640*, *Motos KTM LC4 E 640*, etc. Por otro lado, *Google* tiene una opción de búsqueda que permite buscar intervalos de números, añadiendo dos puntos entre las dos cifras del intervalo, por ejemplo, la búsqueda de *125..512 Mb*, recuperará *128 Mb*, *256 Mb*, *512 Mb*, etc.

La técnica de introducir números, sin embargo, sí que podría resultar más útil con las herramientas de gestión de corpus que compilan y analizan páginas descargadas de Internet (como *TerminoWeb*, *Corpógrafo* o *WordSmith*), ya que en estos casos, aunque el corpus está formado por páginas de Internet, suele estar restringido previamente a un área temática más concreta. En *Corpógrafo*, si creamos un corpus con páginas web de informática (descargándolas a partir de alguna página web de informática proporcionada por el usuario) y buscamos el número 640, sí que accedemos a medidas que contienen términos propios de la informática, como *640 Kbytes de memoria RAM*, *memoria RAM hasta 640 MB*, etc. Pero en este caso, para que en los resultados aparezca el término *memoria RAM*, el usuario tiene que haber proporcionado a *Corpógrafo* alguna página que contenga información sobre la memoria RAM. En el caso de haber cargado en *Corpógrafo* una página del campo de la informática, pero de cualquier otro tema, es muy probable que el término *memoria RAM* no estuviera presente. Además, las páginas que descarga *Corpógrafo* de Internet contienen menos palabras que un corpus ya compilado, por tanto, las posibilidades de obtener resultados también se reducen.

#### 3.4.4.1.3 Búsqueda mediante dos o más expresiones

Las consultas en Internet también pueden ser **combinaciones de dos o más formas y números**. En el caso de las herramientas de búsqueda en Internet las expresiones mixtas estarán formadas por formas y números; pero no por lemas o categorías gramaticales, ya que las páginas de Internet no están etiquetadas gramaticalmente ni lematizadas.

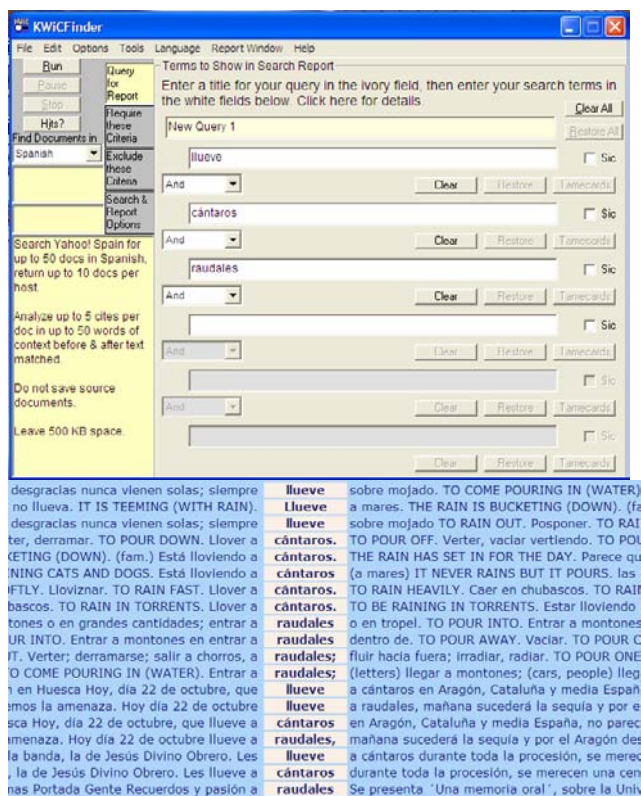
En Internet es posible introducir expresiones indicando que estén **presentes todos los elementos** (con el uso por ejemplo del operador AND), que estén **presentes algunos de los elementos** introducidos (con el uso del operador OR) y que estén **ausentes algunos**



**elementos** (con el uso del operador NOT). También puede especificarse en qué lugar de las páginas web deberán incluirse o excluirse dichos elementos.

La búsqueda avanzada de *Yahoo!* permite añadir o excluir palabras de la búsqueda, que deberán aparecer, o no, en los títulos o en cualquier parte de las páginas recuperadas. *WebCorp* ofrece la opción *word filter* para incluir o excluir palabras de la búsqueda. En *Web Concordancer*, se pueden incluir o excluir palabras de la búsqueda en las opciones de búsqueda avanzada. En *KWiCFinder* se pueden añadir o excluir palabras que deberán aparecer, o no, en el título o en la url de las páginas recuperadas.

En la Ilustración 82 hacemos una consulta en *KWiCFinder* en la que deben estar presentes todas las formas exactas *llueve*, *cántaros* y *raudales* en páginas de Internet en español. En el resultado vemos concordancias en las que aparecen dichas palabras de búsqueda. Accedemos a secuencias como *llueve sobre mojado*, *llover a cántaros*, *llueve a raudales*, etc.



**Ilustración 82. Búsqueda de combinaciones de formas exactas en KWiCFinder.**

Además, Internet permite la **combinación continua de expresiones**. Puede servir cuando se quiere acceder al contexto de una expresión fija. Es posible introducir combinaciones continuas en los buscadores por medio de su introducción entre comillas.

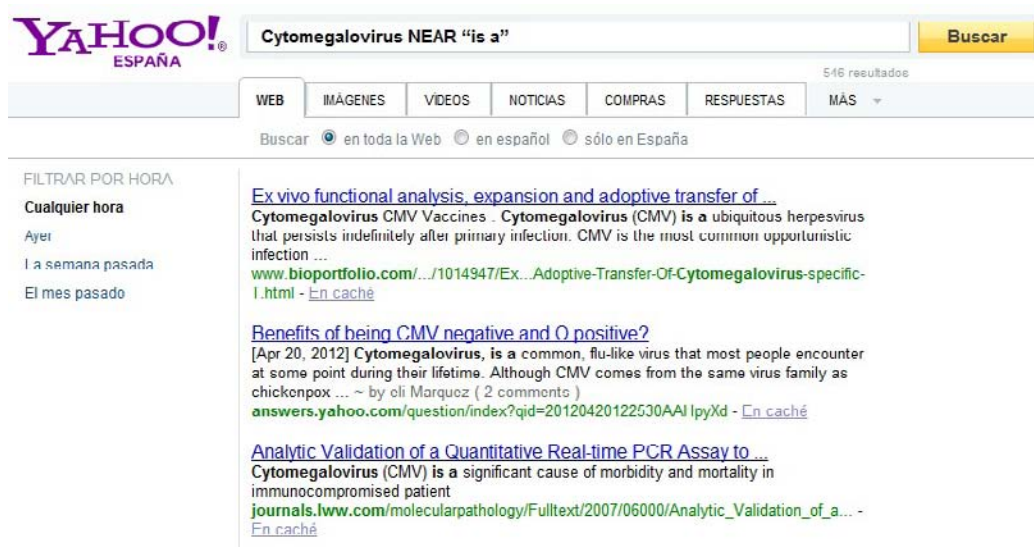


Por ejemplo, podemos buscar en los buscadores *Google* o *Yahoo!* la combinación “prensado en seco”.

Una **combinación discontinua con delimitación de proximidad** en Internet puede buscarse combinando formas y caracteres comodines. Además algunos buscadores permiten operadores como NEAR. Puede ser útil si se quiere buscar una secuencia y solo se conoce alguna de las palabras que la componen.

Por ejemplo, pueden introducirse combinaciones discontinuas en los buscadores utilizando comodines dentro de las comillas para sustituir palabras. Pueden utilizarse tantos comodines como palabras o posiciones quieran sustituirse. Respecto al resto de recursos de búsqueda en Internet analizados, *WebCorp* admite el uso de comodines para sustituir palabras en una combinación discontinua. Podemos introducir, por ejemplo, la combinación *no \* ojo* y se recuperan concordancias en las que *no* aparece en primera posición seguido de una palabra cualquiera y a continuación *ojo*: *no tiene ojo, no típico ojo, no pegas ojo, no pegar ojo, no quitar ojo, no pierde ojo*, etc. *WebCONC* admite el uso de comodines y expresiones regulares en las búsquedas para buscar combinaciones discontinuas, pero únicamente en la opción de búsqueda en la que el usuario especifica su propio corpus de búsqueda, no en la búsqueda directa en Internet. *KWiCFinder* no permite la introducción de combinaciones discontinuas para buscar páginas directamente en Internet, pero sí que permite introducir variaciones de combinaciones continuas con el uso de llaves. Además, en *KWiCFinder*, sí que es posible buscar combinaciones discontinuas con el uso de comodines, pero cuando las páginas ya están recuperadas.

Asimismo, algunos buscadores permiten el uso del operador NEAR. Por ejemplo, si introducimos en *Yahoo!* la expresión *Cytomegalovirus NEAR “is a”*, en los resultados encontramos contextos definitorios en los que el término *Cytomegalovirus* aparece en un posición cercana a *is a*, como por ejemplo: *Cytomegalovirus (CMV) is a ubiquitous herpesvirus that persists indefinitely after primary infection, Cytomegalovirus (CMV) is a significant cause of morbidity and mortality in immunocompromised patient*, etc.



**Ilustración 83.** Búsqueda de *Cytomegalovirus NEAR "is a"* en Yahoo.

En Internet también es posible la **introducción de combinaciones de expresiones en lenguaje natural**<sup>76</sup>. Aunque el usuario no introduzca operadores junto con las expresiones de búsqueda, la mayoría de buscadores (como por ejemplo *Google*) tendrán en cuenta todas las expresiones introducidas como palabras clave de búsqueda, pero excluirán aquellas palabras vacías de significado o *stopwords* (como preposiciones, artículos, etc.) que deberán introducirse dentro de una expresión entrecomillada o con un operador de inclusión delante (AND, +) para que el sistema las tenga en cuenta en la búsqueda.

#### 3.4.4.1.4 Búsqueda mediante el historial de consultas

Las herramientas de búsqueda en Internet, como la mayoría de recursos electrónicos que hemos analizado, almacenan las consultas introducidas por los usuarios, con lo que se encuentran disponibles en el caso de que sea necesario repetirlas. Asimismo, los navegadores de Internet cuentan con un historial de búsquedas que contiene todas las páginas a las que se ha accedido.

Algunos agentes inteligentes, como *Copernic Agent* almacenan las búsquedas y sus resultados por categorías y carpetas, de manera que el usuario puede acceder a ellas posteriormente.

<sup>76</sup> De hecho, según Baeza-Yates y Ribeiro-Neto (1999: 104) las búsquedas booleanas son una abstracción de las búsquedas en lenguaje natural. Véase Baeza-Yates y Ribeiro-Neto (1999: 287-288) para más información sobre las consultas en lenguaje natural: «[...] a natural language query entered by the user is treated as a bag of words, with stopwords removed, for the purposes of document match. However, some systems attempt to parse natural language queries in order to extract concepts to match against concepts in the text collection [...]».

#### 3.4.4.1.5 Filtros

Los filtros añaden una condición de búsqueda a la consulta introducida. En Internet existen, por ejemplo, filtros de **área geográfica, idioma, fecha, lugar del texto, formato de archivo**, etc.

Otros filtros que encontrábamos en los corpus, como el de área temática, no están desarrollados por completo en Internet. Normalmente, en los buscadores no es posible utilizar filtros de área temática para restringir la búsqueda. La única posibilidad es introducir palabras clave propias de un campo temático. Por ejemplo, en los buscadores podemos incluir palabras temáticas que deberán aparecer en el título de las páginas recuperadas, con el filtro `title:`, o en la dirección de la página, con el filtro `inurl:`. También podemos restringir la búsqueda a un determinado dominio de Internet que el usuario sepa que contiene textos de una determinada área temática, con el filtro `site:`. Algunos buscadores temáticos, como por ejemplo *Google Académico*, limitan la búsqueda a textos académicos, libros, artículos, etc.

Asimismo, la herramienta *WebCorp*, en una versión anterior, disponía de una opción de filtrado por área temática. El usuario seleccionaba las áreas temáticas de una lista desplegable. Estas áreas temáticas supuestamente coincidían con las categorías en las que estaban clasificadas las páginas en los sistemas de búsqueda que utiliza *WebCorp*. La selección de áreas temáticas en *WebCorp* no se encontraba suficientemente desarrollada lo que provocaba que, en ocasiones, la búsqueda no produjera resultados. Por ejemplo, si buscábamos en *WebCorp* la palabra *conexión* y restringíamos al área *Computers*, se obtenía un único resultado en el que *conexión* aparecía como parte de una dirección de correo electrónico. Actualmente se ha eliminado la función de restricción por área temática de la interfaz de búsqueda.

El filtro de **área geográfica** restringe la búsqueda a textos de una determinada zona. Este filtro existe en las herramientas de análisis de Internet, restringiendo a páginas de un determinado país, con la extensión propia de un país. El filtro de **idioma** en Internet permite limitar las búsquedas a páginas del idioma que indiquemos, por ejemplo, solo páginas en español.

El filtro de **fecha** restringe la búsqueda a páginas publicadas en un período o fecha, por ejemplo en *Google, Yahoo!, WebCorp* y *KWiCFinder*. El filtro de **lugar del texto** restringe la búsqueda a los títulos, las listas, las tablas, etc. Existen filtros en los buscadores que permiten buscar en el contenido de las páginas, en los títulos o en las url.

En las búsquedas en Internet pueden utilizarse, además, los filtros de **formato de archivo** ya que las páginas web pueden tener diferentes formatos. Las búsquedas avanzadas de los buscadores y las herramientas de compilación y análisis de textos de Internet permiten restringir la búsqueda a páginas de un formato determinado, por ejemplo PDF, PS, DOC, HTML, etc.

#### 3.4.4.2 El segundo elemento: el recurso o instrumento

En Internet el instrumento en el que se realiza la búsqueda son las propias páginas web o textos que se encuentran en línea. En un principio, los buscadores no permiten, a diferencia de los corpus, preseleccionar textos de una determinada temática o establecer si la búsqueda va a realizarse en páginas monolingües o bilingües. La búsqueda se efectúa en el conjunto de páginas de Internet.

Es posible restringir los resultados introduciendo en la consulta palabras clave de idiomas (*English, español*) o incluso por medio de filtros restringir la búsqueda a páginas de un idioma. Aunque todo ello no garantiza que las páginas recuperadas sean únicamente de ese idioma.

En esta línea, se observan esfuerzos por elaborar buscadores que preseleccionen los textos de búsqueda, como por ejemplo textos bilingües paralelos de Internet. Las herramientas *Linguee, Keybot* y *WeBiText* incorporan buscadores inteligentes que rastrean Internet en busca de textos bilingües paralelos que se almacenan y se alinean automáticamente y pueden consultarse a través de su interfaz de búsqueda.

#### 3.4.4.3 El tercer elemento: el resultado

El resultado es el elemento que se obtiene de Internet con una determinada búsqueda. El resultado puede ser: 1) una lista de páginas web y una lista de concordancias monolingües o bilingües, 2) una lista de palabras, 3) una lista de agrupaciones de palabras, 4) una lista de palabras relacionadas semánticamente, 5) una lista de colocaciones, 6) una representación gráfica de las relaciones de una palabra, 7) una imagen o lista de imágenes y 8) archivos multimedia. Como veremos, en el caso de las búsquedas en Internet, los resultados de las búsquedas terminológicas presentan diferencias respecto a los corpus.

##### 3.4.4.3.1 Lista de páginas web y lista de concordancias

En Internet, los buscadores no recuperan concordancias en formato KWIC, a diferencia de los corpus. Los resultados son listas de páginas web que incluyen un resumen

o fragmento que suele corresponderse a un párrafo de la página web en el que, normalmente, las palabras de la búsqueda aparecen destacadas, pero no centradas. Para poder acceder al contexto completo o realizar búsquedas en una página recuperada por un buscador, es necesario entrar una por una en las páginas ofrecidas en los resultados y, una vez accedemos a una página, las palabras de la búsqueda no aparecen destacadas. La opción caché de *Google* sí que destaca con colores las palabras de la búsqueda dentro de las páginas recuperadas.

En los programas de análisis de la Web como corpus que hemos analizado, *WebCorp*, *WebCONC*, *KWiCFinder* y *Web Concordancer*, las concordancias sí que aparecen en formato KWiC. *WebCorp* y *KWiCFinder*, además, ordenan las concordancias alfabéticamente en función de las palabras a izquierda y derecha de la palabra de búsqueda; *WebCorp* también ordena las concordancias por fecha. *Web Concordancer* permite, además, descargar las páginas de las que se han extraído las concordancias, en html o txt, lo que facilita la compilación de un pequeño corpus para su posterior análisis con herramientas de gestión de corpus más avanzadas.

A modo de ejemplo, realizamos una búsqueda de expresiones muy similares en un buscador normal de Internet y en un buscador que permite la consulta de Internet como corpus. En el buscador *Google* buscamos la combinación “llueve a \*” y recuperamos un listado de páginas web en las que aparecen expresiones como *nunca llueve a gusto de todos* o *llueve a cántaros*. Las palabras de búsqueda aparecen destacadas en el resumen de las páginas recuperadas, pero para acceder a los contextos completos, el usuario tiene que entrar en cada página recuperada. Sin embargo, una búsqueda similar de las palabras *llueve*, *cántaros*, y *raudales* en la herramienta de análisis de la web como corpus *KWiCFinder*, recupera los resultados a modo de concordancias KWiC, donde las palabras de búsqueda aparecen dentro de su contexto y destacadas en el centro, lo que facilita el análisis lingüístico de los resultados (véase la Ilustración 84).

[CHARLY MORLOCK: Nunca llueve a gusto de todos](#)  
 3 Mar 2010 ... Nunca llueve a gusto de todos, que gran verdad: si hace pocas semanas comentaba lo bien que viene el agua al campo y las reservas ...  
[www.charlymorlock.com/.../nunca-llueve-gusto-de-todos.html](#) - [En caché](#)

[Diario de Jerez - Nunca llueve a gusto de todos](#)  
 8 Mar 2010 ... Nunca llueve a gusto de todos. ... Diario de Jerez, Jerez, Jerez, Nunca llueve a gusto de todos. Nunca llueve a gusto de todos ...  
[www.dianodejerez.es/.../llueve/.../todos.html](#) - hace 17 horas - [En caché](#)

[La Nueva España - Diario Independiente de Asturias - Occidente ...](#)  
 9 Mar 2010 ... ¿Y si llueve a cántaros? CELSA DÍAZ ALONSO Otra vez. Otro nuevo retraso anunciado en la lenta autovía de nuestros dolores. ...  
[www.lne.es/occidente/2010/03/09/llueve.../883677.html](#) - hace 8 horas

[llueve a cántaros / llover a cántaros - WordReference Forums - \[ Traducir esta página \]](#)  
 12 entradas - 9 autores - Última entrada: 11 Dic 2009  
[llueve a cántaros / llover a cántaros Spanish-English Vocabulary / Vocabulario Español-Inglés](#)  
[forum.wordreference.com/showthread.php?t...](#) - [En caché](#) - [Similares](#)

desgracias nunca vienen solas; siempre no llueve. IT IS TEEMING (WITH RAIN).	<b>llueve</b>	sobre mojado. TO COME POURING IN (WATER).
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	a mares. THE RAIN IS BUCKETING (DOWN). (fe
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	sobre mojado TO RAIN OUT. Posponer. TO RAI
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	TO POUR OFF. Verter, vaciar vertiendo. TO POL
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	THE RAIN HAS SET IN FOR THE DAY. Parece qu
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	TO RAIN HEAVILY. Caer en chubascos. TO RAI
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	TO BE RAINING IN TORRENTS. Estar lloviendo .
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	o en tropel. TO POUR INTO. Entrar a montones
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	dentro de. TO POUR AWAY. Vaciar. TO POUR G
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	fluir hacia fuera; Irradlar, radiar. TO POUR ONE
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	(letters) llegar a montones; (cars, people) lleg
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	a cántaros en Aragón, Cataluña y media Españ
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	a raudales, mañana sucederá la sequía y por e
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	en Aragón, Cataluña y media España, no parec
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	mañana sucederá la sequía y por el Aragón des
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	a cántaros durante toda la procesión, se merec
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	durante toda la procesión, se merecen una cen
desgracias nunca vienen solas; siempre ter, derramar. TO POUR DOWN. Llover a cántaros. (fam.) Está lloviendo a cántaros. (a mares) IT NEVER RAINS BUT IT POURS. las- raudales dentro de. TO POUR AWAY. Vaciar. TO POUR ONE	<b>llueve</b>	Se presenta "Una memoria oral", sobre la Univ

**Ilustración 84. Comparación de resultados en Google y KWicFinder.**

Por otro lado, en las búsquedas en Internet, por regla general, no se puede asegurar que una lista de concordancias sea monolingüe, o pertenezca a un solo idioma o variedad idiomática, ya que los filtros de restricción de idioma tienen sus limitaciones y puede que se recuperen, entre los resultados de una búsqueda, páginas en diferentes idiomas o en diferentes variantes de un mismo idioma, por ejemplo español de España y de México.

Las concordancias bilingües también presentan problemas en Internet, especialmente en lo que respecta a la detección por parte de los buscadores de textos bilingües. Por este motivo, en los últimos años hemos visto que están apareciendo interfaces que asisten a los usuarios en la generación de consultas que ayudan a la recuperación de páginas bilingües en los buscadores de Internet (ejemplo de ello son *WWW Search Interfaces for Translators* y *Term-minator*). Otras herramientas están desarrollando sus propios motores de búsqueda que incorporan sistemas de búsqueda inteligentes para la detección de páginas bilingües de Internet que se indizan y se alinean automáticamente (por ejemplo como hemos visto en *Keybot* a partir de la estructura de las páginas web) con el fin de proporcionar al usuario concordancias bilingües en páginas de Internet, que además el usuario puede validar o descartar, de manera que el conocimiento de los usuarios enriquece la herramienta de cara a futuras consultas por otros usuarios y también fomenta que el algoritmo de aprendizaje

automático identifique criterios que determinan la calidad de los textos conforme a las reacciones de los usuarios (como por ejemplo en *Linguee*).

### 3.4.4.3.2 Lista de palabras

Las listas de palabras pueden ser útiles para tener una visión rápida de la terminología propia de un ámbito. La mayoría de herramientas de búsqueda en Internet no generan listas de palabras. Los buscadores no producen listas de palabras. De las herramientas de análisis de la Web como corpus analizadas, la única que genera listas de palabras de las páginas recuperadas es *WebCorp*. Las palabras de las listas de *WebCorp* pueden ordenarse alfabéticamente o por frecuencia. Los programas de gestión de corpus que descargan páginas de Internet sí que permiten la generación de listas de palabras, por ejemplo *WordSmith Tools 4*, que también puede generar listas de palabras clave si se cuenta con un corpus de referencia. Asimismo, *TerminoWeb* tiene una función de extracción terminológica, que genera listas de palabras frecuentes del corpus.

En la Ilustración 85 vemos una lista de palabras frecuentes de páginas sobre el campo cerámica recuperadas por *WordSmith Tools 4* de páginas descargadas de Internet. Entre los resultados vemos que las palabras más frecuentes son del lenguaje general. Pero dentro de la lista encontramos más palabras propias de la cerámica, como *cerámica*, *ceramic*, *tile*, *brick*, *porcelain*, *mosaico*, *horno*, *terracota*, *arcilla*, *decoración*, *porcelana*, *pottery*, *engobe*, etc.

N	Word	Freq	%	Texts	% emmas	Set
1	#	465	6.11	22	75.86	
2	THE	138	1.81	14	48.28	
3	AND	127	1.67	12	41.38	
4	A	122	1.60	20	68.97	
5	DE	115	1.51	10	34.48	
6	OF	104	1.37	12	41.38	
7	CERAMICA	97	1.27	22	75.86	
8	IN	75	0.99	15	51.72	
9	TO	72	0.95	13	44.83	
10	FOR	52	0.68	11	37.93	
11	DOCUMENT	50	0.66	2	6.90	
12	CERAMIC	46	0.60	10	34.48	
13	SRC	43	0.57	10	34.48	
14	Y	40	0.53	7	24.14	
15	LA	39	0.51	9	31.03	
16	IMG	34	0.45	6	20.69	
17	GIF	33	0.43	3	10.34	
18	S	32	0.42	8	27.59	
19	OUR	30	0.39	9	31.03	
20	EN	29	0.38	8	27.59	
21	ON	29	0.38	5	17.24	
22	TILES	29	0.38	6	20.69	
23	COM	26	0.34	12	41.38	
24	MM	25	0.33	5	17.24	
25	QUALITY	25	0.33	8	27.59	
26	MENU	24	0.32	1	3.45	

Ilustración 85. Lista de palabras de un corpus sobre cerámica recuperado desde Internet con *WordSmith Tools 4*.

### 3.4.4.3.3 Lista de agrupaciones de palabras

Las agrupaciones de palabras son secuencias de dos o más palabras que aparecen con frecuencia en un corpus. Las agrupaciones de palabras pueden ayudar a tener una visión rápida de cómo se combina la terminología propia de un ámbito.

De los buscadores de Internet analizados, solo *WebCorp* genera listas de agrupaciones de hasta 5 palabras. Los programas de gestión de corpus que descargan textos de Internet, como *Corpógrafo* y *WordSmith Tools 4* generan listas de agrupaciones de palabras, también llamadas n-gramas o clusters.

Por ejemplo, en *WebCorp* generamos una lista de agrupaciones de dos palabras, 2-gramas, a partir de una página que contiene un texto sobre la memoria RAM: <http://www.islabit.com/177/memoria-ram-arquitectura-y-funcionamiento.html>. La lista contiene las agrupaciones de palabras: *de memoria, la memoria, la placa, placa base, memoria RAM*, etc.

### 3.4.4.3.4 Lista de palabras relacionadas semánticamente

Hemos visto que algunos corpus han incorporado la función de búsquedas semánticas. Esta función consiste en que se puede acceder a los sinónimos de la palabra introducida y realizar distintos tipos de búsqueda en relación con dichos sinónimos. Esta técnica de búsqueda puede ser útil para acceder a una palabra a partir de otras palabras que están relacionadas semánticamente.

Las herramientas de búsqueda en Internet no permiten, de momento, la búsqueda de sinónimos, aunque se están produciendo avances en este sentido. *Google*, por ejemplo, tiene en cuenta en sus búsquedas los sinónimos de las palabras introducidas cuando aparecen precedidas de una vírgula (~). Aunque esta técnica de búsqueda funciona en cierto sentido con el inglés, pero no funciona con otras lenguas.

Asimismo, tanto en *Google* como en *Yahoo!* se ha incorporado la función de *Búsquedas relacionadas*, lo que significa que el buscador sugiere al usuario palabras de búsqueda relacionadas con las que ha introducido.

*TerminoWeb* contiene una función de búsqueda de patrones semánticos, con la que pueden estudiarse las relaciones semánticas que rodean a una palabra o palabras del corpus. El sistema de búsqueda se basa en estos patrones para seleccionar las páginas web que recupera de Internet.



### 3.4.4.3.5 Lista de colocaciones

Una colocación es una palabra que aparece frecuentemente cercana a otra palabra. La utilidad de las colocaciones reside en que se puede acceder a palabras que aparecen en un contexto inmediato a otra palabra sin tener que leer todas las concordancias de un corpus en las que aparece dicha palabra. En definitiva, puede ayudar a agilizar el proceso de búsqueda.

De las herramientas de búsqueda en Internet que hemos analizado, las únicas que generan listas de colocaciones son *WebCorp*, *KWiCFinder*. En la Ilustración 86 hemos generado colocaciones de *hormigón* con la función *exclude stopwords* activada. Un análisis de las colocaciones permite observar que en primera posición a la derecha de *hormigón* son frecuentes los adjetivos *armado* (38 casos), *impreso* (10), *proyectado* (11).

**Post Search Options**

Show URLs:  Span: 50 characters i

Show Collocates:  Exclude Stopwords:  i

Sort by: Default v Position: 1 v i

Make all above options case insensitive:  i

Filter by Date: No filter v Range: YYYYMMDD - YYYYMMDD i

[Start a new search...](#)

#### Collocates

Word	L4	L3	L2	L1	R1	R2	R3	R4	Total
Hormigón	8	14	8	3	3	8	8	11	64
hormigón	15	4	6	0	0	6	9	12	52
armado	1	2	0	2	38	1	1	0	45
se	3	7	0	0	12	10	5	4	41
que	8	2	3	1	5	2	3	6	31
una	0	14	1	0	1	7	3	4	30
solera	0	0	28	0	0	0	0	0	28
concreto	0	0	0	15	7	1	2	1	26
hormigonera	14	0	0	0	0	3	1	4	22
camion	0	7	0	0	0	0	11	1	19
Planta	4	0	6	0	2	3	3	0	18
mezcladora	0	4	4	0	4	2	2	2	18
Bomba	0	1	14	0	1	1	0	0	17
cemento	1	0	2	2	9	2	0	0	16
enfriamiento	0	0	14	0	0	0	2	0	16
estructuras	0	0	12	0	0	0	1	0	13
Pavimentos	0	6	6	0	0	0	1	0	13
Impreso	1	0	1	0	10	0	0	1	13
celular	0	1	0	1	7	0	2	1	12
pretensado	1	1	0	1	6	0	3	0	12
montada	0	0	0	0	12	0	0	0	12
mezclador	0	0	11	0	0	0	0	1	12
mezcla	1	2	6	0	0	0	1	2	12
sistema	7	0	0	0	1	0	1	2	11
Sistema	6	0	0	0	3	1	1	0	11
más	0	2	6	0	1	0	0	2	11
proyectado	0	0	0	0	11	0	0	0	11
bloques	0	10	0	0	0	0	0	0	10
Cemento	0	1	0	6	1	1	0	1	10
España	0	2	0	1	1	2	2	2	10
construcción	4	1	1	2	0	0	1	0	9
prefabricado	2	1	0	0	5	1	0	0	9

[...]

Ilustración 86. Resultado de colocaciones con *WebCorp*.

Los programas de gestión de corpus que descargan textos de Internet también permiten la generación de colocaciones, ejemplo de ello son *TerminoWeb* y *WordSmith Tools 4*.

#### 3.4.4.3.6 Representación gráfica de las relaciones de una palabra

En el análisis de los recursos de Internet, hemos observado que la herramienta de análisis de corpus *Jaguar* genera grafos para representar la densidad terminológica de los corpus que recupera de Internet. Este grafo permite al usuario detectar con un golpe de vista la terminología del corpus y las relaciones entre los distintos términos. Pinchando en cada uno de los nodos del grafo (correspondiente a una unidad terminológica) el usuario accede a los contextos de dicho término en el corpus y puede decidir añadir este término o descartarlo de la lista de terminología del ámbito del corpus. Con los términos añadidos a la lista, el usuario puede realizar una nueva búsqueda en Internet para restringir el corpus de especialidad. Este proceso puede repetirse tantas veces como el usuario considere necesario. Véase el apartado 3.3.4.3.7 donde aparece un ejemplo de grafo.

#### 3.4.4.3.7 Lista de imágenes

Algunas herramientas de búsqueda en Internet también pueden obtener como resultado de la búsqueda una imagen o una lista de imágenes. Este es un resultado que no encontramos en nuestro análisis de técnicas de búsqueda en corpus, pero sí en los diccionarios electrónicos. Por ejemplo, en *Google*, y en la mayoría de buscadores de Internet, existe una función de búsqueda de imágenes, con la que, a partir de una palabra introducida por el usuario, el buscador recupera imágenes que aparecen en páginas en las que se encuentra dicha palabra y, que, por lo tanto, dichas imágenes podrían ser la representación de la palabra introducida. Las imágenes pueden resultar útiles para averiguar a qué hace referencia una palabra, en el caso de que no sea posible encontrar su definición en el diccionario o un sinónimo. Por ejemplo, *Google* recupera una lista de imágenes a partir de la introducción de la palabra *cigüeñal* (véase la Ilustración 87).

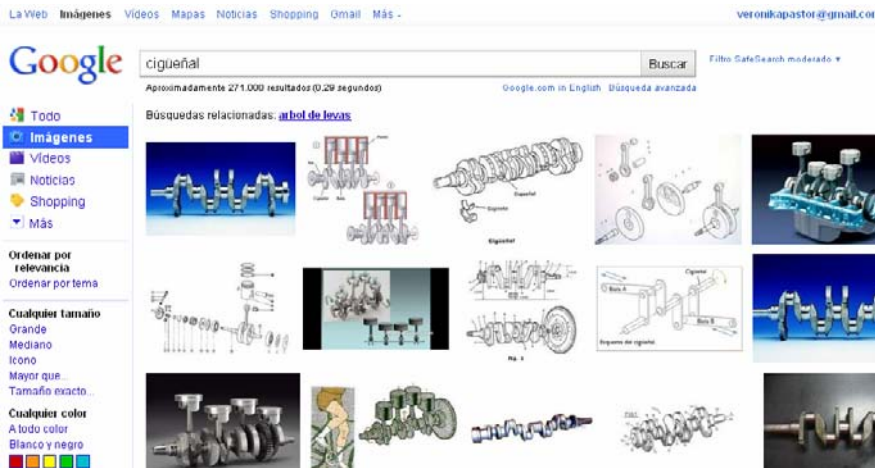


Ilustración 87. Resultado de imágenes de *cigüeñal* en *Google*.

### 3.4.4.3.8 Archivos multimedia

Un archivo multimedia combina diferentes medios (texto, imágenes, animación, sonido, video, etc.). Además de imágenes, la Web es una fuente de archivos multimedia que contienen información sobre los temas más diversos. Esta es otra diferencia que presenta Internet respecto a los corpus, que solo contienen información de texto, y los diccionarios, que pueden contener imágenes y archivos de sonido, pero no es común que incluyan archivos multimedia.

Por ejemplo, siguiendo el ejemplo anterior, con la búsqueda en *Google* de *cigüeñal* no solo podemos acceder a imágenes que nos ayudan a la comprensión del concepto, sino que podemos encontrar videos que nos explican su composición, funcionamiento, etc. (Ilustración 88).



Ilustración 88. Resultado de videos de *cigüeñal* en *Google*.

### 3.4.4.3.9 Información estadística

En los resultados de búsqueda en Internet también es posible obtener información estadística. Por ejemplo, los buscadores ofrecen en la lista de resultados el número de páginas web que se han recuperado y que contienen la expresión de consulta introducida.

Asimismo, otras herramientas de búsqueda en Internet permiten comprobar la tendencia de aparición de determinadas expresiones en páginas aplicando un filtro, como por ejemplo páginas de un dominio regional o de una fecha en concreto. De las herramientas que hemos analizado, esta información la proporcionan *Google trends* y *Diatopix*.

Por ejemplo, en *Google trends* si introducimos *ordenador* y *computadora* para comparar su distribución por regiones y fechas, la información estadística de los resultados nos permite comprobar que *ordenador* se emplea con mayor frecuencia en páginas de España y *computadora* en páginas de Méjico. También se observa la tendencia de uso por años (véase la Ilustración 89).

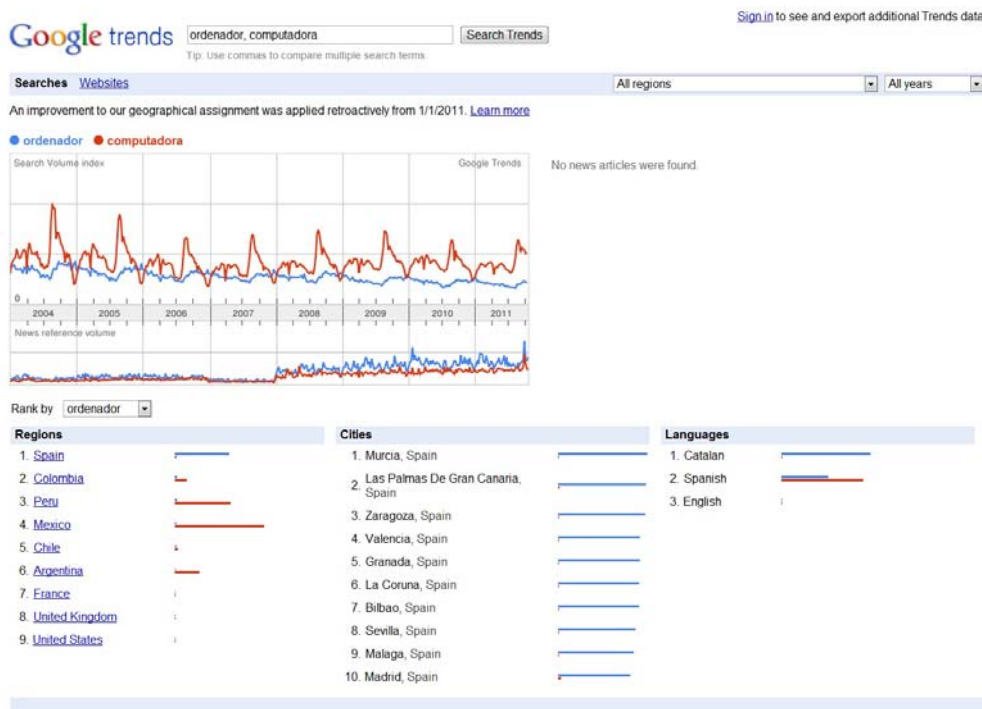


Ilustración 89. Resultados estadísticos sobre *ordenador* y *computadora* en *Google trends*.

### **3.4.5 Recapitulación de nuestra propuesta de clasificación de técnicas de búsqueda en Internet**


En los apartados anteriores hemos tratado las técnicas de búsqueda de terminología en Internet. Hemos visto que Internet puede consultarse como si se tratara de un gran corpus, sobre todo a partir de la aparición de nuevas herramientas, paralelas a los tradicionales buscadores de Internet, que han aportado nuevas formas de búsqueda que permiten consultar Internet de un forma más similar a cómo se haría en un corpus y, también, a partir del diseño de módulos en las herramientas que permiten descargar textos directamente desde Internet para analizarlos con programas de gestión de corpus.

La aparición de estas nuevas herramientas demuestra un interés por perfeccionar cada vez más las herramientas de búsqueda para potenciar las ventajas que ofrece Internet y disminuir las desventajas, especialmente en lo que se refiere a la realización de búsquedas terminológicas.

No obstante, pese a los avances que se han producido, y la creciente utilización de Internet como corpus, hemos visto que las búsquedas directas de terminología en la Web no han alcanzado el nivel de las búsquedas que facilitan en la actualidad los programas de gestión de corpus.

Hemos presentado una clasificación de técnicas de búsqueda en Internet tomando como punto de partida un sistema de clasificación que ya habíamos utilizado para ordenar las técnicas de búsqueda en diccionarios electrónicos y corpus. A partir de este análisis hemos generado una nueva clasificación de técnicas de búsqueda en Internet, que resumimos en la Tabla 23.

CONSULTA	RECURSO O INSTRUMENTO	RESULTADO
<ul style="list-style-type: none"> <li>•Expresión léxica (forma exacta o truncada y expresión aproximada)</li> <li>•Número</li> <li>•Dos o más expresiones (formas, números o expresiones mixtas, solo de formas y números)</li> <li>•Presencia de todos los elementos</li> <li>•Presencia de algunos elementos</li> <li>•Presencia de algunos elementos y ausencia de otros</li> <li>•Combinación continua o discontinua de expresiones</li> <li>•Consulta en lenguaje natural</li> <li>•Historial de consultas</li> <li>•Filtros               <ul style="list-style-type: none"> <li>-Área temática</li> <li>-Idioma</li> <li>-Área geográfica</li> <li>-Lugar del texto</li> <li>-Fecha</li> <li>-Formato de archivo</li> <li>-Dominio</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>•Textos/páginas de Internet</li> </ul>	<ul style="list-style-type: none"> <li>•Lista de páginas web y lista de concordancias (monolingües o bilingües)</li> <li>•Lista de palabras</li> <li>•Lista de agrupaciones de palabras</li> <li>•Lista de colocaciones</li> <li>•Lista de palabras relacionadas semánticamente</li> <li>•Representación gráfica de las relaciones de una palabra</li> <li>•Lista de imágenes</li> <li>•Archivos multimedia</li> <li>•Información estadística</li> </ul>



TÉCNICAS DE BÚSQUEDA

Tabla 23. Resumen de la clasificación de técnicas de búsqueda en Internet.

### 3.5 Propuesta conjunta de clasificación de técnicas de búsqueda en diccionarios, corpus e Internet

En este capítulo hemos examinado un conjunto de recursos (concretamente diccionarios, corpus e Internet) desde una perspectiva técnica para averiguar hasta qué punto han desarrollado las posibilidades de búsqueda. A partir de este análisis, hemos elaborado clasificaciones de técnicas de búsqueda en recursos de distinto tipo siguiendo los mismos criterios de categorización.

Nuestra clasificación de técnicas de búsqueda basada en la consulta, el instrumento y el resultado cubre las principales búsquedas que ofrecen los diccionarios, los corpus e Internet y las sistematiza. Es una clasificación que permite ser aplicada a la valoración de

cualquiera de los recursos que hemos examinado. Además es flexible, pues permite la incorporación de nuevos elementos.

Con el objetivo de poder utilizar esta clasificación en análisis posteriores, hemos asignado códigos formados por letras y números que identifican las distintas técnicas de búsqueda. De esta manera, la consulta se representa con la letra A, el recurso o instrumento con la letra B y el resultado con la C. En la Tabla 24 presentamos todas las técnicas de búsqueda y sus códigos.

Dentro de la consulta (A) distinguimos cinco criterios. El criterio A1 es el tipo de expresión, que puede ser A.1.1 expresión léxica (lema A.1.1.1, forma exacta A.1.1.2, forma truncada A.1.1.3, forma aproximada A.1.1.4, combinación de caracteres sin importar el orden o anagrama A.1.1.5 y expresión semántica o conceptual A.1.1.6), A.1.2 expresión gramatical y A.1.3 número. El segundo criterio A2 es el número de elementos de la consulta: A.2.1 un elemento, A.2.2 dos o más elementos.

En el caso de que la consulta esté formada por dos o más elementos, se añade un tercer criterio A3: la presencia total A.3.1 (presencia de todos los elementos introducidos), presencia parcial A.3.2 (presencia de algunos de los elementos introducidos), ausencia A.3.3 (ausencia de los elementos introducidos) y consulta en lenguaje natural A.3.4. Un cuarto criterio A4 es la forma de combinación: A.4.1 presencia de la combinación continua de expresiones y A.4.2 presencia de la combinación discontinua de expresiones con delimitación de proximidad. Además, existe un quinto criterio A5 que son los filtros de la consulta.

Dentro del instrumento (B) encontramos B.1 el campo de entrada (por ejemplo en un diccionario), B.2 el campo de contenido o texto (B.2.1 campo de definiciones, campo de relaciones semánticas B.2.2.1, o sintagmáticas B.2.2.2, foro interno B.2.3 y corpus interno B.2.4), B.3 campo de índice temático, B.4 campo de acceso a recursos externos, B.5 corpus monolingüe, B.6 corpus bilingüe (B.6.1 comparable y B.6.2 paralelo), B.7 corpus etiquetado y B.8 textos o páginas de Internet.

Dentro del resultado (C) diferenciamos: C.1 ficha con información de una palabra, C.2 representación gráfica de las relaciones de una palabra o grafo, C.3 denominación o lista de denominaciones, C.4 imagen o lista de imágenes, C.5 archivos de audio o multimedia, C.6 lista de concordancias (C.6.1 monolingües y C.6.2 bilingües), C.7 lista de palabras que aparecen en un conjunto de textos, C.8 lista de agrupaciones de palabras, C.9 lista de colocaciones, C.10 lista de palabras relacionadas semánticamente, C.11 lista de

palabras relacionadas sintácticamente, C.12 lista de categorías gramaticales, C.13 lista ordenada de páginas web con resumen y C.14 información estadística.



RECURSO ANALIZADO	CONSULTA (A)	INSTRUMENTO (B)	RESULTADO (C)
<p><b>Diccionarios electrónicos</b></p> <p><b>Corpus</b> (corpus en línea que incorporan sus propias interfaces, programas de gestión de corpus, programas TAO)</p> <p><b>Internet</b> (Buscadores, Web as Corpus)</p>	<p>Criterio A1. Tipo de expresión</p> <p>A.1.1 Expresión léxica</p> <p>A.1.1.1 Lema</p> <p>A.1.1.2 Forma exacta</p> <p>A.1.1.3 Forma truncada</p> <p>A.1.1.4 Forma aproximada (de parecido ortográfico o fonético)</p> <p>A.1.1.5 Combinación de caracteres sin importar el orden (anagrama)</p> <p>A.1.1.6 Expresión semántica/conceptual</p> <p>A.1.2 Expresión gramatical</p> <p>A.1.3 Número</p> <p>A.1.3.1 Número exacto</p> <p>A.1.3.2 Número truncado</p> <p>Criterios A.2 Número de elementos</p> <p>A.2.1 Una/uno</p> <p>A.2.2 Dos o más</p> <p>Criterio A.3 Presencia total/parcial/ausencia</p> <p>A.3.1 Presencia de todas las expresiones introducidas</p> <p>A.3.2 Presencia de alguna de las expresiones introducidas</p> <p>A.3.3 Ausencia de alguna de las expresiones introducidas</p> <p>A.3.4 Consulta en lenguaje natural</p> <p>Criterio A.4 Forma de combinación</p> <p>A.4.1 Presencia de la combinación continua de expresiones</p> <p>A.4.2 Presencia de la combinación discontinua de expresiones con</p>	<p>B.1 Campo de entrada</p> <p>B.2 Campo de contenido o texto</p> <p>B.2.1 Campo de definiciones</p> <p>B.2.2 Campo de relaciones léxicas</p> <p>B.2.2.1 Relaciones semánticas</p> <p>B.2.2.2 Relaciones sintagmáticas (colocaciones)</p> <p>B.2.3 Foro interno</p> <p>B.2.4 Corpus interno</p> <p>B.3 Campo de índice temático</p> <p>B.4 Campo de acceso a recursos externos</p> <p>B.4.1 Diccionario externo</p> <p>B.4.2 Buscador de Internet</p> <p>B.5 Corpus monolingüe</p> <p>B.6 Corpus bilingüe</p> <p>B.6.1 comparable</p> <p>B.6.2 paralelo</p> <p>B.7 Corpus etiquetado</p> <p>B.8 Textos/páginas de Internet</p>	<p>C.1 Ficha con información de una palabra</p> <p>C.1.1 Significado</p> <p>C.1.2 Información gramatical</p> <p>C.1.3 Información etimológica</p> <p>C.1.4 Información de uso en contexto / contextos</p> <p>C.1.5 Información de equivalencias</p> <p>C.1.6 Información de colocaciones y palabras relacionadas</p> <p>C.2 Representación gráfica de las relaciones de una palabra / grafo</p> <p>C.3 Denominación o lista de denominaciones que tienen una entrada en el diccionario</p> <p>C.4 Imagen o lista de imágenes</p> <p>C.5 Archivos de audio o multimedia</p> <p>C.6 Lista de concordancias</p> <p>C.6.1 Lista de concordancias monolingües</p> <p>C.6.2 Lista de concordancias bilingües</p> <p>C.7 Lista de palabras que aparecen en el conjunto de textos</p> <p>C.7.1 Lista de palabras frecuentes</p> <p>C.7.2 Lista de palabras clave</p> <p>C.8 Lista de agrupaciones de palabras</p> <p>C.9 Lista de colocaciones</p> <p>C. 10 Lista de palabras relacionadas semánticamente</p>

	delimitación de proximidad Criterio A.5 Filtros A.5.1 Categoría gramatical A.5.2 Área temática A.5.3 Idioma A.5.4 Tipo de texto A.5.5 Área geográfica A.5.6 Autor A.5.7 Fecha A.5.8 Lugar del texto A.5.9 Formato de archivo A.5.10 Dominio/URL		C.11 Lista de palabras relacionadas sintácticamente C.12 Lista de categorías gramaticales C.13 Lista ordenada de páginas web con resumen C.14 Información estadística
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

**Tabla 24. Resumen y códigos de nuestra clasificación conjunta de técnicas de búsqueda (en diccionarios, corpus e Internet).**

Un buen uso de las técnicas de búsqueda en diccionarios, corpus e Internet puede ayudar a los traductores en su trabajo. En este contexto, es necesario continuar indagando qué técnicas de búsqueda incorporan los distintos recursos y reflexionar sobre cómo puede enseñarse a los traductores a utilizar dichas técnicas. Según hemos visto, los autores abogan por la enseñanza a los traductores de técnicas avanzadas de acceso a la terminología en recursos de distinto tipo, diccionarios, bases de datos, corpus e Internet (Clark et al. 2002).

Por un lado, la forma de clasificación que hemos presentado en este capítulo facilita el aprendizaje de técnicas de búsqueda en diccionarios, corpus e Internet, ya que las técnicas disponibles en recursos diferentes aparecen agrupadas en función de los mismos criterios, por lo que se fomenta la asimilación y el aprendizaje. Si un usuario domina las opciones de búsqueda de un recurso y aplica el mismo tipo de clasificación a otro recurso que desconoce, el proceso de aprendizaje será más rápido ya que podrá aplicar los conocimientos de búsqueda en una herramienta que conoce a la búsqueda en una herramienta que desconoce. Asimismo, al docente, las clasificaciones estándar de técnicas de búsqueda le aportan una guía para la enseñanza, así como homogeneidad y claridad en las explicaciones de técnicas de búsqueda similares en recursos diferentes.

Por otro lado, estas clasificaciones también facilitan la evaluación y comparación de las herramientas, ya que podemos ver fácilmente qué técnicas de búsqueda desarrolladas en un recurso todavía no se han implementado en otro recurso y sería deseable que se incluyeran para mejorar el rendimiento de las búsquedas. En este sentido, nos ayudará a vislumbrar las mejoras que pueden realizarse en los recursos.

Así, para mejorar los resultados de las búsquedas, aparte de enseñar a los usuarios a utilizar las herramientas, es necesario continuar perfeccionando los recursos. La creación de clasificaciones estándar de técnicas de búsqueda facilitaría esta tarea a los desarrolladores de nuevas herramientas y a los encargados de mejorar las que ya existen.

Hemos adelantado algunos de los avances, como el aprovechamiento del medio electrónico para la generación de recursos plenamente electrónicos que incorporen técnicas de búsqueda innovadoras, algunas de las cuales hemos observado que, en principio, favorecen una aproximación onomasiológica (Geeraerts 2000; Oppentocht y Schutz 2003: 222).

Entre ellas, hemos visto la consulta dentro de los campos de contenido de los artículos de un diccionario para recuperar lemas o entradas, por ejemplo la búsqueda a partir de palabras clave en las definiciones u otros campos de contenido como los de

relaciones (tanto semánticas como sintagmáticas), y a partir también de información de tipo gramatical, etimológica, o de otra índole.

Además, se está fomentando el desarrollo de metodologías que se basan en la combinatoria léxica para la elaboración de diccionarios. Por ejemplo, hemos examinado diccionarios que utilizan la Lexicología Explicativa y Combinatoria (Mel'čuk et al. 1995). Dentro de estos diccionarios también se aboga por el acceso onomasiológico a las colocaciones.

Igualmente, se advierte un interés por la integración de los recursos para sacar el máximo partido de la información que ofrecen las distintas herramientas, por ejemplo el acceso a la consulta en corpus o en foros dentro de los diccionarios, o la posibilidad de integrar la búsqueda en Internet dentro de cualquier recurso (diccionarios, programas de gestión de corpus o programas de traducción asistida).

En definitiva, el análisis de *técnicas de búsqueda* efectuado en este capítulo nos ha servido para conocer lo que ofrecen los recursos desde un enfoque técnico. Esto constituye un punto de partida para vislumbrar, en el siguiente capítulo, desde un enfoque centrado en los usuarios, las necesidades y los usos que realizan los traductores de los recursos para explotar y obtener el máximo beneficio de los mismos dependiendo de la situación; es lo que denominaremos *estrategias de búsqueda*.

## **PARTE III. ESTRATEGIAS DE BÚSQUEDA**



## 4 Estrategias de búsqueda en traducción y nuestra propuesta de clasificación

[...] I will be shamelessly selfish and ask for the impossible. I will advocate for a dictionary that will always adapt to my needs and always be ready to provide me with exactly the answer that I need and will also agree with. I also expect the dictionary to be able to give satisfactory answers to those questions that I forget to ask. So I will require that the dictionary anticipate my needs and remind me of alternative routes that I can take in my search to find an appropriate way to express what I wish to express.

(Varantola 2002b: 31)

Mientras que en el capítulo anterior nos centrábamos en las técnicas de búsqueda que ofrecen los recursos e identificábamos tres elementos (la consulta, el instrumento y el resultado), en este capítulo incorporamos nuevos factores que intervienen en una búsqueda y que no dependen tanto del recurso, sino del usuario. La combinación de todos estos elementos en una búsqueda es lo que denominamos *estrategia de búsqueda*.

En la misma línea que sugiere Varatola (2006: 225) será necesario investigar por qué los traductores buscan determinadas palabras o elementos léxicos en diccionarios y corpus, abarcando también Internet como corpus, y qué es lo que realmente quieren obtener o conocer en diferentes situaciones, ya que según esta autora solo así se podrá incorporar en los recursos los conocimientos necesarios y se conseguirá integrar a los traductores en el diseño de estas herramientas.

Para empezar, tratamos una serie de corrientes de investigación que abogan por un enfoque orientado al usuario, en general en el ámbito de la recuperación de la información, y especialmente en el ámbito de la lexicografía (apartado 4.1).

Seguidamente, examinamos un conjunto de estudios sobre el uso que realizan los traductores en diferentes recursos, diccionarios (en el apartado 4.2), corpus (en el 4.3) e Internet (en el 4.4), aunque, como veremos, en una misma estrategia de búsqueda pueden intervenir recursos de distinto tipo. En cada uno de los apartados hacemos primero una introducción a las estrategias de búsqueda de los traductores en cada uno de estos recursos y, después, en los subapartados resumimos uno a uno los estudios por separado haciendo hincapié en los objetivos, los sujetos, las metodologías empleadas y brevemente los resultados obtenidos. El tratamiento que hemos dado a cada estudio que incluimos ha sido bastante exhaustivo porque queríamos reflejar todo el análisis de estudios que hemos realizado, ya que cada autor aporta estrategias de búsqueda diferentes y los estudios se han realizado adoptando distintos puntos de vista. Así, el lector de este trabajo puede remitirse,

si así lo desea, a las estrategias de búsqueda concretas propuestas por un autor determinado. No obstante, ofrecemos recapitulaciones más breves, al final de cada uno de los apartados, que recogen las ideas más importantes del análisis sobre estrategias de búsqueda, en diccionarios (al final del apartado 4.2), corpus (apartado 4.3) e Internet (apartado 4.4).

De forma paralela, al análisis de la bibliografía sobre estrategias de búsqueda de los traductores en diccionarios, corpus e Internet, hemos sintetizado los resultados del análisis de estos estudios en el anexo 2, de manera que hemos detectado una serie de tendencias en las estrategias de búsqueda de los traductores que se repiten en la investigación. Finalmente, de la síntesis y ordenación de estas estrategias de búsqueda detectadas en los estudios, hemos elaborado una propuesta preliminar de clasificación de estrategias de búsqueda en traducción, que presentamos en el apartado 4.5.

Con todo ello, en este capítulo veremos, por un lado, en qué punto se encuentra la investigación en materia de estrategias de búsqueda en traducción y, por otro, clasificaremos una serie de estrategias de búsqueda que se han detectado en la bibliografía, con lo que extraeremos conclusiones que podremos tomar como punto de partida para el diseño de nuestro estudio empírico sobre estrategias de búsqueda onomasiológica en traducción.

## **4.1 Enfoque orientado a los usuarios**

Los usuarios se han considerado desde hace tiempo un factor relevante en la elaboración de los recursos, en general en el ámbito de la recuperación de la información, y también concretamente en la elaboración de recursos lexicográficos, como diccionarios.

### **4.1.1 Recuperación de la información orientada a los usuarios**

En la recuperación de la información, las medidas utilizadas habitualmente para evaluar el éxito de una búsqueda son la precisión (*precision*) y la exhaustividad (*recall*). La precisión se calcula dividiendo los documentos relevantes recuperados entre el total de documentos recuperados. Por tanto, mide la proporción de documentos recuperados que son relevantes a la búsqueda. La exhaustividad se deduce de los documentos relevantes recuperados divididos entre los documentos relevantes que se incluyen en la base de datos consultada. Es, por ende, la proporción de documentos relevantes que se han recuperado.

No obstante, según Baeza-Yates y Ribeiro-Neto (1999: 83) la precisión y la exhaustividad se basan en el principio de que la cantidad de documentos relevantes a una



consulta es siempre la misma, por lo que no depende del usuario. Sin embargo, cada usuario puede realizar una interpretación distinta de qué resultados considera relevantes dependiendo de sus necesidades, por lo que estos autores proponen que se empleen también medidas de evaluación centradas en el usuario.

La interpretación de los resultados es un elemento esencial de toda estrategia de búsqueda. Cuando el usuario introduce una consulta en un recurso lo hace para encontrar el resultado que le permita alcanzar su objetivo. El hecho de que la búsqueda satisfaga o no las expectativas o necesidades del usuario al consultar el recurso determinarán en última instancia el éxito de la búsqueda. De esta manera, la investigación en el desarrollo de interfaces interactivas para los usuarios se basan en que la recuperación eficaz de la información depende en gran medida de la obtención de *feedback* por parte del usuario (Baeza-Yates y Ribeiro-Neto 1999: 96).

En este sentido, los autores investigan formas de expandir las búsquedas para acceder a resultados relevantes para los usuarios. Algunas estrategias que se estudian en los sistemas para la expansión de búsquedas son la utilización de sinónimos, tesauros, *stemming*, etc. Asimismo, la estrategia de reformulación conocida como *user relevance feedback* consiste en que, dada una lista de documentos recuperados por el sistema en una búsqueda, el usuario tiene la posibilidad de indicar aquellos documentos que considera relevantes a su búsqueda y seleccionar términos o expresiones importantes. En una nueva búsqueda se da prioridad a estos términos validados por el usuario, de manera que previsiblemente se obtendrán más resultados relevantes y menos no relevantes (Baeza-Yates y Ribeiro-Neto 1999: 118).

#### 4.1.2 Lexicografía orientada a los usuarios

La importancia de los usuarios en el diseño de diccionarios se refleja en algunas teorías lexicográficas<sup>77</sup> que ensalzan esta idea (Bergenholtz y Tarp 2003; Wiegand 1998, 2001). Gouws (2011) hace un repaso de la historia de la lexicografía teórica y observa tres fases, una primera fase centrada en los datos lingüísticos que se incluyen en los

---

<sup>77</sup> Autores como Bergenholtz y Tarp (2010: 37) indican que, aunque terminografía y lexicografía especializada son lo mismo, los terminólogos y terminógrafos no han centrado la discusión en las necesidades específicas de los usuarios, como muestran las publicaciones sobre terminología, salvo algunas excepciones como L'Homme (2006) y Bergenholtz y Nielsen (2006), según apuntan estos autores; véase también Faber et al. (2010) y Cabré (2007). Para Bergenholtz y Tarp (2010: 37): «In both fields the user is often mentioned, also the promise of user-friendly results, but only seldom is a real discussion of dictionary functions to be found».

diccionarios, que este autor ha denominado *linguistic colonialism*. Una segunda fase, con Wiegand a la cabeza, en la que se produce un distanciamiento del tratamiento basado en el contenido lingüístico, se concibe a la lexicografía como una disciplina independiente, no como una subdisciplina de la lingüística, y se centra principalmente en la estructura de los diccionarios. En una tercera fase, promovida por Bergenholtz y Tarp, se propone un acercamiento funcional, en el que los usuarios son el centro de atención.

Algunos autores no conciben la existencia de una teoría lexicográfica, sino que afirman que la lexicografía es un oficio, el arte de preparar diccionarios (Atkins y Rundell 2008: 3; Béjoint 2010: 381).

Independientemente de que se pueda concebir una teoría lexicográfica o no, o de las divergencias que puedan existir entre distintas teorías (por ejemplo entre la teoría general de la lexicografía y la teoría de las funciones lexicográficas, que veremos más adelante), lo que nos interesa en este apartado, y en lo que hemos visto que coincide la bibliografía, es en la importancia de abordar la elaboración de diccionarios, y de los recursos, desde un acercamiento basado en los usuarios.

En este sentido, se insiste en indagar cuáles son las necesidades de los usuarios de cara a elaborar diccionarios acordes a dichas necesidades. Barnhart (1962: 161) menciona la importancia de que los diccionarios respondan a las necesidades de los usuarios. Martin y Al (1988: 394) indican tres aspectos que deben tenerse en cuenta en la elaboración de diccionarios para que resulten útiles a sus usuarios: el conocimiento del usuario, sus estrategias (de búsqueda), sus finalidades/necesidades. Sobkowiak (1999: 275) sugiere que los diccionarios sean intuitivos (user-friendly) y que implementen funciones que permitan su personalización, flexibilidad e interactividad con el usuario. Por su parte, Atkins y Rundell (2008: 5) indican que el diseño de diccionarios que tengan en cuenta a los usuarios va más allá de la elaboración de diccionarios intuitivos, deberá tenerse en cuenta al usuario y sus necesidades en todos los aspectos de la elaboración de un diccionario, tanto en el contenido como en el diseño.

[...] the most important single piece of advice we can give to anyone embarking on a dictionary project is: know your user. [...] This doesn't imply a superficial concern with 'user-friendliness', but arises from our conviction that the content and design of every aspect of a dictionary must, centrally, take account of who the users will be and what they will use the dictionary for.

(Atkins y Rundell 2008: 5)

En este sentido, Atkins y Rundell (2008: 24-25) plantean una serie de cuestiones previas que deben plantearse en la elaboración de un diccionario, respecto de las propiedades del diccionario:

1. **la(s) lengua(s) del diccionario:** es...
  - a. monolingüe
  - b. bilingüe: en ese caso, es...
    - (1) unidireccional o
    - (2) bidireccional
  - c. multilingüe
2. **la cobertura del diccionario:** es...
  - a. lenguaje general
  - b. material enciclopédico y cultural
  - c. terminología y sublenguajes (p. ej., un diccionario de términos legales, cricket, enfermería)
  - d. área específica del lenguaje (p. ej., un diccionario de colocaciones, verbos frasales o modismos)
3. **el tamaño del diccionario:** es...
  - a. edición estándar (o 'collegiate')
  - b. edición abreviada
  - c. edición de bolsillo
4. **el formato del diccionario:** es...
  - a. impreso
  - b. electrónico (p. ej., en DVD o para PDA)
  - c. en línea
5. **la organización del diccionario:** es...
  - a. de la palabra al significado (la más común)
  - b. del significado a la palabra (cuando la búsqueda de una palabra lleva a otras palabras relacionadas semánticamente)
6. **la(s) lengua(s) de los usuarios:** el diccionario es para...
  - a. un grupo de usuarios en el que todos hablan la misma lengua
  - b. dos grupos específicos de hablantes
  - c. estudiantes que hablan lenguas distintas a la del diccionario
7. **los conocimientos de los usuarios:** son...
  - a. lingüistas u otros profesionales del lenguaje
  - b. estudiantes avanzados
  - c. escolares
  - d. niños pequeños
  - e. estudiantes de lenguas
8. **para qué utilizan el diccionario:** para una o ambas de las siguientes opciones...
  - a. decodificar, que es...

- entender el significado de una palabra
- traducir de una lengua extranjera a la lengua materna
- b. codificar, que es...
  - utilizar una palabra correctamente
  - traducir un texto de la lengua materna a una lengua extranjera
  - enseñar idiomas

(Atkins y Rundell 2008: 24-25, traducción propia del inglés)

Todo ello explica que se efectúen demandas, como la de Delavigne (2008: 167-168), que señala que la última revisión de un diccionario deberían realizarla los propios usuarios del diccionario, o incluso De Schryver y Prinsloo (2000: 197) que sugieren que la opinión de los usuarios se tenga en cuenta en todas las etapas del proceso de elaboración de un diccionario.

What is needed is a dictionary, obviously electronic, that is truly adaptive — meaning that it will physically take on different forms in different situations; and one that would do so as intelligently as possible — meaning that it would have the ability to study and understand its user, and based on that to learn how to best present itself to that user. With this, the field has moved to a very different paradigm indeed, to that of *adaptive and intelligent lexicography*, for short aiLEX.

(De Schryver 2009)

Atkins y Varantola (1998a: 29) hablan de las expectativas de los usuarios, *user expectations*. Varantola (2002b: 31) propone que los diccionarios se adapten a las necesidades de los usuarios y que sean capaces de responder a las preguntas que el usuario realiza e incluso a aquellas que no realiza, también se refiere a la elaboración de diccionarios *inteligentes* (Varantola 2006: 223). Habla de un diccionario personalizado a los perfiles de los usuarios (*user profiles*), pensando aquí en un usuario que precisa del diccionario en su actividad profesional, por ejemplo el traductor o el comunicador técnico (Varantola 2002b: 31).

#### 4.1.2.1 Teoría general de la lexicografía

Wiegand (1998), dentro de su teoría general de la lexicografía (*Allgemeinen Theorie der Lexikographie*), señala que los diccionarios deberán estar orientadas a las necesidades de los usuarios. A continuación presentamos un resumen proporcionado por este autor sobre las particularidades de lo que denomina *acciones de uso* (*Benutzungshandlungen*) y la terminología que emplea para referirse a ellas.

Particularidades de las acciones		Preguntas	Terminología	
Sujeto de la acción		¿QUIÉN?	<i>Usuario, aprendiz de usuario (del diccionario)</i>	
Efecto de la acción (lo que determina el tipo de acción)		¿QUÉ?	<i>Acción de uso</i> <i>acciones del tipo UTILIZAR UN DICCIONARIO</i>	
Modalidad de acción		¿CÓMO?	<i>Habilidades del usuario</i>	
Contexto externo de la acción	Contexto de la acción	¿EN QUÉ CONTEXTO?	<i>Contexto de uso (comunicativo o no comunicativo)</i>	
	Circunstancias de la acción	Momento de la acción (detalles de la fecha)	¿CUÁNDO?	<i>Momento del uso</i>
		Duración de la acción	¿CUÁNTO TIEMPO?	<i>Duración del uso</i>
		Lugar de la acción	¿DÓNDE?	<i>Lugar del uso</i>
Contexto interno de la acción	Objetivo del sujeto de la acción (se consigue si se obtienen los resultados deseados)	¿PARA QUÉ?	<i>Objetivo del usuario, objetivo parcial del usuario</i>	
	Motivo de la acción	¿POR QUÉ?	<i>Motivo del uso</i>	
	La causa de la acción es la circunstancia que se considera el motivo para la ejecución de la acción	¿POR QUÉ CAUSA?	<i>Causa del uso</i>	
	Resultados de la acción	¿CON QUÉ CONSECUENCIAS?	<i>Resultados del uso</i>	

**Tabla 25. Resumen de las particularidades de las acciones de uso (de un diccionario) y la terminología utilizada para referirse a ellas (Wiegand 1998: 292, traducción propia del alemán).**

Wiegand (1998: 284-291) explica estas acciones con un ejemplo. John está leyendo un texto en alemán mientras prepara una clase. En este texto aparece la frase «Mike war ein alter Hase, und man konnte ihm nichts vormachen». John no comprende el significado de la expresión *ein alter Hase sein*. En unos segundos, encuentra *\*ein alter Hase sein* en la tercera línea del artículo de la entrada *Hase* del diccionario *DUDEN-GW*, donde aparece una definición de esta expresión.

Según este autor, cada *acción (Handlung)* la desarrolla un *sujeto de la acción (Handlungssubjekt)*. El sujeto es la respuesta a la pregunta ¿quién? En este caso, John.

Cada acción tiene un *efecto (Handlungsergebnis)* que responde a la pregunta ¿qué?, ¿qué ha hecho John? John ha utilizado el diccionario. Sin este efecto, la acción no se habría ejecutado. Además, la acción se ejecuta para obtener unos *resultados (Folgen)*, en este caso comprender el significado de *ein alter Hase sein*. Puede haber preguntas más concretas, ¿qué tipo de diccionario? John ha utilizado un diccionario monolingüe general, ha utilizado el *DUDEN-GW*.

Lo siguiente es la *modalidad de acción (Handlungsmodalität)*, el ¿cómo? Las respuestas pueden ser variadas. Se pueden proporcionar datos a modo de *protocolo de uso del diccionario*. Por ejemplo, John ha cogido el diccionario, ha buscado la letra A, ha buscado la entrada correspondiente al lema *alt* y ha leído su artículo. Después, ha leído el

prólogo del diccionario, lo que le ha llevado a buscar el lema *Hase*. Ha revisado el artículo de *Hase* y ha encontrado la entrada *ein alter Hase sein*. Respondiendo a esta pregunta se proporciona una *descripción detallada* y una *interpretación* de la búsqueda que John ha llevado a cabo. De este modo, no se concibe una *única acción* (John ha utilizado el diccionario) sino *distintas acciones*, que aparecen ordenadas de un modo determinado.

Wiegand centra su explicación sobre las acciones de uso del diccionario en la *modalidad de acción* y explica que la buena investigación sobre el uso del diccionario deberá basarse en datos empíricos de descripciones detalladas de acciones de uso reales para conseguir un panorama general de las posibilidades de ejecución de acciones existentes hasta el momento sobre tipos y subtipos de acciones de uso del diccionario.

Otros elementos que intervienen en la acción de uso son el *contexto externo de la acción* (*äußerer Handlungskontext*) en el que se determina que la acción puede ser comunicativa, por ejemplo en la lectura de un texto, en una conversación, en el proceso de formulación de un texto. La acción también puede ser no comunicativa, por ejemplo en la obtención de conocimientos cognitivos de un tema. En el ejemplo de Wiegand, si se hace la pregunta ¿en qué contexto?, la respuesta es en la preparación de una clase, en la lectura de un texto en alemán.

En cuanto a las circunstancias (*Handlungsumstände*), se hacen varias preguntas, ¿cuándo? El viernes 20 de mayo de 1984, en la lectura de un texto en alemán, mientras John preparaba una clase; ¿cuánto tiempo?, por ejemplo, ¿cuánto tiempo ha necesitado John el diccionario? Solo unos segundos; ¿dónde? En su escritorio.

En el *contexto interno de la acción* (*innerer Handlungskontext*), encontramos el *motivo* (*Grund*) y el *objetivo* (*Ziel*) de la acción. El motivo se obtiene con la pregunta ¿por qué?, ¿por qué ha utilizado John el diccionario? Porque desconocía el significado de la expresión *ein alter Hase sein* (*motivo*), porque tenía problemas para la comprensión lectora (*circunstancia*). El objetivo se consigue con la pregunta ¿para qué?, ¿para qué ha utilizado John el diccionario? Para solucionar sus problemas de comprensión lectora, para encontrar una definición de *ein alter Hase sein*, para comprender un texto en alemán.

El sujeto de la acción tiene objetivos y la acción tiene resultados, o consecuencias. No todos los resultados son intencionados, puede haber resultados no intencionados. Se entiende que una acción ha tenido éxito, es decir, que se ha alcanzado el objetivo, si el sujeto obtiene los resultados deseados.

En definitiva, para Wiegand (2001: 235) toda situación de uso del diccionario comprende los siguientes tres elementos esenciales:

- un tipo de usuario
- un tipo de acciones de uso
- un tipo de diccionario

#### 4.1.2.2 Teoría moderna de las funciones lexicográficas

Por su parte, los investigadores del Center for Lexicography de Aarhus School of Business desarrollan la teoría moderna de las funciones lexicográficas (modern theory of lexicographic functions), cuyos avances desde la década de los 90 aparecen resumidos en Bergenholtz y Tarp (2003)<sup>78</sup>.

Según Bergenholtz (2003: 3) la lexicografía es una disciplina de derecho propio, no una disciplina lingüística ni una aplicación de la lexicología. Por otro lado, la lexicografía especializada, la terminografía, es parte de la disciplina de la lexicografía. Además, el centro de atención de la elaboración de diccionarios de una temática determinada deben ser las necesidades de los usuarios en situaciones específicas (Bergenholtz y Nielsen 2006; Nielsen 2009: 38).

Bergenholtz y Tarp (2003: 172) conciben los diccionarios como productos de uso que se elaboran para satisfacer ciertas necesidades humanas. Por tanto, toda consideración teórica o práctica debe tener en cuenta estas necesidades: lo que se necesita para resolver un problema determinado, que ocurre en un grupo concreto de usuarios, con características específicas, en situaciones concretas.

[...] dictionaries are considered utility products that are made in order to satisfy certain human needs. Consequently, all theoretical and practical considerations must be based upon a determination of these needs, i.e. what is needed to solve the set of specific problems that pop up for a specific group of users with specific characteristics in specific user situations.

(Bergenholtz y Tarp 2003: 172)

En este sentido, según la teoría de las funciones lexicográficas, los lexicógrafos deben estudiar las actividades humanas para detectar las posibles necesidades que deben satisfacerse por medio de un diccionario. Las necesidades humanas no son algo abstracto, sino que siempre están vinculadas a un grupo específico de usuarios en una situación determinada. El lexicógrafo debe establecer un perfil de un grupo de usuarios y una tipología de situaciones en las que aparecen problemas o necesidades que pueden resolverse con un diccionario.

---

<sup>78</sup> Véase también Tono (2010) que realiza una revisión de la teoría de las funciones lexicográficas.

Estas son las características generales que deberían conocerse para poder establecer un buen perfil de usuario, *user profile* (Bergenholtz y Nielsen 2006: 285). Para tipos de diccionarios concretos, puede haber otras que también sean relevantes.

1. ¿Qué idioma es su lengua materna?
2. ¿Qué nivel de dominio posee de su lengua materna?
3. ¿Qué nivel de dominio posee de una lengua extranjera?
4. ¿Qué experiencia tiene en la traducción entre las dos lenguas?
5. ¿Qué nivel de cultura general y conocimientos enciclopédicos posee?
6. ¿Qué nivel de dominio posee de un tema específico?
7. ¿Qué nivel de dominio posee de un lenguaje específico en su lengua materna?
8. ¿Qué nivel de dominio posee de un lenguaje específico en una lengua extranjera?

Esta información está vinculada a situaciones concretas que deben detectarse para determinar las necesidades específicas de un usuario. La teoría funcional de la lexicografía distingue al menos dos tipos principales de situaciones de los usuarios (cognitivas y comunicativas). En función de las necesidades identificadas en un grupo de usuarios, se puede determinar si dichas necesidades pueden satisfacerse con la consulta de un diccionario y qué tipo de información habría que proporcionar en el diccionario.

El primer grupo son aquellos usuarios que necesitan obtener información adicional sobre un tema, por ejemplo cultura general, información enciclopédica, información de un tema específico (biología, geología, etc.); o información sobre una lengua en el proceso de aprendizaje de la lengua, por ejemplo, para aprender una lengua extranjera. Estas situaciones se conocen como orientadas al conocimiento.

En este tipo de situaciones, donde el diccionario tiene una función cognitiva, el esquema de la consulta es el siguiente que propone Bergenholtz (2005: 20): *user* → *dictionary* → *user*<sup>79</sup>. Aquí los usuarios tienen necesidad de conocimientos enciclopédicos relacionados con el lenguaje, el lenguaje especializado, la cultura o cualquier área de conocimiento.

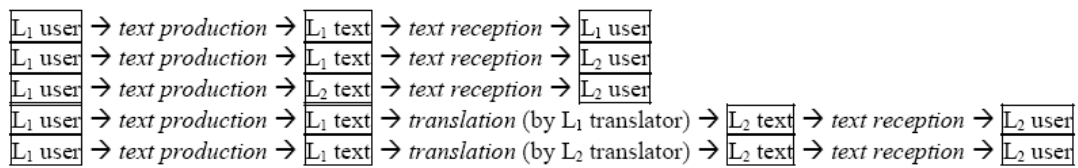
El segundo grupo comprende situaciones en las que hay comunicación oral o escrita y el diccionario interviene cuando se produce un problema de comunicación. Estas situaciones se conocen como orientadas a la comunicación. En este modelo, la comunicación puede darse en la recepción y producción de textos y, en ocasiones, en la

---

<sup>79</sup> Traducido del alemán al inglés en L'Homme y Leroyer (2009: 269).



traducción de dichos textos. Estos autores distinguen un total de cinco formas distintas de comunicación, representadas en la Ilustración 90.



**Ilustración 90. Modelo de comunicación según Bergenholtz y Tarp (2003: 174).**

En las funciones comunicativas, según este modelo, además del usuario y el diccionario, interviene un tercer elemento, un segmento de texto que es el que motiva una necesidad de búsqueda en el diccionario, y que cambia por completo el esquema de consulta: *user* → *text segment<sub>x</sub>* → *user* → *dictionary* → *user* → *text segment<sub>x</sub>* (Bergenholtz 2005: 20, traducido del alemán al inglés en L'Homme y Leroyer 2009: 270).

Bergenholtz y Tarp (2003: 175) indican seis *problemas de comunicación* que pueden resolverse con la ayuda de un diccionario.

1. Producción de textos en la lengua materna (o primera lengua)
2. Recepción de textos en la lengua materna (o primera lengua)
3. Producción de textos en una lengua extranjera (segunda, tercera lengua, etc.)
4. Recepción de textos en una lengua extranjera (segunda, tercera lengua, etc.)
5. Traducción de textos desde la lengua materna (o primera lengua) a una lengua extranjera (segunda, tercera lengua, etc.)
6. Traducción de textos desde una lengua extranjera (segunda, tercera lengua, etc.) a la lengua materna (o primera lengua)

Hay también otros problemas, como por ejemplo la traducción de textos entre dos lenguas extranjeras o la revisión de textos, que se considera una variante de la producción de textos. Identificadas estas situaciones, pueden determinarse las siguientes *necesidades de los usuarios*.

1. Información sobre la lengua materna
2. Información sobre una lengua extranjera
3. Comparación entre la lengua materna y extranjera
4. Información sobre cultura y conocimientos del mundo en general
5. Información sobre un área específica
6. Comparación de un área específica en la cultura materna y extranjera
7. Información sobre un lenguaje específico en la lengua materna

8. Información sobre un lenguaje específico en una lengua extranjera

9. Comparación de un lenguaje específico en la lengua materna y extranjera

Dependiendo de estas necesidades, el lexicógrafo debe decidir qué información incorporar en el diccionario. Cuando el lexicógrafo conoce el grupo de usuarios, sus características, situaciones y necesidades relacionadas, puede determinar las funciones lexicográficas.

La función lexicográfica (*lexicographic function*) de un diccionario es proporcionar ayuda a un grupo de usuarios específico con unas características concretas con el fin de satisfacer todo un abanico de necesidades que surgen en un tipo de situación. Un diccionario puede tener más de una función. Los diccionarios tienen una finalidad auténtica (*genuine purpose*) que se corresponde con el total de funciones de un diccionario y de áreas de especialidad que comprende (Bergenholtz y Tarp 2003: 176).

El elemento determinante de la función del diccionario es la situación de uso. Por ejemplo, existen diferencias significativas entre un diccionario para la producción textual en la lengua materna, y un diccionario para la traducción a una lengua extranjera, e incluso un diccionario para la producción en la lengua materna por parte de usuarios adultos o niños. Usuarios diferentes, con distintas competencias y en situaciones dispares varían la manera en la que deben satisfacer sus necesidades lexicográficas por medio de los diccionarios.

Las funciones orientadas a la comunicación más relevantes de un diccionario son:

- ayudar a los usuarios en la resolución de problemas relacionados con la comprensión textual en la lengua materna;
- ayudar a los usuarios en la resolución de problemas relacionados con la producción textual en la lengua materna;
- ayudar a los usuarios en la resolución de problemas relacionados con la comprensión textual en una lengua extranjera;
- ayudar a los usuarios en la resolución de problemas relacionados con la producción textual en una lengua extranjera;
- ayudar a los usuarios en la resolución de problemas relacionados con la traducción de textos de la lengua materna a una lengua extranjera; y
- ayudar a los usuarios en la resolución de problemas relacionados con la traducción de textos de una lengua extranjera a la lengua materna.

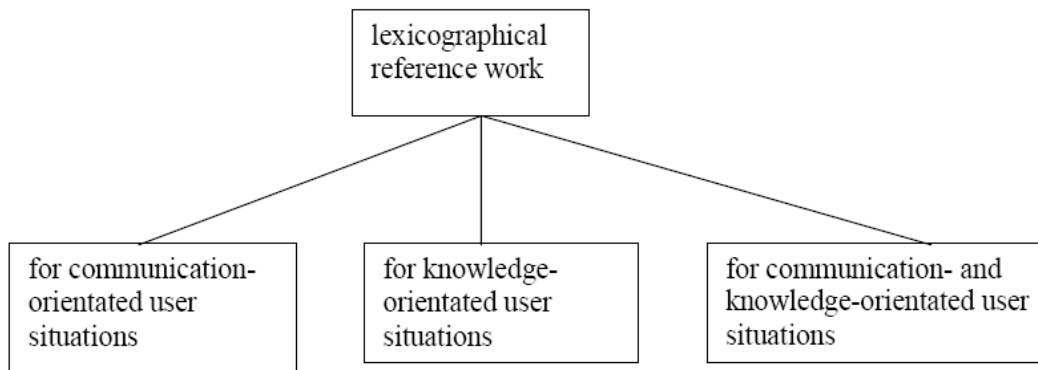
Las funciones orientadas al conocimiento más importantes son:

- proporcionar información enciclopédica y de cultura general a los usuarios;

- proporcionar información sobre el área de especialidad al usuario;
- proporcionar información sobre la lengua al usuario.

Según estos autores, en un diccionario todo está influenciado por las funciones. Estas funciones son los elementos básicos de la teoría y práctica lexicográfica, sin las cuales no se puede concebir un diccionario, tanto en lo que respecta al contenido como a la forma del diccionario. Un usuario consulta un diccionario para obtener información que le resulte útil para resolver un problema concreto o aumentar sus conocimientos de un tema. El diccionario deberá satisfacer las necesidades del usuario en dichas situaciones. Las necesidades pueden ser muy simples y pueden cubrirse con poca información lexicográfica, o pueden ser muy complejas y requerir diferentes tipos de datos lexicográficos.

Así, estos autores clasifican los diccionarios, según sus funciones, en obras de referencia lexicográfica para situaciones orientadas a la comunicación, situaciones orientadas al conocimiento y situaciones orientadas a la comunicación y el conocimiento.



**Ilustración 91. Clasificación de los diccionarios según las funciones lexicográficas (Bergenholtz y Tarp 2003: 182).**

En este sentido, un diccionario etimológico, por ejemplo, es útil si el usuario quiere conocer la historia de una palabra, por tanto, es un diccionario para situaciones orientadas al conocimiento; sin embargo, este tipo de diccionario no resultaría útil para producir o traducir un texto, por lo que no podría considerarse un diccionario para situaciones orientadas a la comunicación.

Por otro lado, Tarp (2007; 2008c) añade un nuevo tipo de situación importante para la lexicografía (aparte de las cognitivas y comunicativas), las situaciones operacionales (*operational situations*). Estas son situaciones en las que los usuarios necesitan recibir instrucciones, y que tratan de satisfacerse con recursos, como por ejemplo folletos,

manuales o guías de usuario. Este autor indica que este tipo de recursos pueden aprovecharse de las mejoras de un tratamiento lexicográfico. Considera que, aunque no todos los manuales o guías de usuario están diseñados o concebidos como productos lexicográficos, no hay razón para no hacerlo ya que, al igual que los productos lexicográficos, estos recursos tratan de ofrecer un acceso sencillo y rápido a las consultas de determinados tipos de usuarios con unas necesidades de información específicas. Por ejemplo, las guías de usuario no suelen estar pensadas para una lectura de principio a fin, sino para consultas puntuales en las que se obtiene información instructiva. Tarp (2009a: 26; 2010: 51-52) también se refiere a la situación interpretativa donde el usuario interpreta las señales y símbolos del mundo que le rodea.

Respecto a los usuarios, Fuertes-Olivera y Arribas-Baño (2008: 3) señalan que las necesidades conceptuales y lingüísticas pueden diferir en los usuarios que utilizan diccionarios especializados, como vemos en la Tabla 26.

<b>User Type</b>	<b>Conceptual information in L1</b>	<b>Conceptual information in L2</b>	<b>Linguistic information in L1</b>	<b>Linguistic information in L2</b>
Expert (bilingual dimension)				•
Semi-expert (monolingual and bilingual dimension)	•	•	•	•
Layman and beginner (monolingual dimension)	•			
Translator and interpreter (monolingual and bilingual dimension)	•	•	•	•
LSP student		•		•

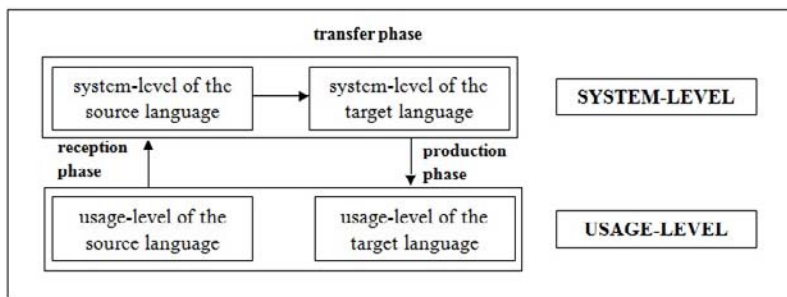
**Tabla 26. Tipología de necesidades lingüísticas y conceptuales de distintos grupos de usuarios de diccionarios de lenguas para fines específicos (según Fuertes-Olivera y Arribas-Baño 2008: 3).**

Por ejemplo, en Tarp (2010: 43-44) se resumen necesidades en el proceso de aprendizaje y Gouws (2010: 55-58) distingue varios tipos de aprendices.

Por otro lado, Bergenholtz y Tarp (2010: 32-33) se centran en el proceso de traducción, donde observan necesidades de tipo cognitivo y comunicativo, que se dan en distintas fases y subfases. En primer lugar, el traductor precisa adquirir conocimientos generales sobre el campo de la traducción. Para ello, debe leer sobre el área temática en cuestión. En segundo lugar, el traductor pasa por diversas fases para satisfacer ciertas necesidades. La primera fase es de lectura y comprensión del texto en lengua extranjera que, según estos autores, los traductores resuelven accediendo a explicaciones y

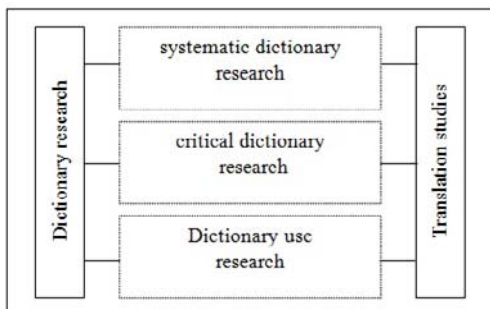
equivalentes en su lengua materna. La segunda fase es de transferencia, donde necesitan una solución bilingüe de los términos y colocaciones en la lengua origen con sus equivalentes en la lengua meta y, en el caso de que existan dos posibles equivalentes, es necesario que se aporte información sobre el significado para que el traductor pueda seleccionar el más adecuado. En la tercera fase, de producción en la lengua meta, el traductor necesita información gramatical y sintáctica, que incluya colocaciones, de los términos en lengua meta. En este sentido, según estos autores, esta función se satisface mejor cuanto más información se incluya de tipo gramatical, sintáctica, de colocaciones y ejemplos, una información que, según declaran, no se incluye en los diccionarios para traductores, lo que reduce sobremanera la calidad.

Según Fata (2010: 88), el diccionario puede proporcionar información relevante en tres fases del proceso de traducción, comprensión, transferencia y producción, como muestra la Ilustración 92.



**Ilustración 92. Modelo de proceso de traducción (Fata 2010: 88), según Muráth (2000: 159), basado en Gerzymisch-Arbogast (1996: 245).**

Así, Fata (2010: 84) propone que en la elaboración de un diccionario especializado bilingüe para traducción (*bilingual specialised translation dictionary* o *bilingual specialised dictionary for translation purposes*) es necesaria la cooperación entre las disciplinas de la lexicografía y los estudios de traducción (véase la Ilustración 93).



**Ilustración 93. Subcampos de la lexicografía en la intersección de la investigación en diccionarios y de los estudios de traducción (Fata 2010: 86).**

En la misma línea, Nielsen (2010: 72) propone que los lexicógrafos presten mayor atención a los resultados de la investigación en traducción para la producción de diccionarios especializados de traducción, además sugiere que las investigaciones se centren en preguntar a traductores expertos.

Lexicographers should pay more attention to the results and development in translation research so that they can produce specialized translation dictionaries that help learners translate as many of the features of LSP texts as possible, including linguistic structures and genre conventions, both of a verbal and non-verbal nature. One reason why lexicographers have been slow to adopt the findings of translation research is that they have asked learners and not translation experts. If lexicographers give more consideration to the advances in translation studies in future, they may be able to make improved dictionaries that help learners make high-quality translations of LSP texts.

(Nielsen 2010: 72)

La teoría moderna de las funciones lexicográficas se ha utilizado como base para la elaboración de diccionarios, tanto impresos como electrónicos. Por un lado, dentro de los diccionarios impresos, Tarp (2005) presenta un diccionario para situaciones orientadas a la comunicación y el conocimiento, *The Encyclopedic Dictionary of Gene Technology English-Spanish / Diccionario Enciclopédico de Ingeniería Genética Español-Inglés (EDGT)*. El autor explica que es un diccionario bidireccional y multifuncional, que puede ser útil para varios tipos de situaciones: recepción de textos en español, producción de textos en español, recepción de textos en inglés, producción de textos en inglés, traducción de español a inglés, traducción de inglés a español, información sobre temas de biología molecular, información sobre el uso del lenguaje molecular en español e inglés.

Para ello, el diccionario cuenta con una sección de introducción a la biología molecular, tanto en inglés como en español. Dicha introducción tiene 43 páginas organizadas en 4 capítulos e incluye una tabla de contenidos. Los términos más relevantes aparecen en negrita como referencias cruzadas vinculadas a dos listas alfabéticas de terminología inglés-español, español-inglés. Este sistema de referencias cruzadas también funciona en el *Nordic Dictionary of Lexicography* (Bergenholtz y Nielsen 2006: 291). Los artículos del diccionario *EDGT* contienen información sobre la categoría gramatical, colocaciones y ejemplos reales de textos extraídos de un corpus.

Este último diccionario, según señala Tarp (2005), puede ser útil para diversos usuarios, por ejemplo un periodista que quiere escribir un artículo para una compañía que trabaja en tecnología genética, y que consulta la introducción sobre biología molecular para adquirir conocimientos generales tanto lingüísticos como enciclopédicos. También puede utilizar el diccionario si está leyendo un texto de tecnología genética y desconoce el significado de algún término. Un farmacéutico también sería un usuario potencial de este

diccionario ya que, como usuario semiexperto, puede necesitar adquirir conocimientos del tema de la energía molecular con ayuda del diccionario. Además, puede utilizarlo un experto en el campo de la tecnología genética que tiene que preparar una comunicación en una lengua que no es la que habitualmente utiliza para la comunicación en dicho campo de especialidad. Otro tipo de usuario es el traductor no experto en el tema, que prepara una traducción sobre la tecnología genética. Según este autor, la consulta de la introducción puede ayudarle a preparar la traducción.

Para Bergenholtz y Nielsen (2006: 295) el *Legal Dictionary English–Danish* puede resultar útil en dos fases del proceso de traducción, la primera fase en la que el traductor adquiere conocimientos sobre la temática previamente a la traducción (función orientada al conocimiento) y la segunda fase en la que el traductor utiliza los datos del diccionario para traducir el texto (función orientada a la comunicación). En este sentido, las funciones de este diccionario son comprender textos jurídicos en inglés y traducirlos al danés. Los artículos contienen referencias culturales con referencias cruzadas a las secciones relevantes de un texto explicativo sobre el área temática (que está escrito en danés). Por ejemplo, este texto explica el concepto de *House of Lords* y los jueces que lo componen (función cognitiva), pero además sugiere posibles traducciones al danés de los términos con marcas culturales (función comunicativa). Véase el siguiente ejemplo.

§ 1 House of Lords

This is the highest civil court in England and Wales, and the House of Lords thus corresponds to the Danish Højesteret, cf. § 20.

(Bergenholtz y Nielsen 2006: 296)

Asimismo, el texto explicativo de *House of Lords*, aunque aportado en inglés por los autores en el ejemplo, se encuentra escrito en danés en el diccionario por dos razones: en primer lugar para facilitar la comprensión por parte del traductor danés semiexperto de las explicaciones dadas sobre un concepto de una cultura extranjera que le es ajeno (recepción textual) y, en segundo lugar, para que el texto proporcione equivalentes a modo de explicación que pueden emplearse en la traducción al danés.

Según estos autores, la solución ideal para resolver las funciones orientadas a la comunicación en una traducción es proporcionar estos textos explicativos en los dos idiomas, y presentarlos en dos columnas paralelas, con lo que se podrán comprobar los términos específicos y su uso (Bergenholtz y Nielsen 2006: 297).

Por otro lado, las tecnologías que se están desarrollando actualmente en el ámbito de la informática y las ciencias de la información pueden suponer una revolución en la teoría y práctica lexicográfica (según Tarp 2009b: 47-61; Tarp 2011b: 63).

Para este autor, la teoría funcional de la lexicografía, en su versión más madura, se ocupa de *tipos* de usuarios, *tipos* de situaciones de los usuarios, *tipos* de necesidades de los usuarios y *tipos* de datos que pueden satisfacer dichas necesidades. Sin embargo, también tiene en cuenta que en realidad no es un tipo de usuario el que realiza la consulta en un diccionario, sino un *usuario individual*, con unas *necesidades de información individuales*, en una *situación individual* y que realiza una *consulta lexicográfica individual* para acceder a *datos concretos* que satisfagan su necesidad individual. Por tanto, aunque justifica la necesidad de una tipología de usuarios, también considera que hay que tener muy en cuenta el tema de la individualización en la satisfacción de las necesidades de los usuarios.

Para Bergenholtz (2011: 30) el pensamiento lexicográfico no funcional lleva a la elaboración de diccionarios polifuncionales, que ofrecen a los usuarios la mayor cantidad de información posible para satisfacer cualquiera de sus necesidades. Sin embargo, un diccionario diseñado para ayudar a cualquier usuario en cualquier situación, no satisface a nadie (véase también Bergenholtz et al. 2009: 7). Según Bergenholtz y Bergenholtz (2011: 207), en una consulta los diccionarios polifuncionales recuperan todos los campos de la base de datos, pese a que este planteamiento provoca que al usuario le resulte complicado o imposible encontrar la información concreta que necesita.

Una buena herramienta es aquella capaz de satisfacer las necesidades del usuario permitiendo el acceso rápido a datos relevantes que cubran dicha necesidad; es lo que se denomina *accesibilidad* (Gouws 2010: 61; Tarp 2010: 41). Estos autores consideran que es posible diseñar muchos diccionarios monofuncionales, cientos o miles, utilizando una misma base de datos y diferentes posibilidades de búsqueda y presentación<sup>80</sup>. Nielsen y Almind (2011: 142) indican lo siguiente: «Databases are vessels that contain data and nothing else. They have no functionality *per se*». Por ejemplo, para resolver un problema de recepción, seguramente solo será necesario obtener información sobre el significado. Hoy en día algunos diccionarios considerados de producción son en realidad diccionarios

---

<sup>80</sup> En este caso es importante, como advierte Tono (2009: 60-61), que las etiquetas que utiliza el lexicógrafo para denominar a cada uno de los diccionarios sean lo suficientemente claras para que el usuario sepa cuál es el diccionario que debe consultar para cada necesidad o situación.



de recepción que ofrecen un exceso de información, mientras que la idea de la accesibilidad sería precisamente lo contrario, reducir la cantidad de información para cubrir las necesidades de producción de selección léxica (Prinsloo et al. 2011: 215).

Tarp (2011b: 68-69) explica que existe confusión terminológica para referirse a los diccionarios que se centran en funciones concretas para los usuarios, algunos lexicógrafos consideran que una herramienta que permite el acceso a dos o más tipos de datos monofuncionales es una agrupación de dos o más *diccionarios monofuncionales* que utilizan la misma base de datos, otros se refieren a un *diccionario pluri-monofuncional*, e incluso algunos indican que se trataría de un solo *diccionario multifuncional* porque se encuentra en un mismo sitio web y los datos se extraen de una misma base de datos.

Por tanto, este autor sugiere acabar con este desacuerdo proponiendo un nuevo escenario para la discusión, que denomina la individualización de la satisfacción de las necesidades de los usuarios. En este sentido, ya no se hablaría de acceso monofuncional o multifuncional, sino individualizado (aunque cada acceso individualizado tendría asignado un tipo de función específica); según este autor tampoco tendría sentido hablar de diccionarios individualizados dentro de una misma herramienta, por lo que en lugar de diccionarios monofuncionales o individualizados prefiere referirse a «one multifunctional dictionary with individualized search options within the framework of its defined functions» (Tarp 2011b: 69). La idea del diseño de diccionarios con posibilidades de individualización la comparten otros autores de la teoría funcional de la lexicografía (Bergenholtz 2011: 53). Estas opciones de personalización, como señala Tarp (2011b: 61), deberán ir más allá de la mera función de mostrar u ocultar los contenidos de algunos diccionarios, como por ejemplo en el *Oxford English Dictionary* o el *Macmillan English Dictionary Online* (opciones que se comentan en Lew 2011: 244).

As a consequence, the best dictionary in terms of needs satisfaction is not necessarily a monofunctional dictionary, but any dictionary –whether monofunctional, pluri-functional or multifunctional –that allows either monofunctional access or individualized access in the framework of its specific and foreseen functions.

(Fuertes-Olivera y Bergenholtz 2011: 6)

Asimismo, estos autores afirman que, en un diccionario electrónico, no solo debería tenerse acceso a una cantidad limitada de datos que se obtienen de una base de datos determinada, sino que debería considerarse el acceso a información *ilimitada* proveniente

de Internet<sup>81</sup>, y también deberían aprovecharse las ventajas que ofrecen los corpus electrónicos e Internet (Bergenholtz et al. 2009: 11).

Los recursos lexicográficos se valen de tecnologías que se están desarrollando en el campo de las ciencias de la información para permitir la personalización de los mismos<sup>82</sup>. En este sentido, desde la teoría funcional se sugiere que la lexicografía debería considerarse «part and parcel of a broader consultation discipline, or science, integrated into information science» (Tarp 2011b: 57); «[...] an interdisciplinary social and information science» (Leroyer 2011: 139). Estos autores proponen un cambio en la lexicografía, de un paradigma basado en la lingüística a las ciencias de la información (Leroyer 2011: 121; Tarp 2011a).

Asimismo, se considera que los usuarios de diccionarios son también usuarios de Internet, por lo que están acostumbrados a la carga y descarga de contenidos. Para estos autores, que los usuarios puedan contribuir al diccionario con sus aportaciones favorece la idea de la democratización lexicográfica (Fuertes-Olivera y Bergenholtz 2011: 2; Gouws 2009: 272). Los usuarios pueden añadir información directamente en el diccionario o en el recurso o realizar una referencia externa (Bothma 2011: 80), por ejemplo un diccionario puede referirse a otro diccionario para suplir algunas de sus carencias (Dung 2009: 157). No obstante, para incrementar la fiabilidad de las aportaciones de los usuarios en los recursos colaborativos, los autores proponen que se utilicen criterios editoriales precisos como que toda la información que se añade esté basada en referencias contrastables, como sucede en la enciclopedia *Wikipedia* (Dung 2009: 156-157; Fuertes-Olivera 2009a: 130). Además, será necesario un control por parte de los lexicógrafos de estas aportaciones (Bergenholtz 2011: 53; Fuertes-Olivera 2009a: 130).

En función de cuestiones como quién ha compilado el diccionario, para quién se compila y si el diccionario es gratuito o restringido, Fuertes-Olivera (2010: 209) distingue entre diccionarios en línea colectivos e institucionales y, dentro de los institucionales, los

---

<sup>81</sup> A este respecto, se menciona el acceso desde algunos diccionarios a materiales de referencia libres disponibles en Internet, como la *Wikipedia* o *Wiktionary* (Fuertes-Olivera y Niño-Amo 2011: 176). Meijssen (2009: 91) indica que la *Wikipedia* se encuentra entre las diez páginas web más consultadas del mundo.

<sup>82</sup> Véanse, a modo de ejemplo, las tecnologías propuestas por Bothma (2011: 80-99): búsqueda y navegación, modelado de usuarios, filtrado, adaptación de interfaces, etiquetado de metadatos, vinculación a información libre, sistemas de recomendación, sistemas de anotación. También la arquitectura multicapa en Spohr (2011). Heid (2011) realiza un estudio de usabilidad, propio de las ciencias de la información, a tres diccionarios innovadores, según la teoría funcional de la lexicografía.

divide en diccionarios en línea institucionales comerciales y restrictivos y diccionarios en línea institucionales gratuitos.

Lew (2011: 236-237) recopila algunos ejemplos de diccionarios en línea en los que se permite al usuario realizar aportaciones, en mayor o menor medida, aunque este autor advierte que el exceso de democracia en la elaboración de un diccionario no siempre es beneficioso. Por ejemplo, según este autor, en el *Urban Dictionary* los usuarios pueden votar por las mejores definiciones, aunque las definiciones más votadas no son siempre las que mejor reflejan el significado de una expresión; en *Wiktionary* algunas contribuciones de los usuarios contradicen las guías del diccionario (aspecto que también se observa en Fuertes-Olivera 2009a: 117); *Wordnik* incorpora contenido de citas e imágenes, que obtiene de *Twitter* y de *Flickr*. Otros diccionarios comerciales también quieren implicar a los usuarios, como el *Merriam-Webster's Open Dictionary* y el *Macmillan Open Dictionary*, donde los usuarios pueden mandar sugerencias a los administradores de nuevos vocablos que deberían incluirse en los diccionarios, lo que según Lew permite salvar las distancias entre los diccionarios institucionales y colaborativos. También se incluyen características de redes sociales como en el *Macmillan Dictionary blog*.

En cuanto a la adaptación de los recursos colaborativos a los usuarios, Désilets et al. (2007) muestran que los recursos colaborativos *Wikipedia*, *Wiktionary* y *OmegaWiki* no están diseñados para satisfacer las necesidades de los traductores, aportan una serie de sugerencias en este sentido e indican que *OmegaWiki* tiene mayor potencial para transformarse en un recurso colaborativo útil para traductores.

En relación con los avances en la adaptación de los diccionarios electrónicos a los usuarios, Tarp (2011b: 58-62) hace una distinción de cuatro categorías de diccionarios electrónicos en función de las características que los diferencia de los diccionarios en papel: *copycats*, *faster horses*, *Model T Fords* y *Rolls Royces*.

La primera categoría son los *copycats*, copias electrónicas de diccionarios en papel, escaneados y disponibles por ejemplo en formato PDF. La segunda categoría son los *faster horses*, diccionarios electrónicos cuya única ventaja que recogen del medio electrónico es el acceso rápido a la información. En esta categoría se incluyen, según este autor, la mayoría de diccionarios electrónicos, especialmente los disponibles en Internet. Cuando habla de *faster horses*, Tarp se inspira en la famosa frase que pronunció Ford antes de empezar con la producción de su nuevo coche Modelo T: «If I had asked people what they wanted, they would have said faster horses». Sin embargo, este y otros autores señalan que

«what we need in this respect is neither faster horses nor more horses, but horses of a different colour» (Andersen y Nielsen 2009: 359).

En este sentido, la tercera categoría propuesta por Tarp son los *Model T Fords*, en los que el usuario puede adaptar sus estrategias de búsqueda en el diccionario a diferentes tipos de necesidades y de situaciones del usuario. Según este autor los pocos diccionarios que entran dentro de esta categoría se han desarrollado siguiendo la teoría funcional de la lexicografía, como veremos a continuación.

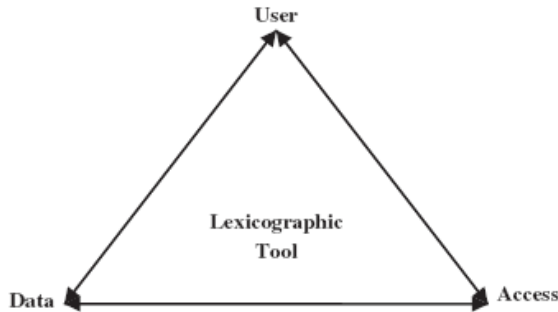
La cuarta categoría, los *Rolls Royces*, podrían ser herramientas adaptadas a usuarios individualizados en situaciones concretas. Según este autor, ninguno de los diccionarios disponibles en la actualidad podría clasificarse dentro de esta categoría y duda que vayan a desarrollarse diccionarios de este tipo en un futuro próximo (Tarp 2008b: 183). Para conseguir la elaboración de diccionarios de este tipo no solo será necesario adaptar las tecnologías disponibles en la actualidad sino diseñar otras nuevas (Tarp 2009a: 29, 2011b: 61).

Asimismo, según estos autores, los cambios que se están produciendo justifican que el objeto de atención de la lexicografía no sean solo los diccionarios, entendidos como diccionarios lingüísticos, sino otros recursos informativos que comparten características funcionales: el acceso a la información basado en las necesidades de los usuarios. Con el aumento de la información, los sistemas de información, los modos de almacenamiento, indexación, intercambio y acceso a la información, existe un riesgo de sobrecarga informativa (*information overload*) donde el usuario no será capaz de acceder a la información adecuada a tiempo y en la cantidad y para la finalidad que precisa. En este sentido, otros recursos de información y referencia podrían beneficiarse de esta nueva perspectiva, como almanaques, atlas, catálogos, directorios, guías, manuales, o incluso la página de inicio de un canal de televisión, de una compañía ferroviaria o aérea, o de una inmobiliaria (Leroyer 2011: 124-127).

La idea de los diccionarios como recursos electrónicos adaptados a los usuarios, que va más allá de la concepción del diccionario lingüístico como único objeto de estudio de la lexicografía, ha llevado a los autores a cuestionar si debería mantenerse el término *diccionario* como término científico (aunque pueda utilizarse en el ámbito comercial) y sea conveniente reemplazarlo por terminología más adecuada, como *lexicographical e-tool*,

*lexicographical information tool*, o *lexicographical consultation tool*<sup>83</sup> (véase Leroyer 2011: 124-127; Tarp 2011b: 69).

En línea con las teorías anteriores, Verlinde et al. (2010: 4) consideran que las obras de referencia (tanto diccionarios generales como especializados, enciclopedias, bases de datos terminológicas y léxicas, glosarios, etc.) deben adoptar un enfoque equilibrado basado en tres elementos (los datos, el acceso y el usuario) (también propuesto por Leroyer 2011: 128).



**Ilustración 94. El triángulo lexicográfico según Verlinde et al. (2010: 3).**

Según estos autores, los enfoques que se centran en los datos se preocupan exclusivamente de la manera de presentar los datos (siguiendo una teoría lingüística), sin embargo no tienen en cuenta que las consultas y el acceso a los datos pueden resultar complejos para el usuario si este no conoce la teoría lingüística que se ha empleado en la elaboración del diccionario, por ejemplo la Teoría Sentido-Texto y las funciones léxicas en la consulta del *Dicouèbe*<sup>84</sup>.

Los enfoques centrados solo en el usuario corren el riesgo de basarse en estudios de mercado donde los usuarios se convierten en consumidores y lo que prima es el beneficio por encima de las necesidades o los perfiles de los usuarios.

Los enfoques centrados en el acceso tienen el problema de que el recurso se convierta en una algama de opciones de búsqueda avanzada, donde el usuario debe convertirse en un especialista para poder utilizarlas. Verlinde y Binon (2009: 86) abogan por adoptar un enfoque más cualitativo que cuantitativo, en el sentido que la prioridad no sea la acumulación de recursos lexicográficos, sino la adaptación de los recursos a las necesidades, dificultades y hábitos de los usuarios.

<sup>83</sup> Anteriormente, se había hablado también de *leximat* (Tarp 2008a: 121).

<sup>84</sup> En este sentido, Almela et al. (2011: 10) proponen conseguir que el diccionario de colocaciones sea más accesible a los usuarios con la reducción al mínimo de los datos metalingüísticos, como en el *Dynamic Combinatorial Dictionary*.

Según estos autores, pese a que el acceso a la información parece el punto fuerte de los diccionarios electrónicos, este es uno de los aspectos que mayores problemas presenta. Por este motivo, abogan por una mayor investigación de la interacción entre los usuarios y los recursos para conocer las necesidades reales de los usuarios en determinados contextos.

La elaboración de distintos diccionarios electrónicos adaptados a diferentes tipos de situaciones de los usuarios a partir de una misma base de datos ya se ha puesto en práctica en algunos diccionarios basados en la teoría funcional de la lexicografía, la mayoría disponibles en el sitio web [ordbogen.com](http://ordbogen.com)<sup>85</sup>.

Por ejemplo, Bergenholtz (2011: 37) habla del acceso a cuatro diccionarios monofuncionales en línea de expresiones fijas, *Meaning of Fixed Expressions* (véase también Bergenholtz et al. 2011), que se obtienen de una misma base de datos. La base de datos contiene diferentes campos (campo principal, significado, significado adicional, gramática, comentarios, enlaces en línea, expresiones fijas, estilo, clasificación de la expresión fija, colocaciones y ejemplos). Dependiendo de la función de cada uno de los diccionarios, el sistema recupera unos campos u otros de la base de datos. En el primer diccionario, haciendo clic en el botón *I read the text, but do not understand the meaning of a fixed expression*, el usuario introduce una expresión o parte de una expresión en la casilla de búsqueda y obtiene la información de los campos de significado. En el segundo diccionario, haciendo clic en el botón *I am writing a text with a specific fixed expression*, el usuario consulta una expresión y recupera los campos que contienen información sobre el uso (significado, gramática, ejemplos, colocaciones, sinónimos, antónimos). En el tercer diccionario, con el botón *I am writing a text and am looking for a fixed expression with a specific meaning*, el usuario introduce en la casilla de búsqueda una o más palabras que reflejan un significado específico y recupera expresiones con dicho significado o parte del significado. En esta búsqueda el punto de partida es el significado para acceder a una denominación. En el cuarto diccionario, con el botón *I want to know as much as possible about fixed expressions*, el usuario introduce una expresión y obtiene toda la información de los tres diccionarios anteriores e información adicional.

---

<sup>85</sup> Para la descripción de estos diccionarios funcionales, nos basamos en las explicaciones proporcionadas por los autores. Las interfaces de búsqueda se encuentran en danés (idioma que no comprendemos) y son diccionarios de pago, solo puede realizarse un número limitado de búsquedas de prueba al día.

En otro ejemplo, los *Accounting Dictionaries* (Nielsen y Almind 2011: 144; Nielsen y Mourier 2005, 2007), se utilizan tres bases de datos monolingües (en inglés, danés y español), cada una de las bases de datos contiene campo de lema, gramatical, definición, sinónimos, antónimos, colocaciones, ejemplos, etc. Para el establecimiento de relaciones bilingües se utiliza el campo de definición, se buscan definiciones equivalentes entre las distintas bases de datos, con lo que se crea una nueva base de datos alternativa, que solo existe de manera virtual. Asimismo, por medio de diferentes opciones de búsqueda personalizadas, se permite la elaboración de doce diccionarios monolingües monofuncionales (y no solo tres diccionarios monolingües multifuncionales), así como nueve diccionarios bilingües monofuncionales (y no solo tres) (para más información, véase Nielsen y Almind 2011: 155-156). En este sentido, habría un diccionario de contabilidad monolingüe con función de recepción, en inglés, otro en danés y otro en español; un diccionario monolingüe con función de producción, en inglés, otro en danés y otro en español; un diccionario monolingüe con función para encontrar una palabra desconocida que refleje un significado conocido, en inglés, otro en danés y otro en español; y un diccionario monolingüe que muestre todos los datos, en inglés, otro en danés y otro en español. Asimismo, se mencionan diccionarios bilingües para ayudar a la comprensión de un término de contabilidad, a su traducción o a la traducción de una colocación o una frase.

Fuertes-Olivera y Niño-Amo (2011) presentan el *Diccionario Inglés-Español de Contabilidad*, un diccionario polifuncional que, según los autores, ayuda al usuario en tareas de comprensión y producción, así como traducción de textos de contabilidad inglés-español y en la adquisición de conocimiento sobre cuestiones de contabilidad.

En otra herramienta, *The Danish Music Dictionary, Musikordbogen* (Bergenholtz y Bergenholtz 2011: 195) se ponen a disposición de los usuarios tres diccionarios monofuncionales de música, uno con función de recepción, otro con función cognitiva y, finalmente, un diccionario polifuncional.

Andersen y Almind (2011: 217) elaboran, a partir de la misma base de datos, tres diccionarios monofuncionales de verbos frasales en inglés (uno de recepción, uno de traducción y uno de producción).

Verlinde et al. (2010) tienen en cuenta estos principios funcionales en la elaboración de la *Base Lexicale du Français (BLF)*<sup>86</sup>. Esta herramienta pretende cubrir diferentes

---

<sup>86</sup> Para acceder al análisis de este diccionario véase el apartado 3.2.3.

necesidades lexicográficas que pueden consultarse en seis casillas de búsqueda diferentes, en lugar de acceder a todas las opciones a través de una sola casilla de búsqueda como en otros diccionarios electrónicos.

Otra de las novedades que mencionan estos autores es la de asistencia inmediata en la lectura, escritura y traducción (Verlinde 2011; Verlinde et al. 2010). Según los autores, esta idea es el futuro de los diccionarios electrónicos, una herramienta interactiva que asista al usuario desde el mismo procesador de textos ofreciendo significados, sinónimos, traducciones, etc., sin necesidad de abrir nuevas aplicaciones.

La teoría funcional de la lexicografía se tiene en cuenta también en la mejora de los modos de acceso a la información en el diccionario electrónico de colocaciones de la informática *DiCoInfo* con el objetivo de que resulte útil al traductor en el uso de estrategias de búsqueda en dos situaciones de uso comunicativas en las que puede encontrarse: 1) en la traducción de la lengua materna a una lengua extranjera; 2) en la producción de un texto en lengua extranjera (L'Homme y Leroyer 2009: 258; L'Homme et al. 2010).

El primer modo de acceso es la búsqueda del equivalente de una colocación, por ejemplo el equivalente en francés de *send something as an attachment (envoyer... en pièce jointe; transmettre... en pièce jointe)*. El segundo es la búsqueda de colocaciones que expresan un significado determinado cuando se combinan con un término concreto, por ejemplo colocaciones que expresan la idea de «uso» cuando se combinan con *file (edit a file)*, con *mouse (move a mouse, click on ... with a mouse, drag and drop ... with a mouse)*, o con *Internet (browse the Internet, search ... on the Internet)*.

El *DiCoInfo* puede incorporar estas dos estrategias de búsqueda ya que contiene colocaciones en más de una lengua y, además, proporciona información semántica de las colocaciones, lo que en este diccionario se denominan funciones léxicas. En cuanto al primer modo de acceso, el *DiCoInfo* puede proporcionar equivalentes entre colocaciones aunque estas colocaciones no se encuentren vinculadas expresamente. Todas las colocaciones en este diccionario han sido descritas de forma sistemática en diferentes lenguas de acuerdo con sus propiedades lingüísticas (sistema de funciones léxicas) y, por tanto, los equivalentes entre colocaciones en distintas lenguas se establecen a partir de aquellas colocaciones que comparten las mismas funciones léxicas (L'Homme y Leroyer 2009: 274-275).

En el segundo modo de acceso, el objetivo es ayudar al usuario en situaciones de producción (en L1 o L2); el traductor puede buscar colocaciones de una palabra que expresan un significado determinado (Jousse et al. 2011). Este segundo tipo de búsqueda



permite a los usuarios acceder desde el significado a la denominación, aunque en las descripciones lingüísticas esta no es la dirección habitual, como afirman L'Homme y Leroyer (2009: 278), «[...] an efficient specialized lexicographic tool should also allow access to data for users searching for the right wording corresponding to a fairly precise idea of the meaning they wish to express». Verlinde et al. (2010: 6) también indican: «Access problems also occur when a user needs the right word in a given context (e.g. collocations) or when he hopes to discover an unknown word by looking up a familiar one, without using a translation dictionary [...]».

Este segundo modo de búsqueda en nuestro trabajo y en otros (por ejemplo Jousse et al. 2011) se entiende como búsqueda onomasiológica y, según indican estos autores, tiene una naturaleza más ambiciosa (en comparación por ejemplo con la búsqueda de equivalentes directos) y precisa de mayor investigación y desarrollo.

Foregrounding the onomasiological selection and presentation of the phraseological information can thus contribute to the enhancement of the functional integration of stylistic data types in the dictionary.

(Leroyer 2006: 191)

Actualmente, en la parte francesa del *DiCoInfo* se están produciendo avances en este sentido, por medio de la agrupación de las funciones léxicas en clases intermedias y clases genéricas. Estas clases intermedias y genéricas permiten clasificar las colocaciones de las fichas del diccionario de forma semántica y facilita el acceso de los usuarios por navegación, descendiendo en la estructura de ramas, de clases generales, a clases intermedias y finalmente a la información sobre las colocaciones (Jousse et al. 2011).

En definitiva, tomando como punto de partida al usuario, se pueden diseñar diccionarios o recursos orientados a usuarios de distinto tipo, como turistas (Leroyer 2009a, 2009b), por ejemplo en el diccionario para turistas *Nordisk leksikografisk ordbok* (Leroyer 2011: 134-137), ciudadanos daneses que quieren comprar una propiedad en Francia, en el diccionario *Ejendomsborg (Dictionary of Real Property)* (Leroyer 2011: 131-134), usuarios que quieren producir artículos científicos en inglés (Leroyer 2011: 137-138), un diccionario de vino para no aficionados a la bebida (Faber et al. 2010), pacientes de oncología (Delavigne 2008), revisores de traducción (Leroyer 2008) y también traductores, que son usuarios potenciales del diccionario de contabilidad de inglés y español (Fuertes-Olivera 2009b: 176; Fuertes-Olivera y Nielsen 2012; Fuertes-Olivera y Niño-Amo 2011).

Como hemos visto en este apartado, la investigación aboga por el estudio de los recursos desde el punto de vista de los usuarios. Además, esta es una tendencia que parece

que ha ido en aumento en los últimos años y que continuará así en el futuro. Por tanto, para la elaboración de recursos centrados en el usuario, será necesario investigar qué necesidades y qué estrategias de búsqueda esperan los usuarios que se faciliten en los recursos, tema que abordamos en los siguientes apartados.

## **4.2 Estudios sobre estrategias de búsqueda de los traductores en diccionarios**

Desde hace algunos años se advierte un creciente interés por el papel de los diccionarios y su uso como un integrante esencial en la fase de búsqueda del proceso de traducción (Roberts 1997). Fenner (Fenner 1989) coincide con esta autora en que este tipo de estudios en el ámbito de la traducción han tendido a subestimarse.

Esto queda patente en el contraste entre el ingente número de estudios sobre el uso del diccionario<sup>87</sup> en grupos de usuarios como nativos y estudiantes de segundas lenguas, mientras que tradicionalmente las investigaciones sobre el uso del diccionario en traducción han sido más limitadas, y además los principales sujetos de estudio han sido estudiantes de traducción, por lo que tampoco se han obtenido resultados acerca de las búsquedas efectuadas en un proceso de traducción *real*, o con traductores profesionales.

[...] the least explored area, viz. the reference skills or strategies translators bring to the task of finding what they are looking for in dictionaries. After all, there are numerous ways of arranging and accessing such information, from simple alphabetic lists to comprehensive thematic networks, and not all of them are mutually exclusive. Only close direct observation of translators in action, under laboratory conditions, will allow us to make valid generalisations about which dictionaries are to be recommended as user-friendly to translators.

(Hartmann 1989b: 16)

Las primeras investigaciones que se ocuparon de indagar sobre el uso del diccionario por parte de los usuarios tomaron como objeto de estudio a nativos de una lengua, especialmente de lengua inglesa. Notables ejemplos de estos trabajos son, entre muchos otros, y por orden cronológico, los de Barnhart (1962), Quirk (1973), Greenbaum et al. (1984), Miller y Gildea (1985), Kipfer (1987). Estos autores están interesados en averiguar principalmente la motivación que lleva a los usuarios nativos a utilizar el diccionario (por ejemplo obtener información sobre el significado de algunas palabras), el tipo de diccionario que prefieren (monolingüe o bilingüe), la información que les resulta más útil de las entradas del diccionario (definición, pronunciación, categoría gramatical,

---

<sup>87</sup> Obras como la de Welker (2010) recopilan estudios sobre el uso del diccionario. Tarp (2009c) efectúa una revisión sobre las metodologías empleadas en este tipo de estudios.

etimología, etc.) y, finalmente, el grado de satisfacción de los usuarios con el diccionario.

A partir de estos primeros estudios sobre el uso del diccionario también suscita interés un nuevo tipo de usuario, el estudiante de lenguas. Así, se publican investigaciones como las de Tomaszczyk (1979), Baxter (1980), Béjoint (1981), Galisson (1983), Bensoussan et al. (1984), Kharma (1985), Snell-Hornby (1987), Neubach y Cohen (1988), Tono (1989), Atkins et al. (1987), Atkins y Knowles (1988), Atkins y Varantola (1998a), Bogaards (1998), Wiegand (1998), Diab y Hamdan (1999), Hartmann (1999), McCreary y Dolezal (1999), Campoy Cubillo (2002), entre otros. Estos estudios están enfocados principalmente a conocer la utilidad del diccionario en el aprendizaje de lenguas, así como a comparar la utilidad del diccionario monolingüe frente al bilingüe. Algunos de estos estudios empiezan, además, a interesarse en cómo acceden los usuarios a la información del diccionario, cómo utilizan los diccionarios, qué estrategias emplean para encontrar la información que requieren (Hartmann 1987: 23), en operaciones efectuadas por usuarios típicos en acciones de consulta típicas (Bogaards 1999: 33). Algunos autores que abordan de alguna manera este aspecto son Béjoint (1981) y Atkins y Knowles (1988). También se ha cobrado interés por investigar el uso de diccionarios electrónicos y en línea (Bergenholtz 2011; Bergenholtz y Johnsen 2005; De Schryver y Joffe 2004; Kopleinig 2011; Müller-Spitzer et al. 2011), aunque estos autores sugieren que la investigación sobre el uso de diccionarios en línea todavía es un tema bastante nuevo (Müller-Spitzer et al. 2011: 203).

Asimismo, algunas de las investigaciones con estudiantes de lenguas, como por ejemplo la de Hartmann (1999), sirven de guía para el diseño de otros estudios posteriores con estudiantes de traducción como el de Sánchez Ramos (2004b). A su vez, personas que en investigaciones previas (Atkins y Knowles 1988) centran sus investigaciones en estudiantes de segundas lenguas, también indagan posteriormente en el uso del diccionario por otros usuarios como lexicógrafos y estudiantes de traducción (Atkins y Varantola 1997, 1998b).

Al mismo tiempo, en algunos de estos estudios con estudiantes de segundas lenguas también han participado traductores, aunque no puede asegurarse que sean representativos del uso de los diccionarios por parte de este grupo de usuarios. La muestra con la que cuentan dichos estudios no es homogénea, ya que combina por ejemplo en un mismo experimento estudiantes de lenguas y traductores. Finalmente, autoras como Mackintosh (1998: 125) consideran que algunos resultados obtenidos con los estudiantes de lenguas extranjeras pueden extrapolarse a los estudiantes de traducción, ya que comparten algunas

características, como la dificultad de decodificar la lengua extranjera (aunque los estudiantes de traducción deberían contar con un mayor dominio de la lengua extranjera), y también los estudiantes de lenguas extranjeras traducen de la L2 a la L1, pero más con el objetivo de aprender una lengua que de aprender a traducir.

Ha sido posteriormente cuando los investigadores han centrado su atención en un nuevo usuario, el traductor, como uno de los usuarios potenciales de distintos tipos de diccionarios. Como resultado de ello, surgen estudios como los de Meyer (1988), Tomaszczyk (1989), Roberts (1990), Atkins y Varantola (1997; 1998b), Varantola (1998), Rogers y Ahmad (1998), Mackintosh (1998), Fraser (1999), Livbjerg y Mees (1999; 2003), Corpas Pastor et al. (2001), Tarp (2002), De la Rosa Martín (2003), Sánchez Ramos (2004b), Prieto Velasco y López Rodríguez (2009) y Jousse et al. (2011), que veremos a continuación por este orden.

En los siguientes apartados resumimos estos estudios (metodología, sujetos, y sobre todo resultados de ejemplos de estrategias de búsqueda) para obtener un panorama general del uso y la utilidad del diccionario para los traductores. Trataremos de averiguar, a partir del análisis de estas investigaciones, qué necesidades motivan a los traductores a utilizar el diccionario, qué información buscan en el diccionario y en qué medida la información que encuentran les resulta satisfactoria para, finalmente, observar qué deficiencias y qué mejoras han propuesto los distintos autores, hasta el momento, en los diccionarios.

#### **4.2.1 Meyer (1988)**

Meyer se centra en el papel del diccionario bilingüe en la traducción inversa. Analiza los errores lingüísticos presentes en 42 traducciones inversas (del francés al inglés) realizadas por estudiantes de traducción francófonos de la School of Translators and Interpreters de la Universidad de Ottawa. El motivo de examinar los errores lingüísticos cometidos por los estudiantes es que estos vocablos o estructuras lingüísticas han planteado problemas y deberían haber motivado la búsqueda en el diccionario bilingüe. La autora consulta estas palabras en los diccionarios bilingües más utilizados por los estudiantes, *Robert & Collins Dictionnaire français-anglais/anglais-français* y *Harrap's New Standard French and English Dictionary*, y observa si estos diccionarios hubieran resultado útiles para resolver los problemas que han provocado el error.

En los resultados, se detectan dos dificultades: la selección de equivalentes en los diccionarios y, una vez elegido un equivalente, su uso o combinación en contexto.

Por un lado, el problema de selección de equivalentes puede estar motivado, en

primer lugar, por la ausencia del término de la LO en el diccionario. En este caso, Meyer observa una falta acusada en los diccionarios bilingües de expresiones multilexémicas (compuestos y modismos) y sintagmas equivalentes a vocablos simples en la LM. Por ejemplo, ninguno de los dos diccionarios consultados contiene el sintagma *moyenne de vie*, que es traducido erróneamente por los estudiantes como *average life*, cuando su equivalente correcto sería *average life expectancy*. Los diccionarios sí que contienen, por separado, las palabras *moyenne* y *vie*, así como expresiones multilexémicas de estructura similar, como *moyenne d'âge/des températures/des gens*, cuyos equivalentes son *average age/temperature/person*. Según la autora, los estudiantes de traducción, fijándose en estos ejemplos semejantes que sí que se encuentran en los diccionarios, pueden haber deducido erróneamente que *moyenne de vie* podía equivaler a *average life*.

En segundo lugar, el problema de selección de un equivalente puede deberse a la ausencia del equivalente correcto en el diccionario que, a su vez, puede estar relacionado con diferencias en la extensión semántica, estilística y en la categoría gramatical. Un ejemplo de diferencia de extensión semántica es la traducción errónea de *encouragements* por *encouragement*, sin embargo el equivalente adecuado en este contexto sería *incentive*, ya que el texto se refiere a un tipo concreto de *encouragement* que se otorga al cumplir un objetivo y que en inglés se conoce como *incentive*. El diccionario *Robert & Collins* no incluye *incentive* como equivalente de *encouragement*, aunque sí que incluye *encouragement* como equivalente de *incentive* en la parte inglés-francés del diccionario. El diccionario *Harrap's* sí que tiene *incentive*, pero solo en uno de sus ejemplos, por lo que no facilitaría la elección de este equivalente como la mejor opción.

En cuanto a la extensión estilística, *hélas* se traduce erróneamente por *alas*, en lugar de por otros equivalentes correctos *unfortunately* o *sadly*. Aunque los diccionarios consultados tienen mecanismos para indicar diferencias estilísticas, ninguno de los diccionarios especifica que *alas* no se utiliza en el inglés contemporáneo, y aparece como el equivalente de *hélas*. Solo el *Robert & Collins*, aunque incluye a *alas* como el equivalente de *hélas*, incorpora ejemplos en los que se utiliza el equivalente *unfortunately*. Finalmente, los diccionarios no suelen incluir información sobre la necesidad de cambiar de categoría gramatical cuando se traduce de una lengua a otra (también conocido como transposición).

En tercer lugar, el problema de selección de un equivalente puede ocurrir cuando el estudiante de traducción no es capaz de escoger el equivalente adecuado de entre la variedad de equivalentes propuestos por el diccionario. Para evitarlo, según Meyer, es

necesario que el diccionario haga distinciones semánticas y estilísticas entre los equivalentes propuestos.

Por otro lado, en el caso de que la selección del equivalente haya sido correcta, Meyer detecta también problemas para la utilización adecuada de los equivalentes seleccionados en contexto, principalmente en lo que respecta a cuestiones de morfología, sintaxis y combinatoria léxica.

En cuanto a la morfología, el error más común señalado es la conjugación errónea de los verbos irregulares y las dificultades de distinguir y utilizar correctamente los sustantivos incontables en inglés. Ambos diccionarios bilingües consultados incluyen información sobre las irregularidades morfológicas de los verbos, pero solo para los términos de la LO, no para los equivalentes, por lo que en la traducción inversa, el traductor tendrá que consultar el diccionario bilingüe en sentido inverso una vez detectado el equivalente o consultar un diccionario monolingüe para acceder a esta información. En cuanto a los sustantivos incontables, ambos diccionarios incluyen información de este tipo, pero no de manera sistemática para todos los sustantivos incontables que se incluyen en el diccionario.

Respecto a la sintaxis, el traductor que traduce en sentido inverso puede necesitar información sobre cómo se combina una palabra con otras palabras que dependen sintácticamente de ella, por ejemplo si algunos verbos rigen una preposición determinada. Los diccionarios bilingües incluyen información de este tipo, pero no de forma sistemática.

Por lo que se refiere a la combinatoria léxica, la autora indica que pueden describirse de forma sistemática con el uso de funciones léxicas, por ejemplo la función léxica Magn, que se refiere al significado «intenso», «muy», puede lexicalizarse dependiendo de la palabra con la que se combine como *raging storm*, *strictly prohibited*, *to bleed profusely*, etc. Los diccionarios bilingües analizados por Meyer y utilizados por los estudiantes del estudio incluyen alguna información sobre la combinatoria de los términos pero, una vez más, no de forma sistemática.

Meyer concluye de este estudio que, pese a los avances producidos, el diccionario bilingüe todavía no satisface por completo las necesidades de la traducción inversa. Señala que será necesario eliminar los errores presentes en los diccionarios y las omisiones, y acabar con la idea de que los diccionarios no pueden ser exhaustivos (esto último vendrá propiciado por los avances en las tecnologías). Además deberán establecerse nuevos principios lexicográficos y los que ya existen requerirán de una mayor formalización. Señala que los creadores de diccionarios bilingües tratan de que un mismo recurso sea

utilizado por un amplio número de usuarios, que se dedican tanto a la traducción directa como inversa, aunque las necesidades de estos dos tipos de traductores sean bastante distintas. Por último, esta autora indica que hasta que se desarrolle un diccionario bilingüe para la traducción inversa, los traductores que se dedican a este tipo de traducción deberán emplear los diccionarios bilingües disponibles de la forma más eficaz posible, lo que implica conocer las deficiencias y solventarlas, así como complementar al diccionario bilingüe cuando sea necesario con el uso de otros recursos, como el diccionario monolingüe, la documentación sobre la temática de la traducción, la consulta a nativos de la LM, etc.

#### **4.2.2 Tomaszczyk (1989)**

Este estudio trata las necesidades en el uso de diccionarios para la traducción inversa de textos científico-técnicos. Los resultados se obtienen de la propia experiencia del autor, que traduce un libro del polaco al inglés en un área de especialidad de la que no tiene conocimientos.

Los resultados indican que el traductor consulta el diccionario para la resolución de problemas con 691 unidades léxicas (una media de tres problemas por página traducida), de las cuales el 81,3% es terminología especializada y el 18,7% dificultades con colocaciones, propiedades sintácticas de las palabras, preposiciones, etc. De todas las consultas, el 54% se realiza porque la unidad léxica es desconocida para el traductor y el 46% para confirmar una intuición sobre la forma, la adecuación contextual, la ortografía, etc.

El diccionario proporciona soluciones para el 58% de los problemas con terminología especializada (donde el 70,9% son unidades simples y 46,4% términos compuestos), en el léxico no especializado la cifra de satisfacción aumenta al 79,1%.

El autor concluye que para conseguir un diccionario ideal para la traducción inversa, el diccionario bilingüe es esencial, aunque el diccionario monolingüe en lengua extranjera también es útil. Además reclama diccionarios especializados restringidos a campos muy específicos y que incluyan, entre otros datos, información sobre el uso de los términos, sus colocaciones, etc.

### 4.2.3 Roberts (1990)

Esta autora efectúa un análisis de diccionarios bilingües, con un objetivo similar al de Meyer (1988), pero desde el punto de vista de un traductor profesional y en la traducción directa. La autora simula el proceso que se seguiría en una traducción y estudia el papel que tendría el diccionario bilingüe en dicho proceso.

Roberts selecciona un texto en francés y su traducción al inglés efectuada por el Servicio de Traducción del Departamento de Comunicaciones de Quebec, y toma como muestra para su análisis los diccionarios bilingües francés-inglés más utilizados, según esta autora, por los traductores canadienses, que coinciden con los del estudio de Meyer, el *Harrap's New Standard* y el *Robert & Collins*.

En la simulación de la primera fase de una traducción, la de comprensión del texto origen, se procede a la lectura del texto, un análisis contextual, documentación y búsqueda de vocabulario o terminología. Como Roberts estudia el proceso desde el punto de vista de un traductor profesional, supone que dicho traductor comprendería la mayor parte del texto de origen y que deduciría del contexto aquellas palabras o términos desconocidos, por ejemplo *plaque tournante* y *transbordement*. El traductor supuestamente querrá confirmar si ha comprendido el significado de dichas palabras, por lo que, en lugar de utilizar el diccionario monolingüe para acceder a su definición, puede consultar el diccionario bilingüe y asegurarse de que realmente ha comprendido el sentido del término fijándose en los equivalentes propuestos. Por ejemplo, la autora busca *plaque tournante* en los dos diccionarios que utiliza para el estudio. En el *Harrap's* le resulta complicado encontrar el término, que se encuentra incluido dentro de la entrada de *plaque*. Una vez localizado el término en el diccionario, la autora solo encuentra el equivalente *pivot*, aunque por el contexto de la traducción sabe que no le sirve, ya que *plaque tournante* tiene dos sentidos, uno literal (que podría equivaler a *pivot*) y uno figurado (que no se incluye en este diccionario). Sin embargo, el diccionario *Robert & Collins* sí que contiene las dos acepciones, y además distingue entre ellas, por medio de etiquetas.

En el caso de *transbordement*, por el contexto de la traducción, se puede deducir que *transbordement* hace referencia a la navegación. Uno de los equivalentes propuestos por los diccionarios es *tran(s)shipment*, que incluye el morfema *ship*, con lo que el traductor puede concluir que este es el equivalente adecuado, aunque no puede estar seguro de que tenga el mismo significado que el término original, ya que el diccionario no incluye ninguna referencia al significado del equivalente. No obstante, algunos diccionarios



bilingües sí que incorporan definiciones de los equivalentes, aunque de forma poco sistemática. Por ejemplo, el *Robert & Collins* en ocasiones incluye entre paréntesis sinónimos o definiciones parciales.

La siguiente fase que simula esta autora es la de transferencia de una lengua a otra. En esta parte la utilidad del diccionario bilingüe parece más evidente, como herramienta para la selección de equivalentes. Roberts toma la traducción al inglés efectuada por el Servicio de Traducción del Departamento de Comunicaciones de Quebec, observa cómo se han traducido los términos originales (*tirer parti de, atout, pièce maîtresse* y *porte*) en el sentido figurado del texto, y comprueba si la búsqueda de estos términos en el diccionario bilingüe hubiera ayudado al traductor a decidirse por el equivalente que finalmente se ha proporcionado (*turn to the advantage, assets, king pin* y *gateway*).

Ninguno de los diccionarios bilingües consultados incluye todos los equivalentes y, aunque admite que los diccionarios bilingües no pueden incluir todos los equivalentes posibles de un término, ninguno de los equivalentes que se buscaron en el estudio eran extraños, con lo que cualquier traductor esperaría encontrarlos en un diccionario bilingüe. Además, para la selección del equivalente adecuado de traducción, los diccionarios bilingües analizados también adolecen de falta de información semántica que permita realizar distinciones semánticas y estilísticas entre los equivalentes. Como solución a este problema Roberts propone la separación de los significados de los distintos equivalentes de un término de la lengua original y que se incluya una pequeña definición de cada equivalente; el uso sistemático de etiquetas que indiquen el ámbito de uso de los equivalentes; la inclusión de actantes o argumentos para diferenciar entre equivalentes que tienen el mismo significado pero se utilizan en distintos contextos; notas de uso, etc.

En la última fase, la de revisión de la traducción, el traductor puede plantearse mejorar la forma de expresarse u otros detalles como la ortografía. Roberts observa que el diccionario bilingüe tampoco le hubiera resultado útil. Pone, por ejemplo, el término francés *transports*, que el traductor podría haber traducido como primera opción por *transport*, y que podría haber cuestionado en la fase de revisión, ya que otro posible equivalente sería *transportation*. En este caso, ni el diccionario *Harrap's* ni el *Robert & Collins* hubieran aclarado la duda del traductor, ya que el *Harrap's* no incluye el equivalente *transportation*, y el *Robert & Collins* propone los dos equivalentes, sin distinguir entre uno y otro.

Roberts concluye que los traductores deberán recurrir a numerosos diccionarios bilingües para tratar de encontrar soluciones a los problemas de traducción que les plantea

un texto, y también deberán hacer uso de diccionarios monolingües en LO y LM para complementar o verificar los equivalentes que les proporciona el diccionario bilingüe. Aunque, como admite la autora, se ahorraría mucho tiempo y esfuerzo si se elaboraran mejores diccionarios bilingües y, en concreto, diccionarios que trataran de satisfacer las necesidades de los traductores.

#### **4.2.4 Atkins y Varantola (1997; 1998b)**

Atkins y Varantola (1997; 1998b) efectúan un estudio sobre el uso del diccionario para resolver problemas lingüísticos en una traducción. Participan dos grupos de usuarios distintos. El primero está compuesto por 71 asistentes a un workshop de EURALEX, la mayoría lexicógrafos, con distintos niveles de lengua y habilidades para utilizar el diccionario. El segundo grupo está formado por 32 estudiantes de traducción finlandeses de nivel intermedio de la universidad de Tampere.

La metodología utilizada es similar para los dos grupos. Los participantes se agrupan por parejas, se les entrega un texto para traducir, que no es necesario que traduzcan, solo se les pide que identifiquen los problemas lingüísticos del texto y traten de encontrar una solución utilizando el diccionario.

Los participantes del grupo de lexicógrafos pueden elegir textos con distintos niveles de dificultad para traducir entre diferentes combinaciones lingüísticas, tanto de forma directa como inversa. Tienen disponibles más de 100 diccionarios para la consulta.

Los estudiantes de traducción deben traducir el mismo texto de forma inversa, del finlandés al inglés, y cuentan con seis diccionarios para la consulta (2 volúmenes de un diccionario bilingüe y 4 diccionarios monolingües de inglés).

Un miembro de la pareja busca en el diccionario, mientras el otro anota información sobre las búsquedas en una plantilla. Toda esta información se recoge en una base de datos, que contiene información de 1000 búsquedas, realizadas por 103 participantes (la suma de los dos grupos).

Se analizan conjuntamente los resultados de los dos grupos, excepto en algunos casos particulares. En un ejemplo de búsqueda, el participante quiere averiguar el equivalente en inglés para la expresión finlandesa *to'er* (expresión coloquial que designa un apartamento de dos habitaciones). En la primera consulta, se busca la palabra *to'er* en un diccionario bilingüe danés-inglés; esta palabra no se encuentra en el diccionario. En una segunda consulta, se busca la palabra *toværelses* en el mismo diccionario bilingüe; la entrada existe pero no contiene información sobre variables informales en inglés. En una tercera consulta,

se busca *two-room flat* en un diccionario monolingüe inglés; no se encuentra. En la cuarta consulta, se efectúa la búsqueda anterior en otro diccionario monolingüe; se encuentra pero no figuran variantes coloquiales en inglés. En una última consulta, el participante busca *flat* en un diccionario bilingüe inglés-danés; se termina la búsqueda, el usuario se muestra insatisfecho con el diccionario, finalmente no ha encontrado lo que buscaba.

Respecto a la información que buscan los usuarios en el diccionario, los participantes tienen que elegir entre una lista de opciones. Los resultados obtenidos se muestran en función de si los participantes han realizado una traducción directa o inversa. La información más buscada por ambos tipos de participantes es, por este orden, equivalente en L2 para término en L1, o viceversa, confirmación de que un término en L2 es el equivalente correcto, información gramatical sobre el término en L2, información sobre colocaciones de un término en L2 y otras informaciones (Atkins y Varantola 1997: 15).

En cuanto a la utilidad de los elementos de la entrada, los participantes dieron prioridad, por este orden, a la traducción o el equivalente de la entrada, ejemplo y equivalente, definición, modismo y equivalente, compuesto y equivalente y otros (Atkins y Varantola 1997: 18).

Sobre las ventajas de los diccionarios, los diccionarios bilingües resultan más útiles (64%) que los diccionarios monolingües (48%). En cuanto a la satisfacción de los resultados obtenidos, por cada 1000 búsquedas, el 37% resultan satisfactorias, el 22% parcialmente satisfactorias, y el 40% insatisfactorias. Atkins y Varantola (1997: 21) comentan que la satisfacción desciende en los usuarios con más conocimientos en L2 probablemente porque estos usuarios buscan información más compleja, son más experimentados y han aprendido a dudar del diccionario, o se han dado cuenta de la importancia de la información sobre colocaciones, o de otros inconvenientes de la traducción inversa.

En este estudio también se compara la satisfacción de los resultados entre el grupo de lexicógrafos y el de traductores. Las autoras comentan que el grupo de traductores muestra mayor insatisfacción con el diccionario probablemente por la formación recibida, que les ha hecho conscientes de los posibles problemas presentes en la traducción, por lo que solo se ven satisfechos cuando confirman la información en varias fuentes (Atkins y Varantola 1997: 26).

Atkins y Varantola centran parte del estudio en el grupo de estudiantes de traducción de Tampere para observar el tipo de búsquedas realizadas cuando el diccionario no ha proporcionado los resultados deseados (ya que de cada 1000 búsquedas, 400 resultan

insatisfactorias). La mayoría de los casos en los que el traductor ha recurrido al diccionario ha sido para buscar términos desconocidos, pero también en gran medida palabras generales cuya traducción depende del contexto, o combinaciones de términos conocidos. Los traductores reclaman en los diccionarios más ejemplos de uso comprensibles e información de colocaciones. En este sentido, las autoras suscriben la propuesta de Varantola (1994) de que un corpus monolingüe es más útil que un diccionario monolingüe o bilingüe para buscar equivalentes en contexto y para obtener información de las propiedades combinatorias de los términos.

En cuanto a la evolución del uso del diccionario monolingüe y bilingüe en las búsquedas de los traductores, se observa un uso muy elevado del diccionario bilingüe en las primeras búsquedas, mientras que el uso del diccionario bilingüe desciende a favor del diccionario monolingüe a medida que se suceden las búsquedas.

Como conclusión de este estudio, las autoras comentan que saber cuáles son las necesidades de los usuarios ayuda a los profesores a enseñar a sus alumnos a utilizar el diccionario; también a los creadores de diccionarios a diseñar diccionarios adaptados a las necesidades de los usuarios. Algunas de las principales necesidades que se han observado en este experimento son las colocaciones léxicas y las combinaciones sintácticas de una palabra, lo que sugiere la necesidad de transformar los diccionarios en papel en diccionarios electrónicos que contengan información de este tipo y que faciliten su acceso a los usuarios.

#### **4.2.5 Varantola (1998; 2003b)**

Esta autora plantea un estudio (véase 1998; 2003b) en el que aborda el uso de recursos de referencia en traducción inversa. Los sujetos son cuatro estudiantes de traducción avanzados.

En cuanto a la metodología, los estudiantes tienen que traducir un texto semiespecializado sobre pesca, de 477 palabras del finlandés (su lengua materna) al inglés (lengua extranjera), sin límite de tiempo. Durante la traducción, los estudiantes deben reflejar el uso de recursos en plantillas, que ya se habían utilizado en otros estudios (Atkins y Varantola 1998b).

Los resultados abarcan 84 búsquedas (*searches*), que forman un total de 128 consultas (*look-ups*), la mayoría en diccionarios bilingües finlandés-inglés (59%) y monolingües en L2 inglés (37%) y 5% en otros recursos (cuatro consultas en un glosario

especializado de pesca, una comprobación en el diccionario monolingüe en finlandés y una en una enciclopedia).

El 87% de las primeras consultas empiezan en el diccionario bilingüe finlandés-inglés y el 12% en el diccionario monolingüe inglés. En el diccionario bilingüe, el 55% busca un equivalente desconocido en L2, el 45% busca una confirmación o información adicional. En el diccionario monolingüe en L2, el 30% trata de confirmar una intuición, el 70% busca ayuda sin especificar sobre el uso de la expresión en L2.

La mayoría de expresiones problemáticas identificadas corresponden a vocabulario semitécnico. Algunos de los problemas planteados son: selección de la expresión adecuada en L2 para el contexto cuando ni el diccionario bilingüe ni el monolingüe proporcionan una respuesta, expresiones en L2 con las que los estudiantes de traducción no están familiarizados, ejemplos de uso de expresiones en L2, los diccionarios consultados no proporcionan el tipo de información requerida, dificultad de seleccionar la expresión apropiada en L2 cuando la información proporcionada es inadecuada y problemas con las colocaciones. Por ejemplo, problemas para expresar la idea en finlandés *heikko ikäluokka* (literalmente «weak + age group»), selección de expresiones estilísticamente inapropiadas (*get smaller* en lugar de *decrease*, *nowadays* en vez de *today* o *currently*, *was a bit over a quarter* en lugar de *accounted for over a quarter*). Finalmente los estudiantes tienen problemas para encontrar paráfrasis idiomáticas que expresen ideas del texto finlandés. Sin embargo, en este caso, no buscan formas de parafrasear en el diccionario ya que por experiencia saben que la posibilidad de encontrar esta información en los diccionarios es muy reducida.

La mayoría de consultas realizadas por los estudiantes son intentos (a menudo fallidos) de encontrar en los diccionarios mejores formas de expresar ideas en inglés. También se observa que las búsquedas por usuarios distintos en el mismo recurso no generan los mismos resultados, lo que sugiere una relación entre las habilidades de uso del diccionario de los usuarios y el éxito de la búsqueda.

Finalmente, se alude a la dificultad de encontrar determinada información léxica en los diccionarios, tanto bilingües, monolingües, generales o especializados. Por ejemplo, problemas para encontrar en los diccionarios impresos colocaciones que resultan difíciles de verificar (como *meagre age groups* o *reduced populations*) y también información estilística sobre expresiones en L2.

La autora concluye que los traductores necesitarán recursos adicionales a los diccionarios impresos, como por ejemplo textos paralelos (concebidos como textos escritos

originalmente y del mismo campo en lengua origen y lengua meta). Estos recursos facilitarían la confirmación de determinadas expresiones a través de la consulta de concordancias en corpus. De hecho, la autora observa intentos de incluir determinados tipos de textos estándar en los diccionarios en papel, iniciativa que podría ser más ambiciosa en los diccionarios electrónicos donde no existen problemas de limitación de espacio, y además también habría que esforzarse por que los traductores tuvieran acceso a textos paralelos generados espontáneamente y no solo textos estándar.

Varantola (2003b: 339) señala que las *necesidades de información* de los traductores en los recursos pueden ser de distinto tipo y que forman un *continuo* desde la búsqueda de un equivalente en lengua materna o lengua extranjera, un término de un campo de especialidad, información estilística o gramatical, información enciclopédica, etc.; pueden ser tanto un problema léxico o terminológico, pero también pueden abarcar fragmentos más largos o cuestiones estilísticas; puede tratarse de un problema múltiple o difuso.

equivalent	grammatical collocation	lexical collocation	examples	idiomatic usage	longer passage	paragraph structure	text structure	stylistic information	encyclopaedic information
------------	----------------------------	------------------------	----------	--------------------	-------------------	------------------------	-------------------	--------------------------	------------------------------

**Ilustración 95. El continuo de necesidades de información (Varantola 2003b: 339).**

#### 4.2.6 Rogers y Ahmad (1998)

Rogers y Ahmad (1998) se plantean cuál sería el tipo de diccionario que cubriría las necesidades de los traductores. Se tienen en cuenta dos cuestiones: la necesidad de información contextual y de conocimientos temáticos.

En lo que respecta a la primera necesidad, los traductores trabajan con formas verbales y expresiones en contexto, no con lemas, por lo que el diccionario debería aportar información sobre las formas y su uso en contexto. Lo explican con un ejemplo: en un diccionario bilingüe *Erkrankung* puede presentarse como equivalente de *disease*, sin embargo, factores de tipo combinatorio pueden afectar a dichas denominaciones, en lo que respecta por ejemplo a cuestiones como el número, su uso en singular o plural (*disease* o *diseases*) o su uso como término principal o modificador de un término compuesto (*acquired disease* o *disease progression*). Una concordancia en un corpus sobre ingeniería genética podría ayudar a resolver estas dudas. Según estos autores representar esta información en un diccionario con un formato de entrada tradicional, tanto en papel como en electrónico, supone enormes dificultades, dada la tendencia a los compuestos en la formación de términos. Una solución sería dar acceso a los traductores a corpus en las entradas de los diccionarios para identificar patrones que no son tan evidentes.

Otra necesidad relacionada son las colocaciones, como información sobre los verbos que se combinan con determinados sustantivos. Por ejemplo, en la traducción del inglés al danés, la traducción del verbo *engage* puede variar dependiendo de sus colocaciones. Este tipo de patrones se encuentran en los textos pero según los autores en pocas ocasiones se codifican en los diccionarios especializados de forma sistemática.

En lo que respecta a la selección de equivalentes, el traductor también debe asegurarse de que el término elegido tiene el mismo significado en el texto meta que en el texto origen, en el campo de especialidad. Sin embargo, los diccionarios bilingües proporcionan escasa información temática, por ejemplo en forma de definiciones. Estos diccionarios normalmente incluyen listas de sinónimos, en ocasiones diferenciados en función del sentido, como en la siguiente entrada:

*judicature* *s* Rechtsprechung; Rechtspflege, Justiz; Richterschaft, die Richter

(entrada de *judicature* en el diccionario *Juristisches Wörterbuch Englisch Deutsch*, Rogers y Ahmad 1998: 197)

En este ejemplo, *judicature* tiene tres sentidos, el segundo y el tercero con dos sinónimos cada uno. A partir de la consulta en este diccionario, el traductor debería recurrir a un diccionario monolingüe de alemán para acceder a las definiciones. Además, según estos autores, si el diccionario se ha compilado desde una perspectiva semasiológica, basada en las palabras, en lugar de onomasiológica, basada en los conceptos, una entrada puede contener numerosos sentidos polisémicos de distintos campos de especialidad, o incluso del lenguaje general, que no se etiquetan de manera adecuada o sistemática. Si los campos de especialidad se etiquetan de forma explícita y consistente en los diccionarios, en un diccionario electrónico el usuario será capaz de limitar la búsqueda a un área temática. En esta necesidad identificada en la traducción, los traductores para poder formarse un *mapa de conocimiento* del campo, recurren finalmente a otros recursos distintos al diccionario.

In order to compile what could be called a *knowledge map* of the domain, translators are in practice thrown back on a wide range of resources far beyond conventional dictionaries, including anything from academic subject experts to Yellow Pages to the receptionist at the local hotel, also a kind of expert for certain queries. In such cases, **the problem is to find a linguistic label in the target language for a new product, service or facility, where bilingual dictionaries, which provide access via the source language word, have failed to provide the answer. The challenge for the designers of a truly electronic dictionary may be put as follows: how to get from what may be described as characteristics of the concept** (as realised, for instance, in an analytical definition) **to the word or term:** what do we call the rectangular piece of plastic which can be inserted into a device on a hotel door in order to gain entry? *key card*.

(Rogers y Ahmad 1998: 199, negrita añadida)

Para estos autores, el diccionario del futuro será un recurso dinámico, más allá del diccionario en papel o de réplicas en formato electrónico. Estos autores vislumbran un nuevo diccionario que integre distintos recursos en uno.

The new dictionary will allow translators not only to draw on electronically-stored data-bases of terms with smarter semantically-relevant navigational paths, it will also support them in creating their own data-bases or simply help them to solve their problems in an ad hoc way by reference to large electronic corpora of text, a source of context-sensitive reference for language use. An integrated interface allowing access to a range of sources, including the Internet, following in the increasingly familiar workbench approach (*cf.* Holmes-Higgin & Ahmad 1992 for terminology; Clear 1993 for lexicography; Johnsen & Buhl 1993 for language for special purposes; Kugler, Ahmad & Thurmair 1995 for translation; Rogers 1996 for L2 learners) would allow the user/translator to extend and complement standard sources.

(Rogers y Ahmad 1998: 193)

Además dicho diccionario deberá ser especializado, ya que la mayoría de traductores profesionales trabajan con el lenguaje especializado. En este sentido, los traductores experimentados, según estos autores, utilizan una gran variedad de fuentes, desde documentales (libros de texto, informes, revistas, artículos académicos, traducciones previas, etc.), diccionarios y enciclopedias, y también expertos. De hecho, Rogers y Ahmad (1998: 195) afirman que el diccionario ideal para el traductor sería un experto bilingüe con acceso gratuito e ilimitado.

#### 4.2.7 Mackintosh (1998)

Mackintosh plantea un estudio sobre el uso del diccionario en la traducción directa, tanto inglés-francés, como francés-inglés, en el que efectúa dos experimentos.

En el primer experimento, los sujetos son 8 estudiantes de traducción de habla nativa inglesa y 7 de habla nativa francesa, que cursan un máster en la School of Translation and Interpretation (STI) de la Universidad de Ottawa, algunos de los cuales tienen algo de experiencia en traducción profesional, lo que se tiene en cuenta en el análisis de los datos.

La elección de estudiantes, en lugar de traductores profesionales, viene motivada, según las palabras de la propia autora, por la mayor disponibilidad de los estudiantes, por no disponer del tiempo y la financiación para acceder a traductores profesionales. Además, desde el punto de vista de la enseñanza indica que le resulta interesante examinar a los estudiantes, ya que estos son más dependientes del diccionario de lo que pueden ser los profesionales.

Los participantes deben traducir un texto especializado con la ayuda de distintos diccionarios en papel (monolingües, bilingües e híbridos, que combinan características



tanto del diccionario monolingüe como bilingüe; generales y especializados) y el banco terminológico TERMIUM.

Se seleccionan textos que propicien el uso de diccionarios de diversa índole, tanto para resolver problemas de comprensión como de expresión, términos simples, términos compuestos, polisemia, términos que se tratan de forma deficiente en los diccionarios, o cuyas definiciones son insuficientes o complejas.

Se utiliza una metodología inductiva, que combina la observación y grabación con Think-Aloud Protocol (TAP). Para simular al máximo una situación natural de traducción, se limita el tiempo para finalizar la traducción a 1 hora y 15 minutos y se pide a los estudiantes que produzcan una traducción, aunque no es obligatorio terminarla, frente a otros estudios como Atkins y Varantola (1998b) donde se pide a los usuarios que usen el diccionario como si tuvieran que realizar una traducción, pero no es necesario que traduzcan.

En cuanto a los resultados (Mackintosh 1998: 131), los tipos de problemas identificados que motivan el uso del diccionario son de comprensión de términos o conceptos en L2, de verificación de significados y de necesidad de equivalentes. Los más comunes son de comprensión en combinación con la necesidad de un equivalente, 172 de 496 búsquedas (34,7%), y de necesidad de un equivalente, en palabras de Mackintosh, cuando el sujeto alega que entiende un concepto en L2 pero desconoce el equivalente adecuado. Este último tipo de búsqueda se produce en 160 de las 496 (32,3%). Finalmente, se descubre que estos dos tipos de problemas más frecuentes se producen sobre todo en textos especializados.

De hecho, también se extrae del experimento que los elementos más consultados de los diccionarios son una definición, un equivalente o ambos, 348 de 362 búsquedas (96.1%) y solo equivalentes, 171 de 362 búsquedas (47,2%).

En cuanto a las búsquedas, aunque no se proporcionan datos detallados de los resultados del experimento, se sugiere que en los problemas de comprensión los estudiantes buscan definiciones, aunque en los campos especializados además de definiciones también consultan equivalentes cuando tienen problemas de comprensión. En los problemas de falta de un equivalente, normalmente buscan solo equivalentes, aunque también combinan en ocasiones la búsqueda de definiciones y equivalentes, sobre todo cuando tienen problemas para aceptar un equivalente porque no están seguros del concepto al que se refiere y si es el mismo que el que expresa el término original, principalmente en

la traducción especializada. La estrategia más satisfactoria es la búsqueda combinada de definiciones y equivalentes.

En cuanto a los diccionarios proporcionados en el experimento el más utilizado es el híbrido especializado, sobre todo en aquellos sujetos con mayor experiencia profesional, lo que sugiere que a los traductores podrían resultarles útiles diccionarios que combinen tanto definiciones como equivalentes en un mismo recurso, en lugar de definiciones y equivalentes aislados en el diccionario monolingüe o bilingüe (Mackintosh 1998: 132).

Respecto a los problemas de insatisfacción detectados en los diccionarios, de 203 búsquedas en las que los usuarios no encuentran la información que requieren, la entrada no se incluye en el diccionario en el 88,2% de los casos, y de 148 búsquedas en las que los usuarios quedan parcialmente satisfechos, el 29,1% no encuentra la entrada exacta sino un término similar o relacionado, el 14,9% quiere verificar el equivalente encontrado en otros recursos, el 12,8% necesita distinguir el significado entre dos equivalentes y el 12,8% no está satisfecho con los equivalentes proporcionados por el diccionario.

Estos datos revelan un alto porcentaje de insatisfacción para resolver problemas de expresión o de producción. De hecho, una información adicional proporcionada por este estudio es que los estudiantes verifican con frecuencia los equivalentes para asegurarse que los han comprendido y que significan lo mismo que el término de partida, así como que el equivalente seleccionado se utiliza en el tipo de texto de la traducción. Además, los estudiantes parecen no quedar satisfechos a no ser que comprueben la información en más de un recurso.

En el segundo experimento se investiga un aspecto más concreto, el metalenguaje de las definiciones de los diccionarios, a partir de las hipótesis generadas del primer experimento, como que los estudiantes, especialmente los anglófonos, evitan las definiciones en la L2, bien porque tienen dificultades en comprender el metalenguaje utilizado en la lengua extranjera, o bien el problema no es tanto el metalenguaje en L2 sino el metalenguaje complejo.

En esta prueba participan 65 estudiantes de traducción francófonos y 41 anglófonos, de grado o graduados. El método empleado en este caso es deductivo. Se proporcionan contextos en L2 con palabras inventadas que representan conceptos conocidos por los estudiantes. Se presenta también una definición en L1 o L2 por cada palabra inventada y, a partir de aquí, se pide a los estudiantes que generen un equivalente en L1 o, en el caso de que no fuera posible, una explicación en L1 basada en la definición. Además, se pide a los

estudiantes que subrayen los términos que no comprenden de la definición. A cada sujeto se le entregan 8 contextos en L2 con 4 definiciones en L1 y 4 definiciones en L2.

El resultado que se deriva es que los estudiantes prefieren las definiciones de los diccionarios en su lengua materna y tienden a evitar las definiciones que se encuentran en lengua extranjera. Asimismo, les resultan también complicadas aquellas definiciones en las que la información esencial no se comprende, como el descriptor y las características específicas.

Finalmente, Mackintosh señala una serie de implicaciones de este experimento para la enseñanza de la traducción, como la identificación de estrategias de consulta en los diccionarios que pueden abordarse en las clases de traducción, la necesidad de enseñar a los estudiantes a compensar la búsqueda de cantidades ingentes de información con las limitaciones temporales.

Respecto a las sugerencias en el diseño de diccionarios, esta autora aboga por diccionarios electrónicos en los que el acceso a la información sea sencillo y rápido, híbridos que contengan tanto definiciones como equivalentes, y acceso a definiciones tanto en L1 como en L2 para asegurar la equivalencia entre conceptos.

#### **4.2.8 Fraser (1999)**

Fraser (1999) realiza un estudio para averiguar cómo utilizan los traductores profesionales los diccionarios, aspecto en el que difieren de los estudiantes de traducción. En investigaciones anteriores (Fraser 1993, 1994) esta autora señala que la principal diferencia reside en cómo los estudiantes de traducción y los traductores profesionales se ocupan de las palabras que desconocen. Los estudiantes utilizan el diccionario, sobre todo bilingüe, mientras que los profesionales son más reacios a utilizar el diccionario y recurren a él con menos frecuencia. Por tanto, en este estudio se pretende observar el papel del diccionario bilingüe en el mundo de la traducción profesional.

Para ello, se estudia a un grupo de 21 traductores profesionales mientras traducen al inglés un artículo en francés sobre la reforma de la educación superior en Francia para su supuesta publicación en un suplemento sobre educación de un periódico anglosajón. Al mismo tiempo que los traductores trabajan en el texto, se les pide que comenten en voz alta el proceso (método TAP). Se estudia principalmente cómo los traductores se enfrentan al vocabulario o terminología del texto.

Los resultados muestran que, en general, los traductores son escépticos ante el uso del diccionario bilingüe (Fraser 1999: 26). A continuación incluimos algunos de los comentarios de los traductores.

'I find quite a lot of the dictionaries these days are [pause] I would say useless'

'It is in here! Amazing! I'm always amazed when something I want is actually in [the dictionary]!'

(Fraser 1999: 26)

Asimismo, los traductores transmiten insatisfacción por los resultados obtenidos con el diccionario bilingüe.

'[for *faire peau neuve*] "Turn over a new leaf" – that's complete cobblers, obviously'

'[for *filière*] It doesn't help very much'

'[for *premier cycle*] Not happy with the dictionary definition here'

(Fraser 1999: 26)

El estudio de Fraser también muestra que, cuando el diccionario bilingüe resulta inservible, los traductores profesionales utilizan otros recursos para acceder a la terminología que desconocen, como por ejemplo un experto, un tesoro, un diccionario monolingüe como el *Petit Robert* o el *Larousse*, o un texto paralelo, como en este experimento, una guía sobre educación superior de Francia.

'Administrative procedures, channels, no... Now, in the context, is it a department, is it a subject area, is it a section? Is it a facul-, no, it's not a faculty... I'm going to use "strand" for now... I'd need to talk to someone working in a university and I'd ask them, do you talk about strands? What's your vocabulary?'

'Faire peau neuve, "to adopt a new image, to change one's clothes, to turn over a new leaf". Right, er, now yes, um, les premiers cycles is talking about the stage in their education... Okay, I've got the feel... "are undergoing a metamorphosis" is a phrase that comes to mind... or "revised" but I don't particularly want to use "revised" again. [Consults computer thesaurus] Reviewed ... "are being reviewed", "are being changed"... "modify" is better, "transfigure, transform"... "In some universities, the first- and second-year courses are being transformed"'

(Fraser 1999: 27-28)

#### 4.2.9 Livbjerg y Mees (1999; 2003)

Livbjerg y Mees (1999) efectúan un experimento en el que tratan de averiguar cuál es el uso del diccionario por parte de estudiantes de traducción y qué influencia tiene sobre la calidad de la traducción. Intervienen 5 estudiantes de traducción avanzados recién graduados de la Copenhagen Business School que realizan una traducción inversa de 126 palabras del danés al inglés. En una primera fase los estudiantes traducen sin ayuda de ningún recurso de referencia ni presión temporal. Una vez finalizada esta fase, y después de un descanso, se pide a los estudiantes que revisen de nuevo su traducción y se les

proporciona un manual de uso, un diccionario bilingüe danés-inglés y un diccionario monolingüe de inglés.

La recogida de datos se basa en los TAP de los estudiantes a medida que traducen, la monitorización con el programa *Translog*, la observación directa y anotación de los investigadores (por ejemplo cada vez que un sujeto consulta un diccionario se anota el diccionario y la página consultada), así como las evaluaciones de las traducciones finales (con y sin diccionario).

Los resultados muestran que en el proceso sin diccionarios, en el 86,5% de los casos en los que el traductor se replantea la solución adoptada, no se producen cambios significativos: en el 21,6% un error se cambia por otro error y una solución correcta por otra solución correcta; en el 64,9% se mantiene el mismo error o solución correcta. Solo se corrigen ocho errores (10,8%) y se producen dos errores que previamente eran soluciones correctas.

El uso del diccionario tampoco mejora el resultado de las traducciones (en el 86,5%). En el 74,3% de los casos, se mantiene la solución adoptada sin diccionario, en el 12,2% una solución correcta o incorrecta lleva a otra solución correcta o incorrecta. Solo en siete ocasiones se corrige el error y en tres una solución correcta se convierte en incorrecta.

Respecto de las consultas en el diccionario, cada estudiante consulta entre 5 y 14 unidades en el diccionario y el número total de consultas por sujeto es entre 10 y 27, ya que algunas unidades se consultan en más de una ocasión. Los diccionarios se emplean principalmente para acceder a vocabulario desconocido, comprobar colocaciones (*open to, up to, Human Rights Commission*), descartar falsos amigos (*inspirere/inspire, debat/debate, op til/up to, resolution/resolution*) y problemas de comprensión.

Las autoras observan una falta de habilidades de los estudiantes para resolver problemas de comprensión y colocaciones con ayuda del contexto, lo que les lleva a utilizar el diccionario. Muestran un ejemplo de estrategia de búsqueda en el uso del diccionario para acceder a colocaciones desconocidas. En palabras de las propias autoras, «she felt that her mental lexicon contained a collocation that she could not raise to the level of consciousness» (Livbjerg y Mees 1999: 143).

En una primera fase del proceso de traducción, una estudiante traduce *debat skabt* por *debate created*, pero tiene la intuición de que existe una colocación mejor que no puede recordar. Con el objetivo de obtener inspiración sobre posibles colocaciones, consulta en el diccionario bilingüe danés-inglés el término *debat*, pero no encuentra soluciones satisfactorias. A continuación prueba a consultar *debate* en el diccionario

monolingüe inglés, con el mismo resultado. El siguiente intento es probar con el otro elemento de la colocación *skabe* que busca en el diccionario bilingüe, pero no resulta. Tiene la intuición de que el verbo *prompted* podría ser una opción para sustituir a *created* en la colocación, así que busca *prompt* en el diccionario monolingüe. Como esta búsqueda no proporciona los resultados deseados, decide mantener su primera opción *debate created*, pero lamenta no haber tenido acceso a diccionarios de colocaciones o de sinónimos que seguramente le hubieran ayudado a encontrar otra solución.

La conclusión aportada es que los estudiantes de traducción deben recibir formación sobre la consulta de las obras de referencia y también sobre su uso adecuado dentro del proceso de traducción para que sepan distinguir entre soluciones que pueden buscar en el diccionario y soluciones en las que tienen que utilizar otras estrategias comunicativas, como la paráfrasis o la omisión.

En Livbjerg y Mees (2003) los objetivos son los mismos. Participan 10 estudiantes de traducción avanzados, que realizan una traducción general inversa del danés al inglés de 126 palabras. El instrumento empleado para reunir información son principalmente los TAP y las autoras indican que también han realizado una monitorización con *Translog*. Se efectúan tres experimentos, 5 de los sujetos colaboran en los experimentos uno y tres y los 5 restantes en el dos. En el experimento uno los estudiantes no pueden utilizar el diccionario mientras traducen pero no tienen límite temporal para terminar la traducción, una vez finalizada y después de una pausa comienza el experimento tres en el que tienen que revisar el texto que han traducido con la posibilidad de utilizar dos diccionarios, uno bilingüe y uno monolingüe; en el experimento dos los estudiantes realizan la traducción desde el primer momento con posibilidad de utilizar los dos diccionarios.

En los resultados, los estudiantes que utilizaron el diccionario para realizar la traducción tardaron más tiempo que los que no lo usaron; consultaron el diccionario de manera considerable, en el experimento dos se empleó en el 53,3% de las unidades de traducción y en el experimento tres en el 42,2%, pese a que la hipótesis era que los estudiantes que habían realizado primero la traducción sin diccionario no sentirían la necesidad de utilizarlo después, una teoría que fue desmentida.

Respecto a la influencia del diccionario en la calidad de la traducción, estas autoras concluyen que no mejora, de 115 unidades consultadas, solo en 21 ocasiones el diccionario permite al traductor cambiar una solución que era errónea por una correcta y, de estas, 9 son mejoras de ortografía que podrían haberse solucionado con el corrector ortográfico, con lo que el total de soluciones satisfactorias proporcionadas por el diccionario es de 12.

A las autoras les sorprende que uno de los estudiantes alegue haber llegado a una solución gracias a la ayuda del diccionario pese a que la elección adoptada no se encontraba en el diccionario. De ahí, concluyen que los estudiantes hacen un uso excesivo del diccionario y que se centran demasiado en las unidades léxicas sin tener en cuenta la situación y el contexto.

Del total de búsquedas efectuadas por los estudiantes en los diccionarios, 59% fueron en el bilingüe y 41% en el monolingüe. En este estudio se incluye, además, a modo de ejemplo, una estrategia de búsqueda de un estudiante para el acceso a un equivalente en inglés de la palabra danesa *menneskerettighedskommissionen* (que planteaba problemas de ortografía, mayúscula o minúscula, elección de denominaciones *commission* o *committee*, gramática, singular o plural, y estructura de la frase *Human Rights Commission* o *Commission on/of/for Human Rights*). La estrategia de búsqueda descrita se componen de distintas acciones de búsqueda, la primera es la consulta de *kommission* en el diccionario bilingüe danés-inglés, pero no obtiene suficiente información para decantarse por *commission* o *committee*, así que consulta estas dos palabras en el diccionario monolingüe lo que le lleva a decantarse por *commission*.

#### 4.2.10 Corpas Pastor et al. (2001)

Este estudio aborda la consulta del diccionario por estudiantes de traducción. Se distribuyen dos cuestionarios. Los autores justifican la elección de la encuesta por ser el método más adecuado por su naturaleza empírica de obtención de datos, así como por su uso habitual en la universidad y añaden que existe una corriente investigadora bastante consolidada que utiliza esta misma metodología.

El primer cuestionario se utiliza como experiencia piloto. Los informantes son 212 estudiantes de la licenciatura en Traducción e Interpretación. A partir de los resultados obtenidos, se refina el análisis con una segunda encuesta para profundizar en los resultados de la primera, que se administra a un grupo de sujetos más restringido, 59 estudiantes de 2º curso de la licenciatura en Traducción e Interpretación. En la primera encuesta las preguntas están totalmente estandarizadas y solo una de las preguntas es parcialmente abierta, mientras que en la segunda encuesta la mayoría de las preguntas son abiertas y se encuentran limitadas solo en parte.

Los resultados de la primera encuesta se muestran por grupos de sujetos en función del curso en el que están matriculados. De 52 alumnos de primer año, el 48,1% indica que consulta el diccionario varias veces al día, el 17,3% una vez al día. El 94,2% señala que los

diccionarios ordenados alfabéticamente resultan más fáciles de consultar, y al 73,1% que tienen una presentación más atractiva, en comparación con el diccionario ordenado por campos conceptuales o por temas. También el 94,2% prefiere el diccionario bilingüe al monolingüe y al uso de la enciclopedia. El 53,8% afirma que la información más consultada son los equivalentes de traducción y el 34,6% las definiciones. Además, los encuestados demandan un mayor número de contextos en los diccionarios.

En cuanto a los alumnos de segundo curso, el 74% asegura que utiliza el diccionario varias veces al día, el 96% indica que los diccionarios ordenados alfabéticamente son más fáciles de consultar, y el 60% que tienen una presentación más atractiva. La primacía del diccionario bilingüe desciende levemente, lo consulta el 58%, frente al 26% que prefiere el diccionario monolingüe y el 10% la enciclopedia. La información más buscada también es la equivalencia (78%) y los fallos más nombrados son la información insuficiente (el 4% muy frecuentemente<sup>88</sup>, y el 34% frecuentemente) y la falta de contextos (64% acumulado entre frecuente y muy frecuentemente).

Los estudiantes de tercero utilizan el diccionario diariamente en un 72%, también prefieren el diccionario ordenado alfabéticamente (un 94% por la comodidad de consulta y un 58% por el aspecto formal). El 86% prefiere el diccionario bilingüe, frente al monolingüe (8%) y se señalan otros diccionarios, como el de sinónimos (4%), antes que la enciclopedia (2%). El 74% busca equivalentes. Los encuestados denuncian que las voces del diccionario y los sinónimos están desprovistos de contexto (28% muy frecuentemente y 24% frecuentemente), o con pocos contextos (38% muy frecuentemente y 30% frecuentemente).

En cuarto curso, desciende el número de encuestados que utiliza el diccionario varias veces al día (60%). El diccionario ordenado alfabéticamente continúa siendo el más utilizado por la comodidad de consulta (93,3%) y por la presentación (56,7%), aunque los porcentajes disminuyen muy levemente a medida que se avanzan los cursos. El diccionario bilingüe continúa siendo el más utilizado (73,3%), seguido del monolingüe (20%), y la enciclopedia (3,3%). Las equivalencias de traducción son la información más consultada (71,1%). Aparece un dato nuevo que llama la atención a los investigadores, el 40% busca definiciones muy frecuentemente, y el 30% frecuentemente. La ausencia de contextos sigue siendo la queja más recurrente (70%).

---

<sup>88</sup> Valores de la escala: muy frecuentemente, frecuentemente, a veces, rara vez, nunca. Solo se han contabilizado los valores correspondientes a muy frecuentemente y frecuentemente.



En la segunda encuesta, se pide a los alumnos de segundo curso que indiquen sus lenguas B (inglés o francés), así como su lengua materna y las lenguas C que están cursando. El 63% de los estudiantes de inglés afirma no haber recibido ninguna instrucción sobre el uso del diccionario, el 34,8% indica haber recibido formación en EGB (58,8%). El 38,5% de los estudiantes de francés señala no haber recibido formación, mientras que el 46,2% declara que sí se le ha instruido en el uso del diccionario por igual en todos los niveles de formación. Los estudiantes mencionan haber adquirido algún diccionario por razones académicas (64,8% los de inglés y 86,7% los de francés). El 57% de los estudiantes indica que tiene diccionarios de sinónimos, tanto en la lengua materna como en la lengua extranjera. Los diccionarios especializados más elegidos son los bilingües, frente a los monolingües. En cuanto al uso de los diccionarios, el 97% de los estudiantes de inglés y el 100% de los de francés señalan utilizar los diccionarios bilingües generales muy frecuentemente, seguidos por los monolingües generales, el 69,9% de los alumnos de inglés y el 100% de los de francés.

Como conclusión, los autores observan que los estudiantes de traducción tienen al menos un diccionario monolingüe y bilingüe general, y que tienen preferencia por el diccionario bilingüe frente al monolingüe, aunque en los cursos más elevados esta tendencia disminuye y, en ocasiones, se invierte.

Asimismo, estos autores constatan un interés mayor por el diccionario semasiológico (ordenado alfabéticamente) que por los diccionarios de carácter onomasiológico (en este estudio los autores entienden por diccionarios onomasiológicos a los ordenados por campos conceptuales o por temas), aunque afirman que los diccionarios onomasiológicos podrían resultar más ventajosos en algunas ocasiones. Como justificación de la elección de los estudiantes de los diccionarios semasiológicos, aluden su desconocimiento sobre la existencia de diccionarios onomasiológicos y de sus ventajas, como la ordenación de los términos dentro de una ciencia o campo, lo que favorece la ubicación, comprensión y utilización del término idóneo. Asimismo, los autores declaran que las editoriales no fomentan la publicación de diccionarios onomasiológicos porque no tienen suficiente aceptación.

La información más consultada en los diccionarios por los estudiantes de traducción son los equivalentes, aunque también se desprende de esta encuesta que la información microestructural que ayudaría a seleccionar el equivalente adecuado es deficiente. Entre los inconvenientes mencionados por los estudiantes, se incluyen la existencia de equivalentes obsoletos, la falta de equivalente de traducción para algunas acepciones de un lema, la

propuesta de equivalentes antinaturales o poco idiomáticos, o divergencias dentro del mismo diccionario. Se destaca que la elección del equivalente puede presentar dificultades debido a que los diccionarios presentan información incompleta sobre el uso real de los términos en LO y LM, no se incluyen contextos, construcciones verbales, restricciones de registro ni ejemplos. Los diccionarios también adolecen de falta de sinónimos, según indican los encuestados, razón por la cual poseen algún diccionario de sinónimos. Además se solicita mayor variedad de modos para acceder a la información en los diccionarios. Algunos estudiantes solicitan versiones electrónicas que faciliten la recuperación de los datos de manera sencilla y rápida.

La información menos consultada en los diccionarios por los alumnos es la etimología, las remisiones internas, las ilustraciones, los gráficos y las fotografías, excepto en los diccionarios especializados, en los que indican que una imagen puede facilitar la comprensión del significado.

Finalmente, se constata una falta de formación respecto al uso del diccionario. Los autores de este estudio abogan por fomentar la instrucción de los estudiantes de traducción en el uso de diccionarios y otras fuentes o recursos, como los corpus paralelos y comparables, las bases de datos, los foros de discusión, etc., de manera que aprendan a utilizarlos por separado o en combinación.

#### **4.2.11 Tarp (2002)**

Tarp (2002) trata de demostrar la utilidad de los diccionarios monolingües en combinación con los diccionarios bilingües en las distintas fases de la traducción directa e inversa. Su hipótesis es que los diccionarios bilingües no son la mejor solución, al menos por sí solos.

Este autor se basa en su experiencia como traductor jurado danés-español. En cuanto al uso de diccionarios, cuando traduce al español consulta más diccionarios monolingües en español que bilingües danés-español; aunque cuando traduce al danés consulta más diccionarios bilingües español-danés y monolingües de español que monolingües de danés, lo que cree que puede deberse a que el danés es su lengua materna por lo que le suscita menores problemas.

En la recepción de un texto de la lengua materna, el diccionario puede ayudar cuando se producen problemas de comprensión proporcionando definiciones, que incluyan sinónimos, hiperónimos, hipónimos, paráfrasis, ejemplos de uso en contexto, explicaciones más o menos largas. Otro problema derivado cuando se consulta el diccionario es la

identificación del lema adecuado que se buscará en el diccionario (ya que hay que tener en cuenta cuestiones de género, formas irregulares, variantes ortográficas).

En la transferencia de un texto de la lengua materna a una lengua extranjera, el diccionario no puede proporcionar equivalencias a nivel de texto, pero puede contener equivalentes de palabras, colocaciones, modismos, etc. En el caso de equivalencias parciales, el diccionario debería proporcionar información sobre las diferencias de significado entre las lenguas para ayudar al traductor a seleccionar el equivalente más adecuado. Según Tarp, en esta fase el traductor necesita confirmar que el lema equivalente es adecuado y que expresa el significado preciso, este problema puede solucionarse proporcionando una definición.

En la producción del texto en la lengua extranjera, el problema es producir un texto en la LM que sea correspondiente al texto de la LO. En este sentido, el diccionario proporciona soluciones de ortografía, género, conjugaciones irregulares, propiedades sintácticas, pragmática, colocaciones, modismos, sinónimos, antónimos, etc. En este caso para consultar el diccionario el traductor debe conocer el lema en la LM, de lo contrario se verá forzado a volver a la LO e iniciar la búsqueda en la LO.

En la recepción de un texto en lengua extranjera el traductor debería tener mayores problemas de comprensión, sobre todo en cuestiones semánticas. A diferencia de la recepción en la lengua materna, la recepción en una lengua extranjera puede solucionarse proporcionando equivalentes en la lengua materna junto con alguna explicación si fuera necesario. Las explicaciones más tradicionales solo serán necesarias en áreas de conocimiento desconocidas por el traductor o en fenómenos culturales que escapan a los conocimientos del traductor. La identificación del lema en lengua extranjera resulta más complicada porque el traductor deberá conocer las reglas en la lengua extranjera.

En la transferencia de una lengua extranjera a la lengua materna el problema es similar al de transferencia de la lengua materna a una lengua extranjera, pero difiere en la solución en el sentido de que los equivalentes en la lengua materna ya proporcionan por sí mismos los matices de significado.

En la producción del texto en la lengua materna los problemas son menores que en la producción en lengua extranjera, ya que el traductor posee un mayor dominio de su lengua materna. En el lenguaje más común, el traductor solo necesita que se le recuerden colocaciones, modismos, sinónimos, antónimos, y en el caso de vocabulario menos frecuente, precisará más información sobre género o conjugaciones irregulares. Sin embargo, como es difícil determinar qué vocabulario es más o menos frecuente, Tarp

recomienda incluir este tipo de información para todos los vocablos del diccionario. En este caso, lo que se consulta también es el lema en la LM con el fin de facilitar y acelerar el proceso de consulta.

La explicación de Tarp revela que las fases del proceso de traducción son similares aunque no idénticas en la traducción directa y en la traducción inversa. En lo que respecta al uso de diccionarios, este autor indica que aunque la fase recepción podría resolverse con la ayuda de un diccionario monolingüe, sería conveniente que se contara con uno bilingüe, ya que los problemas de recepción van a ir seguidos de problemas en la transferencia entre lenguas y en este sentido un solo diccionario bilingüe puede servir para las dos fases, además según este autor los equivalentes en LM normalmente son suficientes para comprender los términos de la LO, por lo que el diccionario monolingüe es opcional, pero no imprescindible en la fase de recepción del proceso de traducción.

En el proceso de producción en LM, según el autor, los diccionarios tanto monolingües como bilingües con lemas en la LM son la mejor solución, ya que para Tarp el traductor ya cuenta con una idea sobre cómo expresarse en la LM y solo necesita confirmar esta idea en cuestiones de ortografía, género, conjugación, colocaciones, sinónimos, etc. En este sentido los diccionarios bilingües de la LO a la LM son una pérdida de tiempo porque además pueden provocar que el traductor emplee un estilo más restringido y la calidad de la traducción se vea afectada.

Por tanto, para Tarp, el diccionario de traducción debería ser bidireccional o contener dos listas, una bilingüe y otra monolingüe: una lista bilingüe de la LO a la LM, y otra lista, tanto monolingüe como bilingüe, que contenga los términos en la LM, aunque reconoce que todavía queda mucho por investigar sobre cómo podrían implementarse estas ideas en un diccionario útil para el traductor.

#### **4.2.12 De la Rosa Martín (2003)**

De la Rosa Martín efectúa una encuesta similar a la de Corpas Pastor et al. (2001) a 9 estudiantes de tercer curso de la licenciatura en Traducción e Interpretación, en una asignatura de traducción especializada. Las preguntas de la encuesta van dirigidas a conocer el uso del diccionario.

Respecto a los diccionarios utilizados, el 88% señala el diccionario monolingüe en español, inglés y otras lenguas, el diccionario bilingüe y el diccionario bilingüe especializado; el 55% la enciclopedia, el 44% el diccionario monolingüe y bilingüe general y el 33% el diccionario monolingüe especializado.

En relación con la frecuencia con la que utilizan estos diccionarios, los diccionarios más utilizados son para el 100% el diccionario bilingüe, el 77% el diccionario bilingüe especializado, el 66% el diccionario monolingüe inglés, el 55% el diccionario monolingüe español y diccionario bilingüe general, el 22% el diccionario monolingüe general y el 11% la enciclopedia.

En cuanto al formato más útil para consultar el diccionario, el 44% selecciona el formato papel en un volumen, el 55% el formato CD-ROM. También se les pregunta qué tipo de diccionario les resulta más fácil de consultar: el ordenado alfabéticamente, o por conceptos o temas. El 100% señala el orden alfabético.

Respecto a la información consultada con mayor frecuencia, el 88% indica las definiciones y los ejemplos, el 66% los equivalentes de traducción, el 44% el campo, el 33% la ortografía, el 11% la fonética.

Sobre las fuentes que utilizan los traductores aparte de los diccionarios, el 66% indica Internet, el 33% los textos paralelos, el 22% los manuales sobre el tema y consulta a expertos y el 11% la prensa, las enciclopedias y los diccionarios de sinónimos.

En relación con los elementos que ayudan a elegir el término adecuado para una traducción, se les pide que distingan entre traducción general y especializada. En la traducción general, el 88% de los estudiantes indica los ejemplos, el 66% definiciones y equivalentes, el 11% la identificación del campo; en la traducción especializada el 88% señala definiciones y equivalentes, el 66% ejemplos, el 33% la identificación del campo.

Respecto a las mejoras en el diccionario, el 33% no contesta, el 11% señala diccionarios en CD-ROM más accesibles, actualizados, con más entradas y ampliaciones, dibujos, tablas, formato más claro y ejemplos en contexto, ya que algunos diccionarios ofrecen equivalentes sin ejemplos de uso. En los diccionarios generales se pide más información sobre registro y acepciones más concretas.

Sobre la importancia del diccionario como obra de consulta, el 33% la considera importante, el 11% imprescindible. En cuanto a si los estudiantes han recibido formación sobre el uso del diccionario, el 66% contesta que sí, el 11% que no y el 22% no responde a la pregunta.

Finalmente, en lo que respecta a las dificultades o problemas de uso del diccionario, el 66% señala que no incluye todos los significados de las palabras y que no siempre encuentra la palabra buscada, el 44% que no ofrece información de uso, el 22% que no proporciona ejemplos, el 11% que no define claramente los significados, las definiciones son ambiguas y no se especifica el registro especializado o general de las palabras.

#### 4.2.13 Sánchez Ramos (2004b)

Este estudio se realiza con 98 estudiantes de primer y segundo curso de la licenciatura en Traducción e Interpretación de la Universitat Jaume I de Castellón que cursan las asignaturas Traducción B-A1 (Español) I y Traducción B-A1 (Español) II (Sánchez Ramos 2004b: 478).

Se diseña un cuestionario a partir del propuesto por Hartmann (1999) con estudiantes de inglés como segunda lengua.

Los resultados indican que los diccionarios más utilizados por los estudiantes de traducción son los bilingües (87,8%), especialmente el *Oxford* y el *Collins*, seguidos de los monolingües en inglés (10,2%) y solo unos cuantos recurren al monolingüe en español (2%), preferentemente el *Diccionario de la Real Academia Española* y el *Diccionario de Uso de María Moliner*.

En cuanto a la frecuencia de uso del diccionario, los alumnos utilizan el diccionario bilingüe todos los días y solo recurren al monolingüe inglés y español una vez a la semana. Asimismo, se observa que el uso del diccionario bilingüe disminuye a medida que transcurren los cursos mientras que el del monolingüe aumenta y empiezan a emplearse otras estrategias, como la deducción por contextos.

La información consultada del diccionario bilingüe son las equivalencias (80,6%) y en el monolingüe las definiciones (74,5%) así como ortografía y uso (22,4%). El principal problema de los diccionarios es que los estudiantes no encuentran (31,6%) o no suelen encontrar (32,7%) las palabras que buscan, y las definiciones no son precisas (26,5%).

De aquí se extraen dos valoraciones: la falta de conocimientos por parte de los alumnos sobre el uso de los diccionarios o las deficiencias de las herramientas actuales para solventar todos los problemas terminológicos de un traductor.

Asimismo, se observa una falta de familiarización de los estudiantes con las herramientas electrónicas. Se mencionan algunos diccionarios en CD-ROM y en línea. Otras herramientas que utilizan en Internet son la prensa en español, corpus en español, libros electrónicos y otras herramientas documentales como foros, correo y portales especializados en traducción.

De los diccionarios en red valoran la accesibilidad, utilidad, organización de las entradas, distribución y rapidez en el acceso. Entre las desventajas, destacan la dificultad a la hora de buscar palabras, problemas con la conexión, la poca fiabilidad de la información, etc.

#### 4.2.14 Prieto Velasco y López Rodríguez (2009: 206-208)

Prieto Velasco y López Rodríguez (2009: 206-208) investigan la utilidad de la base de datos *EcoLexicon* en un grupo de usuarios potenciales, estudiantes de traducción científico-técnica de la Universidad de Granada. Efectúan un experimento en el que los estudiantes tienen que traducir dos textos sobre temática medioambiental en español e inglés con la ayuda de recursos terminológicos sobre medioambiente, además de la base de datos *EcoLexicon*. Después de traducir los textos, tienen que responder a un cuestionario en el que se les pregunta por sus impresiones acerca de *EcoLexicon* en comparación con los otros recursos.

Los resultados muestran que la valoración de los estudiantes es bastante positiva, destacan la relevancia de la información gráfica<sup>89</sup> frente a otros recursos, aunque también la lingüística, conceptual y contextual. Según los estudiantes, *EcoLexicon* resulta más útil que otros recursos como *IATE*, *UFORDAT* y *GEMET*, en los que solo se incluyen equivalentes, sin definiciones o imágenes que faciliten la comprensión de los conceptos.

Se comentan también aspectos mejorables como el exceso de información, ya que en ocasiones las redes de conceptos proporcionadas por el recurso son tan complejas que se dificulta la localización de información relevante. También se sugiere que se incluya un sistema de comunicación de los usuarios con los desarrolladores, a través del correo electrónico para informar sobre posibles errores. Otra recomendación es la traducción de la interfaz a inglés y alemán para que pueda utilizarse por un mayor número de usuarios. Además, se menciona la posibilidad de incluir filtros para los conceptos en función de los niveles de experiencia de los usuarios (de científicos a público general, o usuarios con perfiles diferentes: periodistas, estudiantes, traductores, etc.). Por último, se indica que se puedan recuperar textos de un corpus de medioambiente en el que se permita buscar un término o un concepto en su contexto junto con la imagen correspondiente. Los autores reclaman mayor investigación en este sentido.

---

<sup>89</sup> Este dato coincide con otro estudio con estudiantes de traducción de tercer y cuarto curso de la Universidad de Granada, en el que estos y otros autores (Tercedor et al. 2009) muestran que en la traducción técnica la visualización de imágenes facilita la comprensión de los conceptos y la generación de terminología adecuada en la LM.

#### **4.2.15 Jousse et al. (2011)**

Jousse et al. (2011) presentan un estudio piloto para probar la utilidad de las nuevas funciones incorporadas en el diccionario *DiCoInfo*: el acceso onomasiológico a las colocaciones. La idea es que el diccionario asista en las tareas de producción tanto en L1 como en L2 cuando el usuario conoce el significado de una colocación pero desconoce sus componentes (principalmente verbos con un significado especializado).

Los sujetos del experimento son 7 estudiantes de la Universidad de Aarhus, matriculados en un curso sobre traducción de páginas web corporativas danés-francés, y 3 estudiantes de traducción de la Universidad de Montreal.

Antes de llevar a cabo el experimento, los sujetos reciben formación sobre el uso del *DiCoInfo*. El experimento consiste en cuatro sesiones en las que se especifica una serie de búsquedas que los participantes deben realizar y responder a preguntas sobre el modo de acceso, el tiempo invertido en la búsqueda y la satisfacción (tanto en la obtención de resultados como en la facilidad de consulta).

En los resultados, el 69% considera las búsquedas en el *DiCoInfo* satisfactorias, el 16% muy satisfactorias y el 15% ligeramente insatisfactorias, por lo que los autores suponen que el grado de satisfacción es bastante alto, aunque todavía podría mejorarse.

Los comentarios positivos sobre el diccionario son la calidad de la información y la utilidad para las tareas de producción textual. Un aspecto negativo que se señala es la dificultad en la navegación entre los artículos del diccionario. Se sugiere que se utilicen distintos colores para destacar las secciones, una presentación más simple, mayor velocidad en la carga de los artículos y mejoras en la interfaz de búsqueda. Los autores señalan que uno de los futuros avances que se proponen es adaptar la interfaz de búsqueda a las nuevas funciones.



#### 4.2.16 Resumen de estrategias de búsqueda en diccionarios detectadas en los estudios

En los apartados anteriores hemos visto estudios sobre el **uso del diccionario por estudiantes de traducción y traductores profesionales**. En este sentido, hemos podido extraer unas primeras conclusiones, por un lado, sobre la metodología empleada en estos estudios y, por otro, sobre las diferentes necesidades y estrategias de búsqueda de los usuarios, que resumimos a continuación.

En cuanto a la **metodología**, el instrumento de medida que se emplea tradicionalmente en los estudios sobre el uso de diccionarios es el **cuestionario** cuya función principal es cuantificar qué información del diccionario consultan con más frecuencia los usuarios, así como su satisfacción con los diccionarios. Esta información se complementa en ocasiones con **entrevistas**. También se emplea la **observación directa** y **anotación** en plantillas, **TAP** y **monitorización**.

En este punto, nos gustaría señalar que Bergenholtz (2011: 32-33) critica la metodología mayoritaria del cuestionario cuantitativo, que se basa esencialmente en ofrecer porcentajes y generalizaciones. Este autor indica que en dichos estudios es el investigador el que selecciona un grupo de 20 a 100 estudiantes, que además han recibido instrucción del propio investigador, y que en cualquier caso no podrían considerarse representativos de la población general (inconveniente que también señala Tarp 2009c: 290). Asimismo, argumenta que en lo que respecta a los resultados estadísticos, la obtención de un porcentaje pequeño, por ejemplo del 10%, no significa que dicha necesidad no deba considerarse importante, sino que deberán tenerse en cuenta los resultados sobre las búsquedas de cada uno de los usuarios individuales, que proporcionen información útil para la construcción de los diccionarios. Tarp (2009c: 291) también reclama que muchos estudios se centran en averiguar las necesidades de los usuarios en función de la información que buscan en los diccionarios, por lo que no recogen aquellas necesidades que no satisfacen los diccionarios.

En relación con las **necesidades y estrategias de búsqueda**, observamos diferencias entre distintos tipos de usuario (estudiantes nativos, de segundas lenguas y traductores). En estudiantes nativos el uso mayoritario del diccionario es para la comprensión, la resolución de dudas de significado (Barnhart 1962; Greebaum et al. 1984; Quirk 1973), por tanto la información que más se consulta son las definiciones, aunque también la ortografía, pronunciación, sinónimos, notas de uso y etimología. También se observa un uso del

diccionario en tareas de codificación y expresión, en la escritura (Kipfer 1987). Los usuarios en general critican las definiciones de los diccionarios que parecen ser demasiado complejas y no resultar útiles (Greebaum et al. 1984; Miller y Gildea 1985; Quirk 1973). Estas conclusiones se apoyan en datos empíricos, como los de Miller y Gildea (1985) que demuestran que las oraciones construidas por los estudiantes con la ayuda de las definiciones del diccionario son más deficientes que con la consulta de ejemplos de uso del diccionario y contextos extraídos de un periódico.

Al igual que en los nativos, el principal motivo de consulta del diccionario en estudiantes de segundas lenguas es la comprensión. La información que más consultan estos usuarios en el diccionario son las definiciones.

Sin embargo, los estudios que simulan situaciones de traducción o que examinan a estudiantes de traducción y traductores profesionales muestran tanto una necesidad de comprensión como de expresión, con predominio de la expresión. También se advierten problemas de verificación en LM.

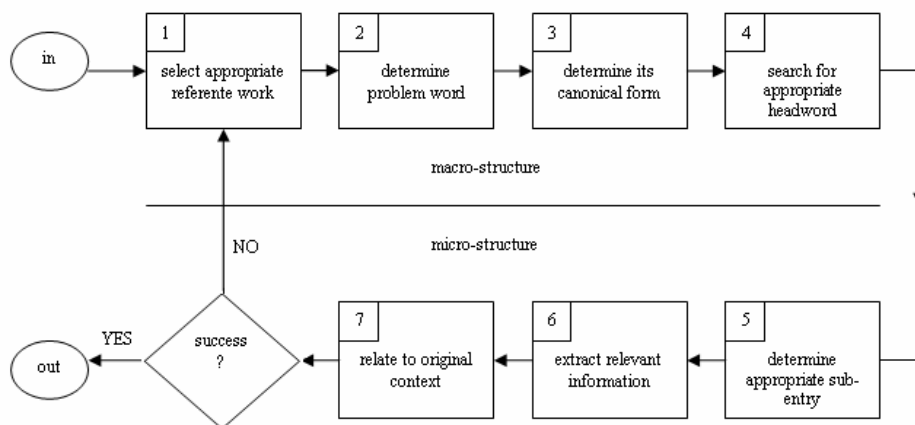
Los traductores necesitan acceder a equivalentes desconocidos en LM. Esta necesidad, según se ha indicado, surge sobre todo en la traducción especializada cuando el traductor conoce una idea de la LO y debe encontrar una denominación en la LM que desconoce (Bowker 2000: 22; Mackintosh 1998; Rogers y Ahmad 1998: 199). Asimismo, una vez localizado un equivalente en LM, el traductor necesita acceder a información de uso y colocaciones de dichos equivalentes en contexto (por ejemplo, Atkins y Varantola 1997; Varantola 1998; Varantola 2003b). Otros datos que precisa el traductor, según estos estudios, son ortografía, sinónimos, confirmación de intuiciones, etc. En los problemas de comprensión, la información que más se requiere son las definiciones. No obstante, las definiciones en los diccionarios para traductores no solo se utilizan para resolver problemas de comprensión de la LO, sino también son necesarias para confirmar equivalencias de la LM en contexto.

Las necesidades de expresión, de selección de equivalentes y colocaciones que expresan un determinado significado se consideran, como hemos visto, de naturaleza onomasiológica (Jousse et al. 2011; L'Homme y Leroyer 2009: 278; Leroyer 2006: 191; Verlinde et al. 2010: 6). Autores como Rogers y Ahmad (1998: 199) o Bowker (2000: 22) aluden a la dificultad de encontrar una denominación en LM cuando el diccionario bilingüe no proporciona una respuesta mediante la búsqueda a partir del término en LO. Para estos autores, los diseñadores de diccionarios electrónicos auténticos se encuentran ante el reto

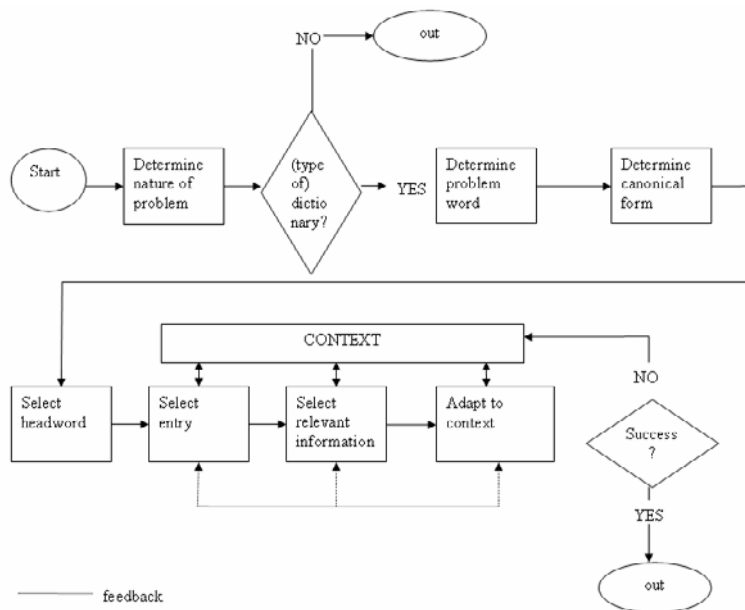
de cómo llegar desde la descripción de un concepto a encontrar la palabra o término en el diccionario.

Otros autores que hemos visto que sugieren situaciones que requieren búsquedas onomasiológicas en los diccionarios por parte de estudiantes de segundas lenguas son, por ejemplo, Baxter (1980: 327-329) y Wiegand (1998: 901-909). Sin embargo, en estos estudios las situaciones que se consideran de búsqueda onomasiológica pasan casi desapercibidas, la mayoría se limita a reflejar la necesidad de acceder a sinónimos y antónimos en los diccionarios de sinónimos y tesauros; Wiegand (1998) llega un poco más allá y sugiere otros tipos de relaciones como los hipónimos, los hiperónimos, los merónimos y los holónimos, y se refiere a otros tipos de diccionarios, aunque sin profundizar en ello, como diccionarios onomasiológicos, visuales, de relaciones semánticas, y también menciona algunas estrategias con las que un diccionario semasiológico sobre todo de tipo especializado podría servir para realizar una búsqueda onomasiológica de relaciones de meronimia, por ejemplo a partir de la búsqueda de partes de compuestos nominales en alemán.

A pesar de las necesidades de tipo onomasiológico sugeridas, muchos estudios se han centrado en las estrategias de búsqueda en el diccionario alfabético y semasiológico, no contemplan en estos casos la utilización de estrategias de búsqueda más complejas. A continuación presentamos dos modelos similares de uso del diccionario presentados por Hartmann (1989c: 105) y Bogaards (1993: 25; 2003: 29). Aquí los traductores tienen que buscar en el diccionario a partir de un determinado lema o entrada.



**Ilustración 96. Modelo de uso del diccionario según Hartmann (1989c: 105).**



**Ilustración 97. Modelo de uso del diccionario según Bogaards (1993: 25; 2003: 29).**

No obstante, los estudios también han mostrado, de alguna manera, un interés por las estrategias de búsqueda y los diccionarios onomasiológicos, aunque la investigación en este sentido se limita en muchas ocasiones a una pregunta sobre la utilización o no de estos diccionarios.

Por ejemplo, a pesar de que los modelos de uso del diccionario que acabamos de mencionar sugieren una búsqueda principalmente alfabética, Hartmann (1989c: 104), en el mismo estudio, indica que el diseño de un diccionario debe venir precedido por un conocimiento de las necesidades de los usuarios en distintas situaciones. Por ejemplo, sería interesante conocer si las búsquedas tradicionales semasiológicas con el fin de decodificar «(“Where does item X come from?”, “How is it used?”, “What is its basic meaning?”)» son más populares que las búsquedas onomasiológicas cuyo objetivo es codificar «(“Which field is X confined to?”, “How can I use it appropriately?”, “What other items may be (more) suitable?”)».

Asimismo, en el marco de la enseñanza de uso del diccionario, Hartmann (1989a: 185) propone enseñar, tanto a estudiantes de lenguas avanzadas, pero sobre todo traductores profesionales, a buscar equivalentes en los tesauros, e incluso diseñar nuevos tipos de *tesauros paralelos*. Por su parte, Béjoint (1989a: 210) explica que los usuarios deberán aprender a seleccionar el diccionario apropiado en cada momento (monolingüe o bilingüe, alfabético o conceptual, general o especializado, para estudiantes de lenguas o hablantes nativos, etc.), dependiendo de sus necesidades de información. Por ejemplo, deberán aprender a distinguir entre diccionarios bilingües para codificación y

decodificación, y deberán recibir formación de cada tipo de diccionario, por ejemplo en el caso de los diccionarios organizados conceptualmente. No obstante, reconoce que en algunas lenguas esta elección se encuentra muy limitada.

Corpas Pastor et al. (2001: 251) señalan que «[...] la búsqueda semasiológica se centra en las definiciones de los lemas, y no en el establecimiento de posibles correspondencias» pese a que reconocen que la búsqueda de equivalentes y su contextualización es la información que más buscan los traductores en los diccionarios. Además, estos autores indican que los estudiantes de traducción solicitan modos más diversos para acceder a la información en los diccionarios. Por un lado, los estudiantes señalan que los diccionarios más útiles o más cómodos de consultar son los alfabéticos pero, por otro lado, afirman que precisan una mayor variedad de formas de consulta, por lo que podríamos inferir que la búsqueda alfabética tampoco les resulta la más apropiada. Esto apoya la teoría de que los diccionarios no se encuentran diseñados de manera que faciliten a los usuarios el acceso a la información.

Por ejemplo, Corpas Pastor et al. (2001) preguntan a un grupo de estudiantes de traducción si prefieren el diccionario semasiológico o el onomasiológico y los resultados indican que los estudiantes de traducción prefieren los diccionarios alfabéticos, aunque los autores reconocen las ventajas de los diccionarios onomasiológicos en algunos casos.

Los autores de este estudio justifican la elección de los estudiantes de los diccionarios semasiológicos por desconocimiento de la existencia de diccionarios onomasiológicos y de sus ventajas, como la ordenación de los términos dentro de una ciencia o campo, lo que favorece la ubicación, comprensión y utilización del término idóneo. Asimismo, declaran que las editoriales no fomentan la publicación de diccionarios onomasiológicos porque no tienen suficiente aceptación, lo que según estos autores se convierte en un «verdadero círculo vicioso» (Corpas Pastor et al. 2001: 254).

Otro problema de la escasa proliferación de este tipo de diccionarios, no obstante, podría deberse a que todavía no se han desarrollado los mecanismos y las metodologías idóneas para elaborar diccionarios onomasiológicos que resulten rentables a los diseñadores en cuestión de tiempo, esfuerzo y coste y que resulten además útiles a los usuarios. De hecho, a pesar de que la lexicografía aboga desde hace tiempo por identificar las necesidades de los usuarios previa a la elaboración del diccionario, no se han planteado apenas estudios centrados en conocer de qué manera podría interesar a los traductores buscar dentro de un diccionario para hacer una búsqueda onomasiológica. Los diccionarios en papel no satisfacen del todo las necesidades de los traductores. Sin embargo, como

hemos visto, el desarrollo de diccionarios electrónicos fomenta la flexibilidad y adaptación a los usuarios, por lo que este tipo de estudios cobran sentido.

De hecho, un aspecto por el que apuestan la mayoría es el desarrollo de diccionarios electrónicos, como instrumento que abarcará no solo mayor información sino información útil (Atkins y Varantola 1997: 36; Corpas Pastor et al. 2001: 253; Mackintosh 1998: 124). Sin embargo, esto contrasta con que muchos estudios con traductores todavía utilizan diccionarios en papel en sus experimentos, e incluso algunos concluyen que los traductores apenas advierten diferencias entre el diccionario impreso y electrónico, como por ejemplo Palomares Perraut y Pinto (2000: 111) y De la Rosa Martín (2003: 295). Estos resultados podrían deberse al año de realización de los estudios. En estudios más recientes (Jousse et al. 2011; Prieto Velasco y López Rodríguez 2009) los usuarios valoran mucho que el diccionario se encuentre en formato electrónico y que incorpore funciones de búsqueda innovadoras.

Por un lado, se precisa la introducción sistemática y actualizada en el diccionario electrónico de equivalentes adecuados y actuales. Además, la información semántica es necesaria para la selección de los equivalentes, los sentidos deberán estar claramente diferenciados. También deberán incluir información sistemática sobre el uso en contexto, información sintáctica y colocaciones. En este sentido, se reclama el acceso a las colocaciones a partir de los significados en los diccionarios. También se requiere que las colocaciones puedan consultarse directamente en los contextos de un corpus a partir del diccionario.

Por eso, muchos coinciden en sus conclusiones en la necesidad del uso combinado e integrado de los diccionarios con otros recursos complementarios, que aporten información que no se incluye en los artículos de los diccionarios. Por ejemplo, Rogers y Ahmad (1998) y Varantola (1998) sugieren posibilitar el acceso a textos paralelos, corpus e Internet para consultar contextos. Se aboga por la elaboración de distintos recursos en uno. Los traductores muestran una tendencia a utilizar más de un recurso en sus búsquedas y a comprobar las soluciones en más de un recurso (Mackintosh 1998: 134). También se comenta la posibilidad del acceso a expertos, ya que según Rogers y Ahmad (1998: 199) el mejor diccionario es un experto bilingüe con acceso gratuito e ilimitado. También los traductores reclaman más y mejores diccionarios especializados manejables que grandes diccionarios generales (Tomaszczyk 1979). Algunos diccionarios permiten el acceso limitado por relaciones semánticas. Las imágenes son útiles para la comprensión de

conceptos en LO y generación de terminología en LM, para la comparación de conceptos entre términos en dos lenguas (Prieto Velasco y López Rodríguez 2009).

Los datos que más verifican los traductores son los equivalentes, sobre todo comparan las definiciones de términos supuestamente equivalentes para comprobar si coinciden los conceptos y, en concreto, en el caso de las expresiones técnicas o especializadas (por ejemplo, Varantola 1998). Por ello, se descubre un interés de los traductores por diccionarios de tipo híbrido y especializado, se trata de diccionarios que combinan tanto equivalentes como definiciones y que incorporan características del diccionario monolingüe y bilingüe en el mismo diccionario. Además, Bergenholtz y Nielsen (2006: 297) consideran, al igual que el estudio de Mackintosh (1998: 135), que las definiciones deben proporcionarse en la L1 para facilitar la comprensión de los usuarios de los conceptos extranjeros. Por ejemplo, en el caso del experimento de Mackintosh (1998: 132) estos diccionarios son los preferidos por parte sobre todo de los sujetos con más experiencia profesional. Fraser (1999) alega que mientras que los estudiantes de traducción prefieren una búsqueda exclusiva de equivalentes, los profesionales son escépticos respecto al diccionario bilingüe y precisan un acceso combinado a equivalentes y definiciones, tendencia que también se observa en estudiantes de niveles avanzados, que además emplean con frecuencia el diccionario de sinónimos (Corpas Pastor et al. 2001). Los profesionales en ocasiones utilizan la información que obtienen del diccionario bilingüe como pistas para realizar otras búsquedas, por ejemplo en tesauros, diccionarios monolingües, textos paralelos o expertos.

En otros estudios, que veremos a continuación, se investiga cómo otros recursos como los corpus e Internet pueden ayudar a los traductores a desarrollar estrategias de búsqueda que les permiten resolver problemas que los diccionarios tradicionales no son capaces de solventar.

### ***4.3 Estudios sobre estrategias de búsqueda de los traductores en corpus***

En la investigación se han identificado distintas aplicaciones de los corpus para usuarios diferentes, como por ejemplo las recogidas por Abaitua (2002: 77-81), Bowker y Pearson (2002: 11) o Laviosa (2003b), entre las que encontramos la elaboración de recursos lexicográficos y terminográficos (Budin y Mörth 2011; Burdach et al. 1990; Faber et al. 2001; Tercedor y López Rodríguez 2008), la edición plurilingüe, la internacionalización de productos, las búsquedas translingüísticas, lingüística forense,

estudios de literatura, estudios sociológicos, por ejemplo para comparar variedades lingüísticas diferentes, o la evolución histórica de una lengua; también son útiles en el desarrollo de gramáticas basadas en corpus, y también en el ámbito de la lingüística computacional, para el desarrollo sistemas de traducción automática basados en ejemplos, o en otras herramientas de procesamiento del lenguaje natural.

También en gran medida los corpus tienen cabida en la enseñanza y aprendizaje de segundas lenguas y de traducción<sup>90</sup>. Este tema se ha abordado en obras y congresos específicos como CULT (Corpus Use and Learning to Translate) y TALC (Teaching and Language Corpora).

Hay que tener en cuenta que las necesidades en el uso de los corpus pueden variar considerablemente dependiendo de los usuarios, por ejemplo las necesidades de corpus por parte de lexicógrafos o traductores (Varantola 2006: 217).

En los últimos años, las investigaciones en materia de corpus han calado con fuerza en los estudios de traducción (Laviosa 1998), en ellas se han abordado distintas perspectivas sobre las aportaciones de los corpus a la traducción. Algunos autores los denominan *corpus orientados a la traducción* o *translation-driven corpora* (Zanettin 2000: 103).

Desde el punto de vista del tipo de corpus que se utiliza en los experimentos han proliferado los estudios teóricos cuantitativos con corpus comparables monolingües. Este tipo de estudios parten de la teoría de los universales de traducción, introducida por Baker (1993), que han tratado de comprobarse por medio de distintos estudios en los que se comparan textos originales en una lengua y textos traducidos en la misma lengua para constatar sus diferencias. A modo de ejemplo, encontramos estudios como los de Zanettin (2000), Zanettin (1999) o Laviosa (1998: 561). Estas investigaciones cuantitativas que se basan en corpus se consideran de inestimable valor para los estudios de traducción, por lo que requerían una mención en este apartado. Sin embargo, no dan cuenta de las estrategias de búsqueda que utiliza un traductor para consultar un corpus, por lo que no formarán parte del objeto de nuestro estudio.

---

<sup>90</sup> Algunos autores proponen una metodología colaborativa de compilación de corpus en el aula de traducción, que prepare a los estudiantes en habilidades tan demandadas en el mundo profesional como son el trabajo en equipo, la gestión de proyectos, etc. (Bowker 2002: 29; Yuste 2002: 35).



En este apartado veremos estudios cualitativos sobre estrategias de búsqueda de los traductores en los corpus. Cada vez son más los autores que tratan de demostrar la utilidad de los corpus en la mejora del resultado de las traducciones y proponen una serie de estrategias de búsqueda que pueden emplearse en estos recursos para el acceso al vocabulario y la terminología que se precisa en una traducción.

En este sentido, los corpus pueden ser más eficaces que los recursos tradicionales, como los diccionarios, que hemos tratado anteriormente. De hecho, Pearson (1996a: 88) suscribe el principio firthiano de que *you shall know a word by the company it keeps*, en el sentido de que el análisis de suficientes ejemplos de una palabra en su uso real pueden ayudar al traductor a entender su significado, localizar palabras relacionadas e identificar colocaciones adecuadas.

Tal y como se afirma por ejemplo en Corpas Pastor (2001b: 98; 2008: 56-77), la prosodia semántica es imprescindible para establecer equivalentes de traducción.

Sin ir más lejos, Sinclair (2000 [1996]) rechaza la palabra como unidad semántica. En su lugar, propone unidades semánticas ampliadas (*extended units of meaning*) que trasciende a la unidad léxica tradicional en sí, pues incluyen sus relaciones de colocación (*collocation*), de coligación (*colligation*), de preferencia semántica (*semantic preference*) y de prosodia semántica (*semantic prosody*). Dichas relaciones, sobre las cuales volveremos más adelante presentan conexiones también entre sí y forman parte del significado de una palabra dada.

(Corpas Pastor 2008: 57)

La recopilación de los distintos usos o estrategias que se han descrito hasta ahora en los corpus podría servir, por un lado, para conocer más en profundidad cómo utilizar los corpus, ya que los corpus son buenos en la medida en que lo son sus usuarios; en palabras de Bernardini et al. (2003: 11) «[...] a corpus is only as good as its use(r)s [...]». En nuestro caso los corpus serán útiles para los traductores en la medida en que los traductores sepan utilizarlos. Pero también, por otro lado, la recopilación de estrategias de búsqueda empleadas por los traductores en los corpus podría llevarnos a extraer conclusiones sobre cómo habría que mejorar los corpus para facilitar las estrategias de búsqueda en corpus, de manera que se satisfagan las necesidades de búsqueda de los traductores.

La concepción de Bernardini et al. (2003: 10) sobre estrategia de búsqueda en corpus trata la *flexibilidad* de los usuarios, concebida como el desarrollo de *estrategias* para el uso de corpus, para acceder a la información que hay en los corpus. Esta idea de flexibilidad o estrategia de búsqueda en corpus, que tradicionalmente se ha desarrollado más en profundidad en otros campos como en el de aprendizaje de lenguas extranjeras (Bernardini 2000a: 109-111, 2000b; Laviosa 2002), bien podría servir para la práctica de la traducción, en concreto en lo que se refiere a las búsquedas terminológicas en traducción. Aquí, Aston

(2009: IX) señala que los traductores deberán *interpretar y evaluar* los resultados que les ofrecen los corpus para llegar a una *conclusión*, deberán desarrollar una *sensibilidad* sobre cómo pueden utilizar los corpus para resolver los problemas de traducción.

Algunos de los estudios que analizamos a continuación abordan la búsqueda en distintos tipos de corpus en el mismo experimento y también los combinan con el uso de Internet, como veremos en el apartado 4.3.1. Otros autores se centran en las estrategias de búsqueda de estudiantes de traducción y traductores profesionales en corpus bilingües comparables y paralelos (apartado 4.3.2), así como en corpus monolingües en LM (apartado 4.3.3). En cada uno de los subapartados presentamos los estudios examinados por orden cronológico y, a su vez, por orden alfabético según los autores.

### **4.3.1 Estudios cualitativos sobre estrategias de búsqueda en corpus de distinto tipo**

Los estudios que presentamos en este apartado muestran ejemplos de estrategias de búsqueda de los traductores en corpus de distinto tipo: corpus monolingüe en LO y en LM, corpus bilingüe paralelo y comparable, e Internet como corpus.

#### **4.3.1.1 Aston (1999; 2000)**

Aston (1999), desde su experiencia, presenta los siguientes usos y búsquedas que un estudiante de traducción puede dar a distintos tipos de corpus. Este autor señala que los corpus ofrecen ventajas respecto a otros recursos tradicionales como diccionarios y glosarios: incluyen información que no se encuentra en los diccionarios y glosarios como colocaciones, frecuencia y uso típico, un mayor número de ejemplos y contextos más amplios, lo que facilita una mejor comprensión de determinados usos lingüísticos y del funcionamiento de los textos (Aston 2000).

El corpus monolingüe general en LO puede utilizarse como complemento de diccionarios para comprender el significado de una palabra del texto origen, por ejemplo, para saber si *overlooking* se emplea en un texto para expresar que un inmueble tiene buenas vistas, irónicamente para dar a entender que tiene malas vistas o de forma neutra y con ambigüedad. Esta es una información que no proporciona el diccionario y puede comprobarse en un corpus, por ejemplo mediante la consulta de *overlooking* en el contexto de *apartment* en el *BNC*. Normalmente *overlooking* tiene un matiz positivo en las concordancias recuperadas, ya que las vistas son a montañas, ríos, océanos, puertos, plazas

o jardines. Solo en algunos casos tiene connotaciones negativas que se presentan con ironía.

Un corpus monolingüe en LM puede emplearse para comprobar la validez de un equivalente en LM. Por ejemplo, al traductor puede interesarle verificar si *dava su* en italiano como equivalente de *overlooking* tiene las mismas connotaciones positivas y comparte también la posibilidad de emplearse de forma irónica. Para ello, se consulta *dava su* en el contexto de *appartamento, camera, casa y finestra*, el resultado es que aparece en contextos con *giardino, mare, montagna, piazza*, etc. En esta búsqueda se detecta de forma casual otra expresión, *con vista su*, que puede comprobarse también en el corpus.

Asimismo, se señalan inconvenientes del corpus monolingüe, como la dificultad de encontrar un corpus suficientemente representativo en cualquier lengua, el carácter polisémico de los términos que reduce la precisión de las búsquedas, ya que los contextos relevantes aparecen junto a los no relevantes, las dificultades de los estudiantes de traducción de interpretar y extrapolar los resultados de las búsquedas a soluciones para la traducción.

Un corpus bilingüe comparable especializado puede ser útil especialmente para reducir el número de resultados no relevantes ya que la búsqueda se limita a un determinado campo. En este sentido el corpus puede ayudar a formular hipótesis de traducción, como en el siguiente ejemplo. En un texto original en italiano aparece *casi mortali per insufficienza epatica*. En un corpus comparable que contiene textos en inglés sobre hepatitis C se pueden identificar los textos en los que aparece *death*. Al leer el primero de los textos recuperados, se identifica la expresión *fatal liver disease*, que puede buscarse ahora en todo el corpus para comprobar su uso.

También se alude al acceso a la información por casualidad (concepto al que este y otros autores denominan *serendipity*). Por ejemplo, la búsqueda de concordancias de *panel painting* en el corpus de la National Gallery incluye entre los resultados tipos de composiciones pictóricas como *tondi*, y técnicas como *pastiglia*, que son términos que pueden resultar útiles en una traducción sobre historia del arte.

Los corpus comparables pueden ser corpus *ad hoc* diseñados para un determinado encargo de traducción. Aston (2000) indica que estos corpus *ad hoc* compilados por el traductor desde Internet pueden ser útiles para evaluar expresiones del texto origen o hipótesis en el texto meta. Por ejemplo, en el caso de la traducción del francés al inglés de un artículo de prensa donde aparece la palabra *clochemerlesques* en el texto origen (ejemplo extraído de Bertaccini y Aston 2001). La consulta de *clochemerl\** en un corpus

de artículos del periódico *Le Monde*, y el buscador *Altavista* da acceso a 20 textos en los que se refleja su significado, adquirido de la novela satírica *Clochemerle* de Chevallier y que se emplea en los contextos periodísticos franceses para aludir al faccionalismo lastimero sin necesidad de una explicación sobre sus orígenes. La búsqueda de *clochemerle*\* en Internet y en un corpus de textos de los periódicos *The Independent* y *The Daily Telegraph* muestra que este término no se utiliza en los textos periodísticos británicos. Del análisis del corpus francés se extraen una serie de adjetivos que aparecen en el contexto de *Clochemerle* como *polémique, querelle, discorde, guérilla, rivalité, conflit, rebelle*. La búsqueda de los equivalentes de estos términos en el corpus inglés *quarrel, squabble, row, bicker, dispute, rebel, battle, rivalry* sugiere que *squabble* expresa el matiz que se está buscando, por lo que la solución que se adopta en el texto meta es *local squabbling*.

El corpus paralelo de textos originales con sus traducciones a otra lengua puede servir para encontrar rápidamente equivalentes de traducción (Aston 1999). El problema es el difícil acceso a corpus de este tipo. Por ejemplo, un corpus paralelo inglés-italiano se emplea para encontrar equivalentes de *establish* como *avviare, creare, elaborare, instaurare, realizzare, verificare*. También puede ser útil para encontrar el equivalente de *relazione* que es *report*, de *favore* que es *favour, votare a favore di o vote for* y *accogliamo con favore* o *we welcome*. El equivalente encontrado en un corpus paralelo puede emplearse como hipótesis de traducción que se contrasta en un corpus monolingüe general o especializado en LM.

Este autor advierte de los riesgos de utilizar un corpus paralelo como único recurso de búsqueda en traducción: se pueden calcar estructuras del texto original o reiterar posibles errores de traducción que existan en el corpus paralelo, por lo que se recomienda su uso combinado con corpus monolingües y comparables. En este sentido, se defiende la creación de corpus paralelos bidireccionales o recíprocos que, como veíamos en el anterior capítulo, contienen corpus paralelos y comparables.

Finalmente, Aston (2000) señala que las memorias de traducción, corpus paralelos analizados con programas de traducción asistida, pueden ser útiles para acceder a expresiones equivalentes que ya se han traducido con anterioridad.

#### 4.3.1.2 Varantola (2000; 2002a; 2003a; 2006)

Varantola (2000) investiga las necesidades de los estudiantes de traducción de la Universidad de Tampere cuando traducen de forma inversa un texto de lengua materna finlandés a lengua extranjera inglés.

A partir de estos estudios, recopila una serie de búsquedas lingüísticas en corpus que esta misma autora considera preliminares y que requerirían una investigación más profunda (Varantola 2002a).

El primer tipo de búsqueda propuesto es la confirmación o verificación de equivalentes. Por ejemplo, el traductor quiere comprobar si los equivalentes que ha encontrado en el diccionario *hatch* y *fry* pueden emplearse en un contexto sobre peces, ya que el diccionario no incluye ejemplos. Este tipo de confirmación puede realizarse en un corpus sobre peces.

Otro ejemplo consiste en que el traductor quiere encontrar un antónimo de *fine-meshed net*. El traductor piensa en algunas opciones como *large-meshed*, *bigger-meshed*, *coarse-meshed*, *loose meshed*, pero ninguna la encuentra en el diccionario monolingüe. Una búsqueda en el *BNC* recupera algunos ejemplos de *large-meshed* y una búsqueda directamente en Internet con la ayuda del buscador *Google* confirma que *large-meshed* es un equivalente adecuado y de uso actual. El problema que señala la autora con este tipo de búsqueda es que el traductor debe contar con una intuición previa para poder buscarla y confirmarla. En algunos casos más complejos, el traductor puede no contar con una intuición inicial y la búsqueda en corpus puede ayudar a encontrar una primera intuición en LM.

En los diccionarios electrónicos actualmente la búsqueda en todo el contenido del diccionario también proporciona resultados a consultas que antes eran difíciles de resolver, por ejemplo en la búsqueda de una denominación en LM inglés de «reel seat belt» (traducción literal del finlandés). Según la autora, este era un nuevo concepto que no se había incluido en los diccionarios bilingües finlandés-inglés y la búsqueda de *reel* no generaba resultados relevantes en los diccionarios monolingües de inglés. Actualmente en *Google* la búsqueda de un término compuesto parcialmente correcto recupera contextos de *inertia reel seat belt*, por lo que en ocasiones no será necesario ni siquiera compilar un corpus, ya que una búsqueda en Internet también puede proporcionar una confirmación. En otro ejemplo, el traductor quiere confirmar si *basic medicine* puede emplearse para designar un concepto en finlandés, sin embargo el traductor encuentra en el corpus otra expresión más adecuada *first-line*, *first-line antihypertensive therapy*.

En ocasiones se precisa encontrar formas de parafrasear en la LM una unidad léxica de la LO. Por ejemplo, el traductor puede saber que la mejor forma de traducir una idea del finlandés al inglés es reformular dicha idea, lo que puede resultar complicado si se traduce hacia una lengua extranjera. Supongamos que el traductor tiene que expresar en inglés la idea del texto finlandés de «weak age groups» de determinados peces en un lago grande. Probablemente el traductor sepa que la expresión *weak age groups* no es la más adecuada, pero no encuentra otra solución. Buscando en un corpus comparable podría haber encontrado la expresión *reduced populations*.

Esta autora también se refiere a la búsqueda por aproximación (fuzzy searches) y encuentros por casualidad (serendipitous finds). Según indica, la solución de *reduced populations* puede haberse encontrado de forma casual al acceder a determinados contextos de un corpus comparable.

Otro caso en el que los corpus pueden resultar útiles según Varantola es en la corrección de textos. Por ejemplo en un estudio pidió a dos grupos de estudiantes que mejoraran la calidad de dos folletos de instrucciones sobre la hipertensión y el SIDA publicados en inglés para la población finlandesa común que no habla finlandés. Con la consulta de corpus se realizaron las siguientes correcciones en el texto sobre hipertensión: *patient* en lugar de *subject*, *irregular heart rhythm* en lugar de *arrhythmia*, *over-the-counter medicine* en lugar de *non-prescription drug*, *elderly* en lugar de *older people*, *drugless treatment* en lugar de *nondrug treatment*.

En el texto sobre SIDA se decidió que los términos *asymptomatic* y *seroconversion* no podían emplearse sin ser definidos ya que no eran lo suficientemente generales; *lymph nodes* sí que tenía un uso generalizado en el corpus, aunque también aparecía (*swollen glands* por lo que los estudiantes decidieron utilizar ambos términos en el texto dejando patente que eran sinónimos; asimismo *mucose* y *mucose membrane* resultaron términos demasiado técnicos, mientras que una búsqueda en el corpus de *transmit* permitió encontrar términos más generales como *lining* y *tissue*. El corpus aporta en este sentido convenciones terminológicas, vocabulario complementario adicional, formas idiomáticas de expresar ideas.

Varantola (2003a) efectúa un estudio basándose en principios similares a los del experimento anterior, pero en esta ocasión se trata de realizar traducciones en grupo de encargos *reales*, en el sentido de que hay que entregar una traducción final al cliente y los traductores cobran por ello. Se preparan cinco textos para traducir, cuatro de finlandés a inglés y uno de inglés a finlandés. En una primera fase, los estudiantes tienen que decidir

cuál va a ser el público meta de su traducción, posteriormente el tipo de problemas que esperan encontrar en la traducción para después compilar un corpus que les ayude a resolverlos. Los problemas de traducción detectados son de tipo terminológico: terminología legal de una cultura específica, terminología especializada, terminología médica menos técnica, terminología de la UE, terminología de otras áreas; información estilística, por ejemplo las convenciones estilísticas de textos instructivos o de *marketing*, estilos informativos más generales; problemas con el texto origen, como por ejemplo que esté escrito en inglés con una fuerte influencia del francés. También se identifican problemas a la hora de compilar el corpus, como el exceso de información en línea sobre un mismo tema.

En general, respecto a las búsquedas las concordancias se emplean para obtener información colocacional y contextual, así como para encontrar porciones de texto más largas que ayuden en la construcción textual y la cohesión; las listas de palabras y las listas por frecuencia resultan útiles para encontrar palabras clave de búsqueda; la información sobre la longitud oracional en el corpus comparándola con la del texto meta indica si el texto meta es adecuado estilísticamente; finalmente se emplean las listas de palabras clave para la selección de términos más adecuados y para cuestiones estilísticas.

En cuanto a la evaluación del programa de análisis de corpus utilizado, los estudiantes critican algunos inconvenientes de *WordSmith Tools*, como que no es sencillo de utilizar y el manual de instrucciones no resulta de mucha ayuda, que en ocasiones requiere tener abiertas muchas ventanas al mismo tiempo y que no es capaz de realizar muchas tareas simultáneamente, sin embargo, también comentan que los corpus les han resultado útiles para tener más confianza en las decisiones adoptadas en la traducción y para encontrar términos y expresiones léxicas que no habrían podido encontrar en diccionarios y glosarios.

Finalmente, Varantola (2006) incluye un ejemplo de búsqueda en la que el traductor necesita comprobar la existencia y su aceptación en contexto del término compuesto *iron-bar lever*. Una búsqueda en el diccionario no proporciona resultados. En *Google* se comprueba que el término existe y que se emplea en el contexto que interesa al traductor.

#### 4.3.1.3 Kübler (2003)

Kübler (2003) efectúa un estudio cuyo objetivo es ayudar a los estudiantes de traducción a que se familiaricen con los corpus electrónicos y los programas de gestión de corpus, así como con el uso de Internet como corpus, para la búsqueda de información lingüística y enciclopédica.

Para ello, compila tres corpus, de los cuales dos son corpus especializados. Para la consulta de los corpus se utiliza un programa que permite buscar con expresiones regulares y etiquetas de categoría gramatical.

Los sujetos del experimento son dos tipos de estudiantes, estudiantes de traducción de cuarto curso y estudiantes graduados, que tienen que traducir en grupo un texto extraído de una página web y desarrollar tareas de distinto tipo (compilación de corpus, creación de una base de datos terminológica, traducción, revisión y preparación de la página web para su publicación). La autora no especifica cómo analiza los resultados, pero proporciona ejemplos de estrategias de búsqueda en los corpus para comprender términos, acceder a información lingüística y equivalentes desconocidos en la LM.

En la comprensión de términos, pone el ejemplo de que el traductor desconoce el significado de *white balance* y que no lo encuentra en diccionarios y glosarios. La búsqueda de *called {0,30} white balance (called a una distancia de 0 a 30 caracteres de white balance)* en un corpus monolingüe recupera concordancias de definiciones como la siguiente: «It's called 'White Balance' (WB, for short), which simply means the camera tries to keep white objects fairly white, so they can take on extreme color casts» (Kübler 2003: 35).

Para la obtención de información lingüística, siguiendo el mismo ejemplo de *white balance*, con la consulta de *(white balance)/(WB)/(whites)* el traductor puede acceder a concordancias de *white balance*, *WB* y *whites*. Ordenar las concordancias por la primera posición a la derecha de la palabra pivote, puede llevarnos a unidades compuestas como *white balance setting*, *white balance mode*, *white balance control*, *white balance compensation*, *white balance system*, *white balance feature*, *white balance department*, *white balance thing*. Ordenando las concordancias por la primera posición a la izquierda de la palabra pivote pueden encontrarse colocaciones como *accurate*, *daylight*, *cloudy*, *manual*, *automatic* o *auto white balance*. Las unidades *white balance department* y *white balance thing* por la intuición del traductor no parecen términos, lo que se comprueba en un corpus de lenguaje general.



En la búsqueda de un equivalente en francés para *white balance*, si el traductor busca una hipótesis de equivalente de *balance*, en francés *équilibre*, en un corpus francés sobre cámaras digitales, no se encuentran resultados. Sin embargo, si se busca la hipótesis de traducción del otro componente del término compuesto *blanc* (por *white*) encontramos *balance des blancs*. También puede accederse a información de términos compuestos y colocaciones, por ejemplo, las colocaciones adjetivales de *exposition* (*exposure* en inglés) son *forte* y *faible*.

Otros ejemplos de búsqueda en corpus paralelos y comparables que proporciona la autora son la búsqueda de un equivalente en francés del término inglés *firewall*. Si se busca el término original en un corpus paralelo, podemos ver una definición de *firewall* y acceder al párrafo equivalente en francés donde se observa que también se utiliza *firewall*. Una búsqueda posterior en la parte francesa de un corpus comparable nos confirma que *firewall* se utiliza en francés, sobre todo en contextos técnicos entre especialistas, pero que en los manuales de usuario dirigidos al público general se emplean con más frecuencia *garde-barrière* y *coupe-feu*.

En cuanto a estructuras verbales y sus argumentos, la autora pone un ejemplo en el que se busca un equivalente en francés de *boot* en el campo de la informática (*Your BIOS may not allow you to boot to a Linux installed there*). La French Commission for Computer Science Terminology trató de imponer el término *amorcer*. Sin embargo, la búsqueda en los corpus y en Internet proporciona otros equivalentes. Según la autora, estos equivalentes pueden verificarse en un corpus paralelo buscando al revés a partir de los equivalentes en LM para comprobar después si estos equivalentes se emplean como traducción del término original *boot*. Por ejemplo, se busca *lancer*, *démarrer* y *bouter* en francés y se comprueba si se emplean como equivalentes de *boot*. En el caso de *lancer* el resultado es negativo, aparece como equivalente de *launch* o *run, issue, type (a command)*.

Por último, se proporcionan ejemplos de búsqueda en Internet para comprobar o verificar intuiciones cuando el corpus no ofrece resultados. Por ejemplo, en el caso de *focus brackets*, *brackets* no equivale a *parenthèses* ya que no aparece en el corpus. Se ha encontrado un equivalente para *focus* que es *mise au point*, pero en el corpus tampoco se encuentra un equivalente para el término compuesto. La búsqueda de *mise au point* en Internet recupera contextos en los que aparece *repères de mise au point*. En el caso de la búsqueda de un equivalente de *Trojan horse*, la búsqueda de *cheval de Troie* en un buscador de Internet recupera contextos de dicho término en páginas de informática que hablan de virus.

#### 4.3.1.4 Moraes (2006)

Moraes (2006) trata la utilidad para la traducción de los corpus monolingües (analizados con *WordSmith Tools*), comparables, paralelos (compilados por el propio traductor o ya disponibles como *COMPARA*), memorias de traducción (analizadas con programas de traducción asistida como *Trados* y *Déjà Vu*) e Internet como corpus con la consulta en *Google* y otras herramientas como *WebCONC*. Aporta una serie de ejemplos de búsqueda basados en su propia experiencia como traductora.

En el primer ejemplo, una compañera traductora necesita comprender el significado de una expresión *figured to* que aparece en el texto original en la frase *the team was figured to win the game*, en un artículo sobre un equipo deportivo en inglés. La búsqueda en *Google* de esta expresión no facilita su comprensión en el contexto indicado. Sin embargo, según la autora una búsqueda en *WebCONC* sí que recupera contextos en el campo del deporte en los que *figured to* se emplea con el sentido de *were expected to*, como en los siguientes casos (Moraes 2006): «We believe that if we stay healthy and win a game we're not figured to win, we could possibly go 5-5», «Pennsylvania's unbeaten football team figured to beat Harvard on Saturday, and both offenses figured to score often», «Thanks. That there's just what we figured to do».

En otro ejemplo, un traductor le pide ayuda sobre cómo utilizar *over* con el significado *with respect to*. Una búsqueda en el corpus *BNC* genera 50 concordancias, demasiadas para analizar por un solo traductor. Sin embargo en la primera y la quinta concordancia aparecen buenos ejemplos de uso de *over* en el sentido mencionado.

Otro ejemplo de búsqueda se da en un corpus monolingüe con el programa *WordSmith Tools* en la búsqueda de ciertas expresiones para comprender el significado. También menciona la estrategia de «adivinar y comprobar» (*guess and check*) que consiste en que en expresiones compuestas por varios términos, el traductor consulta en el corpus uno de los términos que componen el término compuesto, cuya traducción conoce seguro, y accede a las colocaciones de dicho término para tratar de identificar el equivalente del término compuesto completo.

Otra utilidad del corpus monolingüe es la de verificar la ortografía de las palabras, por ejemplo de aquellas palabras que terminan en *\*ic* o *\*ical* (como en el caso de *electric* y *electrical*). Se habla de *electric train*, *electric switch*, *electrical wiring*, *electrical installations*.

Otro ejemplo es la búsqueda de un equivalente en inglés del término portugués *prazo*. La consulta de *prazo* en un corpus paralelo recupera concordancias paralelas entre las que se encuentran *term, deadline, period, period of time* y un cambio en la estructura de la frase con *within*. La traductora señala que una búsqueda en un corpus comparable de los equivalentes propuestos por el corpus paralelo puede ser útil para conocer las estructuras más naturales en la LM.

#### 4.3.1.5 Roberts y Bossé-Andrieu (2006)

Roberts y Bossé-Andrieu (2006) tratan de mostrar la utilidad de los corpus en traducción, tanto corpus compilados con un determinado propósito (que denominan corpus lingüísticos) como el uso de Internet como corpus (que llaman corpus general). Para ello, presentan un experimento en el que se valen de su propia experiencia en la que simulan cómo un traductor freelance puede utilizar los corpus para la traducción de un texto del inglés al francés.

Toman un texto original en inglés sobre el *missile defence program*, y señalan una serie de problemas que plantea su traducción y que resuelven con ayuda de distintos recursos: el corpus paralelo *Hansard*, un corpus compilado con textos del *Wall Street Journal*, un corpus con textos de *English Canadian Press, WordNet* y *Google* utilizado como herramienta de análisis de corpus. Con ello, aportan una serie de ejemplos de búsquedas en distintos recursos: diccionarios, pero sobre todo corpus compilados e Internet como corpus para la búsqueda de terminología en traducción.

La primera dificultad consiste en problemas relacionados con el texto original, entre los que se encuentran problemas de comprensión del texto original, que implica tanto falta de conocimiento enciclopédico, como lingüístico.

En lo que respecta a los conocimientos enciclopédicos, el traductor necesita tener nociones del tema que va a traducir, en este caso el *missile defence program*. Según las autoras, el mejor lugar en el que puede encontrar esta información es en Internet, ya que los corpus lingüísticos no están tan actualizados y están diseñados para proporcionar concordancias y no textos completos. En *Google* introducen la consulta *missile defence program* y obtienen 26.600 resultados, entre los diez primeros se encuentra la información que puede interesar al traductor, por ejemplo el objetivo del *National Missile Defence Program*, sus componentes más importantes y su ubicación en el *Ballistic Missile Defence Program*.

También puede resultar necesario comprender una oración del texto original, por ejemplo «it [National Missile Defence Program] would violate the 1972 anti-ballistic missile treaty and run counter to Canada's commitment as a signatory to the non-proliferation treaty to promote complete nuclear disarmament». En este caso puede que el traductor no comprenda si la frase se refiere a un único tratado (el de 1972) firmado por Canadá, o a dos tratados distintos (*1972 anti-ballistic missile treaty* y *nuclear non-proliferation treaty*) el segundo firmado por Canadá. La búsqueda de *1972 anti-ballistic missile treaty* y *nuclear non-proliferation treaty*, en *Google*, muestra en sus resultados que se trata de dos tratados diferentes sobre armas nucleares y que solo el segundo fue firmado por Canadá.

Otro problema de comprensión sería acceder a la forma desarrollada de acrónimos, abreviaturas o siglas que aparecen en el texto original. Por ejemplo, con las siglas *NMD* la solución la proporciona el propio texto original *National Missile Defence (NMD)*, pero en caso contrario, podría haberse encontrado la respuesta en Internet.

En cuanto a los problemas lingüísticos relacionados con el texto original, podría precisarse conocer el significado de términos en LO. En este caso el traductor puede recurrir a diccionarios, por ejemplo en papel, que están limitados por el espacio, no suelen contener términos especializados, o no contienen todos los matices de significado de una palabra y los diccionarios en CD-ROM suelen ser meras copias de los diccionarios en papel. Sin embargo, según las autoras, el uso combinado de diccionarios y corpus en esta ocasión puede ser la solución.

Uno de los ejemplos propuestos es el término compuesto *missile flight test*, que puede interpretarse de manera distinta en función del orden de sus componentes: «a test flight for missiles» o «a flight test for missiles». En un corpus de textos extraídos del *Wall Street Journal* se encuentra tanto *missile flight test* como *missile test flight*, y ambos se combinan con la palabra *ban* como en el texto original, por lo que en el corpus no se advierten diferencias. La búsqueda en *Google* permite encontrar textos en los que *NMD program missile testing* se denomina *missile flight test* y se utiliza en documentos oficiales del gobierno de EE. UU.; *missiles test flights* también aparece en los resultados de *Google* pero no en páginas oficiales.

Otro ejemplo es la palabra *cornerstone*, que pertenece más al lenguaje general. La búsqueda de la palabra en *WordNet* proporciona tres significados de los cuales se deduce que el primero es el correcto, pero al traductor le puede interesar saber si se puede hablar de «cornerstone of ... regimes». En el corpus de textos de *English Canadian Press* se

busca la palabra *cornerstone* en el contexto de *regimes*. Aunque el contexto proporcionado «[...] the Canadian Assistance Plan (the cornerstone of the welfare system), new regulatory regimes for transportation [...]» no se corresponde con la estructura que el traductor estaba buscando, de forma casual aparece la expresión «the cornerstone of the ... system» que es similar a la que el traductor quería comprobar.

Estas autoras también se refieren a problemas relacionados con el texto meta, en la transferencia a la LM, que pueden ser tanto problemas textuales como problemas lingüísticos y, dentro de estos últimos, la búsqueda de equivalentes.

En el primer ejemplo muestra cómo encontrar un equivalente de *Non-Proliferation Treaty* en el corpus bilingüe paralelo *Hansard*. Con la consulta del término original, se producen 56 resultados de concordancias paralelas, entre las que se encuentran en los contextos franceses *traité sur la non-prolifération* y *traité de non-prolifération*. Para confirmar cuál de los dos equivalentes es el más adecuado, el traductor puede buscarlos en un corpus monolingüe, por ejemplo en un corpus de prensa canadiense en francés, en el que se encuentran 77 resultados de *traité de non-prolifération* y 5 de *traité sur la non-prolifération*. Como se trata de un corpus de artículos de prensa, el traductor considera oportuno comprobar en Internet la denominación oficial del tratado. Se recuperan 46 resultados de *traité de non-prolifération*, de los cuales uno es un texto oficial, y 24 de *traité sur la non-prolifération*, uno también en un texto oficial, por lo que el traductor puede deducir que ambos términos son correctos, aunque por razones de frecuencia elige *traité de non-prolifération*.

En la búsqueda de un equivalente de *cornerstone*, el diccionario bilingüe proporciona el equivalente *pierre angulaire* tanto en sentido literal como figurado, pero no incluye un ejemplo de uso. En Internet, de los 50 resultados comprobados en *Google*, en más de 40 aparece el equivalente en sentido figurado pero no seguido de un sustantivo en plural, en las mismas construcciones que «cornerstone of ... regimes». En el corpus de prensa canadiense en francés puede comprobarse la construcción «*pierre angulaire* + des + sustantivo abstracto». De los 11 resultados de *pierre angulaire des X*, siete incluyen un sustantivo abstracto y en una de ellas el sustantivo es *programmes*.

En la consulta por problemas textuales, el traductor puede observar en el corpus paralelo *Hansard* cómo se presentan las peticiones.

#### 4.3.1.6 Rodríguez Inés (2009)

Rodríguez Inés (2009) efectúa un estudio en el que emplea una metodología para la enseñanza de la traducción con el apoyo de corpus. Por medio de esta metodología, la autora también es capaz de recopilar las estrategias de búsqueda empleadas por los estudiantes para resolver determinados problemas de traducción.

En el experimento participan 26 estudiantes de traducción españoles de último año que cursan una asignatura de traducción especializada al inglés. La investigadora selecciona un texto origen, una lista de corpus para consultar, un cuestionario en el que se identifican determinados problemas de traducción y en el que el estudiante tiene que explicar qué búsquedas ha realizado en los corpus para resolverlos y una escala de evaluación para evaluar tanto el uso de los corpus como las soluciones o equivalentes aportados por los estudiantes en sus traducciones. La profesora evalúa la validez del corpus seleccionado, de la consulta realizada, de las restricciones de búsqueda empleadas, de la ordenación de los resultados, del uso de determinadas funciones como colocaciones, agrupaciones de palabras, palabras clave, etc., y la aceptabilidad del equivalente propuesto.

Finalmente, se distribuyen otros dos cuestionarios, uno para conocer las impresiones de los estudiantes sobre las ventajas de utilizar corpus para traducir y otro para que evalúen la metodología empleada. Los resultados del primer cuestionario revelan que los estudiantes valoran positivamente el uso de corpus en traducción, con los siguientes comentarios (Rodríguez Inés 2009: 147): «I am able to use a monolingual comparable corpus in order to translate a text», «Using corpora has helped me to feel more confident about my translation solutions»; en el segundo cuestionario la evaluación de la metodología también es positiva, señalan que el uso de corpus les ahorra tiempo en la traducción y que aprender a utilizar *WordSmith Tools* les ayuda a sacar el máximo provecho de los textos paralelos.

### 4.3.2 Estudios cualitativos sobre estrategias de búsqueda en corpus bilingües paralelos y comparables

Los estudios que presentamos a continuación se centran en las estrategias de búsqueda de los traductores en corpus bilingües paralelos y comparables.

#### 4.3.2.1 Danielsson y Riddings (1996)

Danielsson y Riddings efectúan un estudio en el que utilizan un corpus paralelo en las clases de traducción. Según indican los autores el estudio está pensado para observar

los beneficios prácticos que puede tener un corpus paralelo en la traducción profesional. Sin embargo, los sujetos del experimento son estudiantes de traducción, que previamente habían realizado al menos dos cursos de los idiomas suizo, inglés, francés y alemán.

Los autores crean el corpus paralelo *PENDANT* compuesto por un millón de palabras aproximadamente, de textos traducidos con el sueco como LO y LM y las principales lenguas europeas como LO y LM. El corpus está compuesto por textos a los que los futuros traductores seguramente deberán enfrentarse en el mundo profesional. Por ejemplo, contiene textos de la empresa Volvo traducidos del sueco al inglés, principalmente documentos de calidad y manuales de mantenimiento de tractores; también incluye informes financieros de la empresa Trelleborg and Investor en sueco e inglés, catálogos con descripciones técnicas de productos de SFK en sueco, inglés, francés, alemán e italiano, informes de prensa anunciando productos de empresas en diferentes idiomas y tratados de la UE en once idiomas.

Este corpus lo emplean los estudiantes para buscar terminología en las traducciones. Los autores ponen a disposición de los estudiantes una herramienta para la consulta del corpus paralelo. Por medio de macros en *Word*, el estudiante de traducción puede seleccionar una palabra del texto origen, que se copia automáticamente en la herramienta donde se efectúa la búsqueda. Una vez el programa presenta la concordancia paralela, el estudiante puede explorar los contextos proporcionados o añadir una equivalencia que haya detectado en el corpus a la base de datos (debajo de las concordancias aparecen unas casillas destinadas a tal efecto). Los equivalentes no se añaden definitivamente a la base de datos sin la supervisión del profesor y, en el caso de que los equivalentes sean términos compuestos, estos se indexan a partir de cada palabra constituyente del término, de manera que en posteriores consultas, el usuario puede acceder a un término compuesto buscando a partir de cualquiera de sus constituyentes, o acceder a todos los términos compuestos que contienen una palabra determinada.

#### 4.3.2.2 Teubert (1996)

Teubert (1996) muestra la utilidad de un corpus paralelo en traducción en casos en los que el diccionario bilingüe no resulta útil para acceder a una equivalencia, o en los que la equivalencia puede ser difusa o variar dependiendo del contexto, y también cuando al traductor puede interesarle conocer otra información, por ejemplo sobre el uso de construcciones predeterminadas en un idioma, colocaciones o el uso de preposiciones.

Por ejemplo, en el caso de *Bericht* en alemán, que puede traducirse al francés por *rapport*, pero también por *compte rendu*, *reportage*, *bulletin*, *récit*, etc.; o el uso de las preposiciones *über*, *von* y *zu* en alemán que no tienen por qué corresponder a *sur*, *de* y *concernant* en francés; o construcciones como *er gibt Bericht* en alemán, independientemente de si va acompañado de un complemento o no, mientras que en francés *il fait son rapport* se emplea sin complemento, y se dice *il fait rapport (de sa discussion)* cuando hay complemento.

Toda esta información puede extraerse de concordancias de un corpus paralelo, como las siguientes:

Er **erstattete** dem Präsident **Bericht über** seine Auslandsreise > Il a **fait** son **rapport** au Président **sur** son voyage à l'étranger.

Er **erstattete** dem Präsident **Bericht** von den Verhandlungen mit der Gewerkschaft > Il **rend compte** au Président des pourparlers avec le syndicat.

Der **Bericht zur** Situation der Frau war sehr ausführlich > Le **compte rendu concernant** la situation de la femme était très détaillé.

Der **Bericht über** das Scheitern der Kommission war sehr ausführlich > Le **reportage sur** l'échec de la Commission était très détaillé.

Er **gibt** seinem Vorgesetzten **Bericht** > Il **fait** son **rapport à** son supérieur.

(Teubert 1996: 255)

#### 4.3.2.3 Schöffner (1998)

Esta autora trata de demostrar la utilidad de los textos paralelos (lo que otros autores han denominado corpus comparables) en clase de traducción. Para ello, se basa en su experiencia en clases de traducción con estudiantes de último año en un curso avanzado de traducción de la Universidad de Aston. En la traducción de textos, pide a los estudiantes que busquen textos paralelos tanto en lengua origen como en lengua meta (que compartan la misma función y características que el texto original) para consultarlos y detectar similitudes y diferencias que les ayuden a producir el texto meta.

En el primer ejemplo, los estudiantes traducen un manual de instrucciones de una radio del alemán al inglés. La consulta de textos paralelos tanto en LO como en LM permite comprobar cuestiones de estilo; en la transmisión de instrucciones al lector, en los textos alemanes se utiliza el imperativo, mientras que en inglés el infinitivo. Por tanto, el estudiante traduce *Gerät einschalten und Lautstärke einstellen* por *Switch on the radio and adjust the volume*.

En el segundo ejemplo, en la traducción de un folleto turístico del alemán al inglés, en los textos paralelos el estudiante descubre fórmulas prefijadas en inglés para transferir ideas del alemán *We will do everything possible to make your stay here a very pleasant one*



por *Wir bemühen uns, Ihnen Ihren Aufenthalt so angenehm wie möglich zu machen*. También averigua que la distancia se expresa en términos espaciales en los textos alemanes y en tiempo en los ingleses, lo cual se refleja en la traducción adoptada *Every second minute you find a new, romantic corner* por *Alle zwanzig Meter eine neue, romantische Ecke*.

#### 4.3.2.4 Zanettin (1998; 2001b; 2002a; 2002b)

Zanettin (1998) aborda una serie de estrategias que pueden emplearse en corpus bilingües comparables en la clase de traducción, pero que también podrían utilizar traductores profesionales como recurso complementario en la actividad de traducción.

Propone ejemplos de cómo un corpus comparable sobre deportes puede utilizarse como recurso para traducir un artículo sobre las Olimpiadas de 1992 del italiano al inglés. El subcorpus italiano está compuesto por 65.000 palabras de artículos sobre deportes de 1992 del periódico *Il Sole 24 Ore*. El subcorpus inglés está compuesto por 250.000 palabras de artículos sobre deportes de 1992 de los periódicos *The Independent* y *The Daily Telegraph*. Zanettin (2001b) presenta otro experimento en el que estudiantes de traducción de grado italianos realizan una traducción inversa de un artículo de prensa sobre los Juegos Olímpicos de 1992 del italiano al inglés con ayuda de este corpus comparable. La traducción final debe hacerse de forma individual pero la búsqueda en corpus se efectúa en grupos.

En Zanettin (1998) se proporcionan los siguientes ejemplos de búsqueda en el corpus comparable. El primero está relacionado con la traducción de lenguaje figurado, como en el caso de *salire il gradino più alto del podio* cuya traducción literal *to climb onto the highest step of the podium* no suena natural. En el corpus italiano la búsqueda de *podio* revela contextos en los que esta expresión significa ganar la medalla de oro. La consulta de *podium\** en el corpus inglés muestra que *podium* se refiere al mismo concepto que *podio*, pero no se utiliza junto con *the highest step* en el sentido de ganar la medalla de oro. Sí que aparecen expresiones como *stand on the podium*, *medal podium*, *two Americans stood on the winner's podium*. Este tipo de búsquedas, según el autor, evitan que se utilice una traducción literal.

El mismo corpus se emplea para comprobar similitudes y diferencias en aquellas categorías muy similares entre lenguas como los nombres propios o los cognados. Por ejemplo, en la búsqueda de *Mitterrand* y de las palabras *francese* y *presidente* en el corpus italiano se deduce que en italiano se habla de *il presidente francese Mitterrand* o *François*

*Mitterrand*, mientras que en inglés se habla de *President Mitterrand* o *François Mitterrand* o de *Mr* o *M Mitterrand*. No se emplea en inglés la palabra *French* para evitar frases muy largas.

Según Zanettin, este corpus también puede consultarse para observar el estilo, como qué verbos introducen el discurso directo, por ejemplo, cuando se citan las palabras literales del presidente. En inglés, se consultan los pronombres *Mr*, *Mrs*, *Ms* y *Pr\** con comillas en el mismo contexto (a unas siete posiciones de distancias a izquierda o derecha). El verbo más utilizado en el corpus de textos en inglés es *said* y también, pero en menor medida, *add*. En el corpus italiano se consulta *ha* (*has*) con comillas en el mismo contexto y se observa que el verbo más utilizado es *dire*, así como *aggiungere* y también *dichiarare*, *affermare*, *chiamare*, *concludere*, *insistere*, *ribadire* (todos en tercera persona del passato prossimo). Además, puede observarse que *ha detto* en italiano suele situarse dentro de las comillas del discurso directo, entre guiones, mientras que *said* en inglés suele aparecer fuera de las comillas al final de frase.

Otro ejemplo se centra en *prices* y *prezzi*, que como equivalentes de traducción no presentan problemas, aunque al traductor puede interesarle conocer información fraseológica sobre los mismos. Consultando en el corpus, se averigua que *prezzi* suele aparecer precedido del artículo *i* (*i prezzi*), mientras que *prices* se emplea sin artículo. Otra información que se obtiene del corpus es que mientras que en italiano se emplean los siguientes grupos nominales *l'aumento dei prezzi* o *la discesa dei prezzi*, en inglés son preferibles los grupos verbales *prices soared* o *prices fell*. También puede accederse a colocaciones verbales de *prices* en inglés como *prices go up*, *shoot up*, *increase*, *soar*, *rise* y *skyrocket*; o *go down*, *abate*, *fall*, *plunge*, *tumble* y *plummet*.

En Zanettin (2001b) incluye más ejemplos de búsqueda en el corpus comparable de deporte. El primer ejemplo es la búsqueda de nombres propios en el corpus. En el texto original se menciona que el deportista Biondi es un *sorvegliato speciale* que se emplea en el lenguaje legal para indicar que una persona se encuentra bajo vigilancia policial, en este caso se emplea de manera metafórica y significa que Biondi se encuentra bajo presión para defender su supremacía. Una búsqueda en el corpus del nombre propio Biondi permite observar una expresión que se emplea para describir al atleta *defending champion* y que el estudiante selecciona para la traducción. Un caso similar es el de Kieren Perkins. El estudiante de traducción pensaba utilizar en la traducción *the Australian Kieren Perkins*, pero una búsqueda en el corpus del nombre propio proporciona otras opciones *Australia's Kieren Perkins* y *Kieren Perkins of Australia*. Posteriormente, una búsqueda de adjetivos

de nacionalidad, nombres de países utilizados como posesivos y la preposición *of* seguida del nombre de un país hace ver al traductor que esta es la forma más utilizada para referirse al origen de los atletas en el corpus y, por tanto, la selecciona para la traducción.

En otro ejemplo, conocer la estructura del discurso puede ayudar a encontrar un equivalente, como en el caso de *In vasca* (literalmente: en la piscina) que se emplea en el texto origen como titular para indicar que el artículo es sobre natación. Una búsqueda de *swim\** en el corpus muestra que un equivalente podría ser *Swimming*, ya que 14 de las 102 concordancias recuperadas se encuentran en titulares.

En cuanto a cuestiones de estilo, en el texto origen aparece la expresión *50 e 100 stile libero* que podría traducirse de manera intuitiva por *50 and 100 freestyle*, lo cual puede comprobarse con una búsqueda de números: *\*0* recupera demasiados resultados, pero *?00* y *\*00m* revela que las distancias se expresan, junto con la unidad de medida *metres* en el corpus inglés: *100 metres freestyle* o *100m freestyle*, *50 and 100 metres freestyle*.

Además, este autor señala que las soluciones en el corpus pueden encontrarse por casualidad, buscando un término puedes acceder a otros que quizá sean útiles posteriormente y que el corpus es especialmente útil para desmentir hipótesis erróneas del traductor. Por ejemplo, en la expresión italiana *re del mezzofondo* (literal: rey de la media distancia). La búsqueda de *king* en el corpus inglés recupera contextos en los que el término se refiere al rey de un país, *the king of Spain*, o de manera metafórica a una actividad pero no a una persona, *basketball is the king in Lithuania*. *Big man* aparece de forma metafórica pero también literal, *top dog* aparece en el corpus pero entre comillas, lo que sugiere que es una expresión del discurso oral. Finalmente, *top performer* aparece en una ocasión en el corpus y es el equivalente seleccionado por el estudiante de traducción.

Zanettin (2002a: 11) extrae una serie de ejemplos de cómo un traductor mediante la consulta de un corpus paralelo puede detectar casos en los que no existe un equivalente en LM para un término en LO (lo que se denomina equivalencia cero). Realiza un análisis de dos traducciones del inglés al italiano de varias novelas de Salman Rushdie efectuadas por dos traductores italianos (Zanettin 2001a). Por ejemplo, en el corpus paralelo se puede observar que el término *edges*, que suele combinarse con distintas preposiciones (*around the edges*, *at the edges*), en italiano no se traduce literal sino que se omite.

...biting the skin **around the edges** of a nail...  
...mordicchiandosi la pelle attorno all'unghia...  
...**around the edges** of Gibreel Farishta's head...  
...intorno alla testa di Gibreel Farishta...  
...**around the edges** of the circus-ring...  
...intorno alla pista da circo...  
...and there was a fluidity, an indistinctness, **at the edges** of them...  
...vicinissime a loro c'erano una fluidità e un'indeterminatezza...  
...the horses grew fuzzy **at the edges**...  
...i cavali diventavano sempre più sfocati...  
...blurred **at the edges**, my father...  
...con la mente annebbiata, mio padre...  
...looking somewhat ragged **at the edges**...  
...con l'aria di un uomo distrutto...  
...Mrs Qureishi, too, was beginning to fray **at the edges**...  
...anche Mrs Qureishi si stava consumando...

(Zanettin 2002a: 11, énfasis añadido)

Zanettin (2002b) trata de mostrar la utilidad de un corpus ad hoc (*DIY corpus*) para la traducción, frente a otros recursos tradicionales como los textos paralelos en papel.

Para ello efectúa un experimento con estudiantes de traducción, de tercer y cuarto año de la Universidad de Bolonia en Forlì. El estudio se desarrolla durante dos semanas, en cinco clases de dos horas, en las que se pide a los estudiantes que realicen un encargo real de traducción, bien como un ejercicio para clase o como una traducción remunerada, no se imponen restricciones sobre tipo de texto ni idiomas de traducción, pero se proporcionan dos textos, uno en inglés y uno en italiano, en el caso de que alguien no encuentre un texto para traducir. Los textos seleccionados son variados (traducción de un artículo de enciclopedia sobre el cáncer de próstata del inglés al italiano, parte de un catálogo textil del italiano al español, parte de un catálogo de candados para bicicleta del italiano al inglés, un artículo sobre terremotos de una revista científica del italiano al inglés, un folleto publicitario sobre diamantes del inglés al italiano).

Los estudiantes traducen individualmente o por grupos de dos o tres personas. Para la consulta utilizan bancos de datos terminológicos, programas de traducción asistida, diccionarios electrónicos y corpus ad hoc descargados por ellos mismos desde Internet y analizados con *WordSmith Tools*. Al final del experimento se les pide que presenten un informe con las ventajas e inconvenientes de utilizar corpus ad hoc para la traducción.

Entre los resultados se incluye el método empleado por los estudiantes para la compilación de los corpus ad hoc. Primero, para la mejor comprensión del texto original, los estudiantes buscan terminología desconocida en LO en glosarios en línea, identifican equivalentes en LM que emplean como términos de búsqueda en buscadores; otros estudiantes utilizan listas de páginas web clasificadas por ámbitos en directorios de Internet. Se crean corpus de entre 10 y 50 textos de 5.000 a 40.000 palabras. Los textos se seleccionan por cuestiones tanto léxicas como por el contenido de imágenes.

En cuanto a las búsquedas en los corpus, se accede a información sobre terminología, fraseología y colocaciones, con consultas alrededor de términos ya conocidos para acceder a términos desconocidos. Por ejemplo, en el texto sobre terremotos, en la búsqueda de un equivalente en inglés para *cedimenti strutturali gravi*, la consulta de *structural\** genera entre las concordancias *heavy structural damage*. Asimismo, para confirmar hipótesis, por ejemplo, sabiendo previamente que *antifurto per bicicletta* es *bicycle lock*, el estudiante busca *lock* en el corpus y ordenando las concordancias por la primera palabra a la izquierda encuentra *cable lock, coiling cable lock, U-lock, disc lock, etc.*

Entre las ventajas y desventajas de los corpus señaladas por los estudiantes en sus informes se encuentra la utilidad de las imágenes y los elementos multimedia de las páginas web, pero la dificultad de encontrar terminología relevante directamente en Internet sin compilar un corpus. Asimismo, la confianza que supone trabajar con un corpus en la adopción de soluciones, especialmente en la traducción inversa y que esta ventaja es mayor en traducciones más largas. No obstante, se advierte de las dificultades de uso de *WordSmith Tools*, por ejemplo en cuanto que no se puede pasar de una concordancia del corpus a la consulta directa de la página web completa, con información de las imágenes. Por esa razón, el autor aboga por la creación de aplicaciones que integren las funciones de buscadores y programas de gestión de corpus.

#### 4.3.2.5 Gavioli y Zanettin (2000)

Gavioli y Zanettin (2000) observan cuáles son las ventajas de los corpus comparables frente a los paralelos como recurso de búsqueda. Los sujetos de su estudio son estudiantes de traducción de último año de la Scuola Superiore di Lingue Moderne per Interpreti e Traduttori (Universidad de Bolonia), con lengua materna italiano, lengua extranjera inglés y habilidades de traducción avanzadas.

Los autores compilan un corpus comparable de artículos de investigación sobre el SIDA, compuesto por 30 artículos en inglés (193.685 palabras) y 30 en italiano (109.904 palabras). Lo utilizan como instrumento de búsqueda en clase de traducción con la herramienta *WordSmith Tools*.

Incluyen una serie de ejemplos de búsqueda en el corpus para verificar el uso de términos en LM. Un ejemplo de búsqueda se propone en la traducción de la siguiente oración: *In questo lavoro sono state esaminate 183 biopsie epatiche di estilista con o senza marcatori HBV*, que intuitivamente podría traducirse por *In this paper 183 **hepatic biopsies** of alcoholic with or without HBV **markers** were examined* (Gavioli y Zanettin 2000: 65). Si se comprueba en el corpus inglés *biops\**, se observa que no hay ninguna colocación de *hepatic biopsy/ies* mientras que aparecen 47 resultados de *liver biopsy/ies*. Por otro lado, las concordancias de *marker/s* revelan que hay un resultado de *patients without hepatitis C markers*, pero parece que son más comunes estructuras del tipo *(to be) positive/negative for HBV/HCV markers*. Una búsqueda de *marker/s* en contexto con *HBV* o *HCV* y con *positive* o *negative* recupera más resultados.

Otro ejemplo es la búsqueda de un equivalente en italiano para *patient(s) experienced* con el sentido de sufrir una enfermedad, normalmente combinado con el adjetivo *severe*. La consulta en un corpus paralelo de 27.000 palabras sobre el SIDA recupera la expresión traducida *con/senza esperienza di/con* que aparentemente se utiliza como equivalente de *experienced* en el sentido indicado. Sin embargo, podría tratarse de un calco, ya que la búsqueda de *sper\** en un corpus italiano de textos originales de 100.000 palabras muestra que *sperimentare* se emplea asociado a la palabra *paziente/i* refiriéndose a datos proporcionados en un experimento (*nella/dalla nostra esperienza*) o a terapias (*la monoterapia intermittente è stata sperimentata...*). La consulta de *pazient\** con *grav\** en un contexto de 20 palabras a izquierda y derecha, en el corpus de textos originales, recupera concordancias de las siguientes expresiones: *i pazienti sono affetti da*, *i pazienti hanno/presentano*, que son más naturales y que son las que se seleccionan para la traducción.

De los datos de este experimento, Gavioli y Zanettin (2000: 79) concluyen que el corpus ideal es un *corpus recíproco* formado por textos originales comparables en dos lenguas y sus traducciones correspondientes.

#### 4.3.2.6 Bowker y Pearson (2002)

Bowker y Pearson (2002: 103-107) proporcionan una serie de estrategias de búsqueda en un corpus paralelo para el acceso a la terminología, tanto por estudiantes de traducción o de segundas lenguas.

Como ejemplo de corpus paralelo utilizan un corpus de contratos colectivos de una institución de Canadá. El corpus consta de 226.601 tokens y está alineado a nivel de oración (Bowker y Pearson 2002: 193).

La primera estrategia propuesta consiste en identificar equivalentes terminológicos, por ejemplo un equivalente en inglés del término francés *plaignant*. La primera idea es buscar el término francés en el diccionario bilingüe, que proporciona tres opciones *litigant*, *plaintiff*, *complainant*, todas válidas en un contexto legal. El diccionario no proporciona contextos ni se indica si alguno de los equivalentes puede utilizarse en un contrato colectivo. En lugar de tratar de adivinar cuál sería el equivalente idóneo, puede buscarse en un corpus paralelo de la misma forma que en el diccionario bilingüe. Las concordancias de *plaignant* no sugieren que ninguno de los términos propuestos por el diccionario sea el adecuado. En los textos canadienses, el equivalente empleado es *grievor*. Estas autoras indican que utilizar un corpus paralelo es un valor añadido para el cliente ya que los términos que se emplean son consistentes con los que se han utilizado en otras traducciones de textos similares.

Otra estrategia consiste en investigar el uso para averiguar cómo expresar ideas de forma natural e idiomática en un área de conocimiento. Por ejemplo, sabemos que *grief* en francés puede corresponder en inglés a *grievance* pero desconocemos con qué términos se combina en una oración. En un corpus podemos observar que *grievance* se combina con *initiated*, *filed*. Según estas autoras, el corpus paralelo es una fuente de valor incalculable para acceder a las colocaciones, ya que los diccionarios especializados raras veces contienen este tipo de información.

El corpus paralelo también se emplea para investigar el estilo. Como señalan Bowker y Pearson (2002: 197) los traductores no trabajan con palabras aisladas, sino con textos completos por lo que resulta necesario realizar traducciones que suenen naturales, además de que contengan la terminología adecuada y se empleen las estructuras gramaticales correctas. Por ejemplo, el corpus paralelo revela que las construcciones que se refieren a *grievance* aparecen en voz pasiva (*grievance initiated by a member*). Además, este tipo de textos se escriben en tercera persona, hay pocas referencias en primera persona en el corpus, y las oraciones suelen ser largas y complejas.

#### 4.3.2.7 Hansen y Teich (2002)

Hansen y Teich (2002: 3) muestran la utilidad para el traductor de un corpus paralelo anotado para la resolución de determinados problemas de traducción mediante la formulación de consultas de tipo gramatical.

Por ejemplo, la consulta [pos="VH.\*"] [pos="RR.\*"] {0,1} [pos="VVN"] permite comprobar que la traducción en el corpus del present perfect y past perfect del inglés al francés han sido traducciones directas o se ha empleado el *passé antérieur* o la construcción *venir de*; en alemán la solución adoptada ha sido la traducción directa, el pasado o el presente de subjuntivo.

Propone además una serie de sugerencias que mejorarían las búsquedas para el traductor en este tipo de corpus, como que pudieran restringirse las consultas en los corpus paralelos tanto en la LO como en la LM, la posibilidad de acceder de manera simultánea a corpus comparables de textos originales para la verificación y el uso de anotaciones más complejas como información semántica y discursiva.

#### 4.3.2.8 Vargas Sierra (2002)

Vargas Sierra (2002) trata de mostrar la utilidad de los programas de concordancias y de la búsqueda en Internet para la traducción especializada como recurso complementario, frente a diccionarios y bancos de datos que, según la autora, tienen un límite sobre todo con respecto a términos más especializados, información fraseológica o colocaciones y registro.

Compila un corpus del sector de la piedra natural compuesto por un corpus paralelo de 14 normas en inglés con su traducción al español y un corpus comparable bilingüe con 38 textos en inglés y 27 en español.

En este corpus presenta una serie de consultas que pueden resultar útiles al traductor para encontrar equivalentes. Por ejemplo, busca equivalentes en español para los términos en inglés *coarsed textured and riven slab*, *slab bedded in mortar*, *capping with mortar*, *squared rough block*. La búsqueda en diccionarios no proporciona un equivalente preciso. La consulta de las concordancias de *baldosa* y *mortero* en el corpus permite encontrar los siguientes equivalentes *coarsed textured and riven slab* > *baldosa partida y de textura gruesa*, *slab bedded in mortar* > *baldosa embebida en mortero*, *capping with mortar* > *refrentado con (capa de) mortero*, *squared rough block* > *bloque escuadrado*.



Finalmente, la autora explica que, si el traductor no tiene claro si los términos obtenidos del corpus son correctos, puede comprobarlos en Internet, que ella misma denomina *mega-corpus*. Le parece útil que en los resultados *Google* destaque las palabras de búsqueda en contexto y que además permita acceder a páginas *En caché*, donde resalta en las páginas web recuperadas las palabras clave en colores diferentes, por lo que se facilita su localización.

#### 4.3.2.9 Pearson (2003)

Pearson (2003) indaga sobre cómo un corpus paralelo puede resultar útil como medio de enseñanza a estudiantes de traducción y como ayuda a la traducción profesional. El uso de corpus paralelos puede ayudar al (futuro) traductor a observar cómo los traductores actúan cuando se ven limitados por la necesidad de mediar entre dos lenguas y culturas. Por ejemplo, en qué medida se han podido hacer transferencias directas de una lengua a otra, o ha sido necesario realizar algún tipo de adaptación, o incluso recurrir a la omisión. Este tipo de situaciones no ocurren en contextos monolingües y, por tanto, esta información no puede encontrarse en corpus comparables. Por ello, Pearson (2003: 17) propone el uso de corpus paralelos para observar qué estrategias han empleado traductores profesionales en la resolución de determinados problemas de traducción.

Esta autora compila un corpus paralelo (187.159 tokens en LO y 168.518 en LM) compuesto por una serie de artículos científicos traducidos del inglés al francés. Asimismo, recopila artículos científicos originalmente escritos en francés. Los textos bilingües los prepara para poder ser analizados con una herramienta de gestión de corpus que le permite extraer concordancias bilingües.

A partir de este corpus se observa cómo se han traducido en LM determinadas informaciones con rasgos culturales de la LO. La autora advierte que la solución adoptada no tiene por qué ser la correcta (ya que puede depender de las directrices de la editorial, el público meta o el fin del artículo), simplemente se trata de una observación de traducciones anteriores.

Un ejemplo es la obtención de información estadística para la adaptación cultural. Por ejemplo, si en el texto origen se menciona que el SIDA ha acabado con la vida de más de 350.000 personas solo en Estados Unidos, esta información deberá adaptarse a la cultura meta: el SIDA ha acabado con la vida de más de 50.000 personas solo en Francia.

Otros ejemplos se refieren a la traducción de nombres propios. Uno de ellos es la traducción de los nombres de unos famosos psiquiatras americanos, que se omiten en francés y se alude de forma general a *les psychiatres* (los psiquiatras). En cuanto a la traducción de nombres de universidades estadounidenses, de 102 nombres de universidades que aparecen en el corpus, 66 se han traducido literalmente empleando el término *université*, en 36 casos se han traducido adoptando otra estrategia o no se han traducido. En los casos en los que el nombre de la universidad se ha dejado sin traducir, normalmente la oración entera en la que aparecía la universidad tampoco se traducía. En este caso concreto se ha mantenido la LO porque la información no es imprescindible para la comprensión en la cultura meta o porque estas universidades no tienen un reconocimiento internacional. En los casos en los que no se ha adoptado una traducción literal, en la traducción se ha incluido el lugar en el que se encuentra la universidad, en vez de su nombre, por ejemplo, la *State University of New York at Albany* se ha traducido por *l'Université d'Albany*.

En cuanto a la traducción de universidades europeas, algunas universidades europeas en los textos originales en inglés americano incluyen la ubicación junto al nombre de la universidad (*University of Lund in Sweden*), mientras que en otras no lo especifican (*University of Oxford*). Esto se debe seguramente a la percepción del escritor de los textos originales sobre el conocimiento de determinadas universidades europeas en la sociedad americana. En la traducción al francés, las referencias a los lugares en los que se encuentran las universidades europeas se omite (*l'Université de Lund*) porque se ha supuesto que el lector francés tendrá suficiente conocimiento de geografía europea para saber la ubicación de dichas universidades.

Asimismo, respecto a la traducción de los nombres de institutos americanos, de 32 nombres de institutos en LO, 20 se traducen de forma literal, 10 no se traducen y 2 se traducen de forma no literal. En este ejemplo, se destaca la decisión de traducir *National Institute* por *Institut américain*, como por ejemplo en *National Cancer Institute (NCI)* que se traduce por *l'Institut américain du cancer*. El propósito en este caso es el de evitar la ambigüedad en el público francés.

Pearson concluye que, aunque la muestra estudiada no permite extraer conclusiones definitivas acerca de patrones o reglas en la traducción de determinadas referencias culturales, sí que se ha visto que un corpus paralelo puede orientarnos para resolver un determinado problema de traducción.

#### 4.3.2.10 Sánchez-Gijón (2003a; 2004a; 2004b)

Sánchez-Gijón (2003a: 307-330; 2004a: 154-281; 2004b) analiza de forma sistemática un corpus multilingüe comparable en inglés, español y catalán en búsqueda de definiciones y relaciones conceptuales, relaciones jerárquicas y no jerárquicas.

A su vez, la autora muestra una serie de estrategias de búsqueda puntuales que pueden emplearse en el corpus, ante problemas que pueden surgir en la traducción: 1) que el traductor conozca una forma lingüística y desconozca el significado o la información conceptual dentro de un área temática, 2) que el traductor desconozca tanto la forma lingüística como la información conceptual y en el proceso de búsqueda adquiera la información conceptual que no conocía y 3) que el traductor desconozca una forma lingüística y conozca el significado o la información conceptual (Sánchez-Gijón 2004a: 202-203).

Los primeros ejemplos de búsqueda se proponen para la traducción de un texto no especializado, un artículo de prensa, donde el emisor es un experto pero el receptor no. El corpus que se emplea para la búsqueda contiene también textos que comparten estas características. En el primer supuesto, ante la unidad lingüística conocida *peak*, se busca información conceptual desconocida. Primero se consulta la forma *peak* en el diccionario monolingüe *Collins Concise Dictionary*, en la entrada de *peak* se observa que se refiere a lo más alto. También se buscan concordancias de *peak* en el corpus. Con la búsqueda de colocaciones se observa si forma parte de una unidad de significado ampliado, por las colocaciones más frecuentes en una primera posición a la izquierda se accede a las combinaciones *shower peak*, *storm peak* y *leonid peak*. También coaparecen expresiones de tiempo *year*, *nov* (*november*), *hour*, *years*, *night*, *time*, *hours*, *when* y los verbos *expect* y *predict* lo que sugiere que es un fenómeno que se calcula previamente. La agrupación *peak of* demuestra que formará parte de un concepto más amplio, en los contextos se encuentra *peak of a storm*, *peak of meteor activity*, *peak of the leonid activity*. A partir de la expresión de tiempo *during the peak* se observa que esta unidad se calcula por la cantidad de meteoros por hora y aparecen 18 resultados de *peak rate*. Todo ello hace concluir al traductor que *peak* es el punto máximo que se calcula o se prevé de un acontecimiento que implica un alto número de meteoros por hora.

A partir de esta información pasa a la búsqueda de equivalentes desconocidos en español. Empieza con una hipótesis de tipo inductivo, basada en la intuición de que *peak* equivale a *pico*. La consulta de *pico* en el subcorpus español genera pocos resultados en comparación con *peak* en el subcorpus inglés, lo que lleva al traductor a sospechar que

puede existir otro equivalente. La segunda hipótesis se genera por un proceso deductivo, se obtiene de información del contexto. Si *peak* se mide por el volumen de meteoros por hora, entonces el equivalente de *peak* deberá aparecer en el corpus como colocación de *meteoro* y *hora*, por lo que se buscan contextos en los que aparezcan *meteoro* y *hora*. En las concordancias recuperadas se emplean las expresiones *máximo*, *tasa máxima*, *de máximo*, *máximo meteórico*, *tasas de actividad máximas*, *tasa máxima* o *en el máximo*. Estas expresiones pueden verificarse en el corpus.

En la búsqueda de equivalentes en catalán, las estrategias de búsqueda son iguales. La primera hipótesis es que *peak* equivale a *pic*, lo que se descarta por el reducido número de resultados. La segunda hipótesis lleva a la búsqueda de *meteor* y *hora* y como hay pocos resultados, se consulta solo *hora*. Se encuentra un resultado en el que aparece *intensitat*, pero el corpus es demasiado pequeño para llegar a una solución concluyente.

Otro ejemplo es *shower* y *storm*. En la búsqueda de información conceptual desconocida, se consulta *shower* en el diccionario. También se accede a las concordancias del corpus (1095 resultados) y a las colocaciones que sugieren que son frecuentes *meteor shower*, *leonid shower* y *leonid meteor shower*. También aparecen indicadores de tiempo *november* y *year*. *Storm* también se consulta en el diccionario, en las concordancias del corpus y las colocaciones. Parecen frecuentes *meteor storm*, *leonid storm* y *leonid meteor storm*, solo hay una expresión de tiempo *year* y una alta aparición de *satellites*. La consulta de los contextos combinando *shower* y *storm* para comprobar si son sinónimos indica que existen diferencias, aparecen contextos como *shower may turn into a storm*, definiciones comparativas y un contexto en el que se indica que *storm* es un tipo de *shower* que se diferencia por la intensidad.

En la búsqueda de equivalentes desconocidos en español, la hipótesis es de tipo introspectiva, parte de los conocimientos del traductor sobre su lengua materna, por lo que puede traducir *shower* por *lluvia* y *storm* por *tormenta*. Por ello, consulta los contextos de *meteor\** con el objetivo de encontrar *lluvia* y *tormenta* en las primeras posiciones a la izquierda. Esta hipótesis se confirma, aparecen muchos resultados de *lluvia de meteoros* y *tormenta de meteoros*. También hay algunos resultados de *chubasco de meteoros* pero siempre en el mismo documento, lo que sugiere que es un recurso estilístico del autor de dicho texto para evitar la repetición.

Con la comprobación de *lluvia\** se accede a un amplio número de resultados de *lluvia de estrellas* y *lluvia de estrellas fugaces* y también *lluvia de meteoros*, por lo que el traductor concluye que son sinónimos en el ámbito de las Leónidas, aunque los contextos

recuperados no le permiten establecer si hay diferencias. La consulta de *tormenta\* de* recupera *tormenta de estrellas y tormenta de meteoros* y también *tormenta de Leónidas y tormenta de las Leónidas*. Esto último sugiere que por el principio de economía lingüística el concepto completo *lluvia/tormenta de meteor\* Leónidas* se sustituye por *lluvia/tormenta de meteor\**, *lluvia/tormenta de las Leónidas* o simplemente *lluvia/tormenta*.

En la búsqueda de equivalentes en catalán, la misma hipótesis de tipo introspectivo sugiere *pluja* y *tempesta*. Se consulta *meteor\** y el resultado confirma *pluja* pero rechaza *tempesta*. La consulta de *pluja/pluges* recupera *pluja/ges d'estels*, *pluja/ges de Leòníds*, *pluja/ges de les Leòníds*, *pluja/ges de meteor\**. La consulta de *tempesta* recupera *tempesta meteòrica*, aunque no es habitual en estos contextos por lo que el traductor deduce que solo la entenderán los expertos o semiexpertos.

En la traducción de un texto especializado, un artículo académico, donde tanto emisor como receptor son expertos, el subcorpus que se emplea para la búsqueda contiene también textos especializados.

En el primer supuesto, ante la unidad lingüística conocida *peak*, un vaciado del corpus especializado inglés comparándolo con el corpus general *BNC* sugiere que *peak* es una palabra clave del corpus especializado. La consulta de *peak\** en el corpus permite observar las colocaciones *activity*, *time*, *location*, *rate* y *ZHR*. No aparecen expresiones genéricas de tiempo como sucedía en el corpus general, siempre se utilizan las expresiones *peak location* y *peak time* para indicar el lugar y el momento de este fenómeno. También se accede a las agrupaciones *peak ZHR* (*Zenith Hourly Rates*) y *peak activity*. En las combinaciones con verbos *peak* aparece como objeto en *peak observed* y como sujeto en *peak occurs*. Ordenando los contextos por la izquierda, se accede a *the storm peak*, *the first peak*, *the second peak*, *the main peak*, *the «traditional» peak* y *with a peak*.

En cuanto a la búsqueda de equivalentes desconocidos el corpus español especializado no es comparable ni cualitativa ni cuantitativamente al inglés, pero pueden emplearse algunas estrategias para acceder a equivalentes. En el caso de *peak*, se localiza a partir de la unidad de medida *ZHR* (*Zenith Hourly Rates*). Como el traductor no conoce el equivalente en español de esta medida, la búsqueda debe empezar por la comprobación de alguna hipótesis previa, por lo que se centra en encontrar el equivalente de *hourly* con la consulta de *hora\**. Ordenando las concordancias por la primera posición a la izquierda y la primera a la derecha, se observa *THZ* (*Tasa Horaria Zenital*). La consulta de *THZ* ayudará al traductor a encontrar el equivalente de *peak*. En los contextos aparecen *máxim\** e *intens\**. Como la equivalencia entre *peak* y *máximo* ya se ha verificado anteriormente, la

búsqueda se centra en *intens\**, que aparece en contextos como *la intensidad llegue a 50.000 meteoros por hora, el pico de THZ más intenso, predicción de la intensidad*, lo que confirma también este equivalente.

Aunque el corpus catalán no cuenta con textos especializados, se utiliza la misma estrategia que en el corpus español para llegar al equivalente de *peak*. También se empieza por la consulta de *hora*. Se accede a un resultado de *Tasa Horària Zenital* y no aparecen las siglas *THZ*.

En otro ejemplo de búsqueda, ante la forma lingüística conocida *activity* se desconoce si en el texto especializado tiene un uso específico. Se consulta *activity* en el diccionario y después en el corpus. En primera posición a la izquierda encontramos palabras clave del campo: los elementos que presentan *activity* son *draconid, meteor, Leonid, shower, storm, ursids*, el modo *enhanced, flat, high(er), low(er), maximum*, etc. y la periodicidad *annual, initial*. Los verbos que con más frecuencia aparecen en primera posición a la derecha *observe* y *predict, drop* y *exceed* indican que se cuantifica y fluctúa.

En la búsqueda de equivalentes en español, la hipótesis introspectiva lleva a la consulta de *actividad*. Se observa que *actividad* aparece en combinaciones como *actividad máxima/buena/suficiente*, etc., también en *actividad meteórica/de las Leónidas/de las Gemínidas*, etc. y se observan otras características como la combinación con expresiones temporales y verbos. La conclusión es que *activity* y *actividad* son equivalentes aunque presentan estructuras sintagmáticas ligeramente distintas. Con la misma estrategia se consulta *activitat* en el corpus catalán, las colocaciones también sugieren rasgos similares, pero no pueden extraerse soluciones concluyentes debido a la naturaleza distinta de los corpus.

En el segundo supuesto, en el que el traductor desconoce las unidades lingüísticas, se propone el ejemplo de las siglas *FWHM*. La consulta en el corpus recupera la forma ampliada *full width at half maximum*, donde además los contextos aportan información sobre que es un aspecto relacionado con *peak* que se mide en tiempo (*FWHM duration, FWHM of about 17 hours*). También aparece la medida *ZHR*.

En la búsqueda de equivalentes en español, la primera hipótesis es buscar la sigla original en el subcorpus en español, pero no produce resultados. La segunda hipótesis es buscar a partir de elementos que aparecen en el contexto, por ejemplo *hora\**, pero tampoco se encuentran resultados válidos. Otra consulta es *dura\** para observar contextos en los que se haga referencia a la duración del fenómeno. Se recuperan expresiones de duración pero

sin la precisión de la sigla inglesa. En cuanto a los equivalentes en catalán las estrategias son las mismas que en español y los resultados tampoco son concluyentes.

Otro ejemplo es *nodal*, que en el corpus inglés aparece siempre designando a un sustantivo que aparece en primera posición a la derecha *the + nodal + [sustantivo]* como en las unidades *the nodal crossing* y *the nodal longitude*. En todos los contextos aparece relacionado con *comet* o *Temple-Tuttle*, que es un cometa. También se analiza *node\** con lo cual se obtiene más información (que es un elemento del cometa, *comet node*, es un elemento que está en órbita con el cometa, *comet's orbit node*, la Tierra lo cruza, *Earth crosses the comet's node*, además lo cruza en un momento y lugar precisos lo que provoca que la actividad de la lluvia de meteoritos aumenta a los niveles más altos).

A partir de esta información, en el subcorpus en español, se consultan los contextos de *Tierra* y *cruzar* (*cruz\**) con la intención de que incluyan información sobre la órbita del cometa. El punto de órbita en el que el cometa está más cerca de la Tierra es el *node*. Para la búsqueda de equivalentes se consulta *nod\**, lo que genera dos resultados. A partir de las características conceptuales en inglés se busca *Tierra* y *órbita*, pero en los contextos no se menciona la distancia entre el cometa y la Tierra. Se hace la misma consulta en el resto del corpus español (no solo en el subcorpus especializado), la noción de *node* aparece parafraseada. En la consulta del corpus catalán a partir de *Terra* y *òrbit\** y de *Terra* y *creuar* los contextos muestran la relación entre la proximidad del cometa a la órbita de la Tierra y las intensas lluvias de meteoritos, pero no se detecta tampoco ninguna denominación específica, por lo que los datos proporcionados por el corpus no son suficientes para proponer un equivalente.

#### 4.3.2.11 Lagoudaki (2006)

Presentamos también algunas de las aportaciones del estudio de Lagoudaki (2006) sobre memorias de traducción desde la perspectiva de los usuarios. Según esta autora, la mayoría de programas de gestión de memorias de traducción están diseñados con un enfoque tecnológico, no están centrados en el usuario y presentan numerosas opciones complejas, de escasa utilidad y difíciles de aprender para los traductores. La mayoría de los diseñadores no tienen en cuenta a los usuarios en las distintas fases del proceso de elaboración de estas herramientas. Solo tienen en cuenta la opinión de los usuarios en la fase de prueba, justo antes de que el programa salga al mercado, cuando el producto está prácticamente acabado y existen muy pocas posibilidades de cambio, o incluso algunas veces el programa se distribuye sin que los usuarios finales realicen ninguna prueba y los

proveedores piden opinión a los usuarios cuando ya no existen posibilidades de mejora hasta la próxima actualización de la herramienta, que normalmente es de año en año. Según esta autora, tener en cuenta la opinión de los usuarios finales, es decir los traductores, en todas las fases del proceso de elaboración de los programas de traducción asistida sería esencial para el desarrollo de herramientas verdaderamente centradas en el usuario.

Se diseña un cuestionario en inglés con preguntas cerradas que se distribuye a través de Internet a listas de distribución de traducción, empresas de traducción y localización, asociaciones de traductores, instituciones académicas que forman a traductores, así como organismos públicos que cuentan con departamentos de traducción. El número final de cuestionarios analizados es de 699. Desde el punto de vista de las estrategias de búsqueda, los resultados muestran que las memorias de traducción se utilizan en interacción con otros recursos en la búsqueda de equivalentes. Así, cuando no es posible encontrar una traducción en la memoria, los traductores acceden, por este orden, a diccionarios electrónicos en CD-ROM (30%)<sup>91</sup>, diccionarios y glosarios en línea (21%), diccionarios en papel (17%), uso de buscadores de Internet, sobre todo *Google*, para encontrar una traducción (15%), glosarios de traducción antiguos (13%), textos monolingües extraídos de Internet (0,9%), preguntas a colegas (0,8%), preguntas a expertos (0,8%), textos bilingües extraídos de Internet (0,7%), textos impresos sobre el tema (0,5%) y preguntas en un foro de Internet (0,3%).

En este sentido, los usuarios abogan por la integración de estos recursos de búsqueda en los programas de traducción asistida. Las opciones más reclamadas son la posibilidad de importar y consultar los diccionarios en CD-ROM desde el programa de traducción asistida y sobre todo la interacción con Internet: localización y descarga de textos paralelos bilingües para la integración en la memoria, así como la búsqueda de páginas web y glosarios en el par de lenguas y tema de la traducción y la posibilidad de añadirlos en una lista de favoritos en el programa de traducción asistida para acceder a ellos a partir del programa.

Las memorias de traducción en el 74% de los casos inicialmente se encuentran vacías y los traductores las llenan a medida que traducen, en el 51% de los casos se alinean traducciones anteriores con sus originales para la elaboración de una memoria, en el 44%

---

<sup>91</sup> La autora reconoce que este dato se debe a que seguramente en el momento de distribución del cuestionario no todos los traductores tenían acceso a Internet desde el trabajo.



el cliente proporciona una memoria y en el 28% la empresa de traducción cuenta con una memoria, el 15% de los usuarios descarga textos bilingües de Internet y los alinea, el 12% alinea corpus paralelos disponibles y el 6% utiliza memorias de traducción disponibles en servidores externos.

En cuanto a las sugerencias aportadas por los usuarios, consideran importante que las memorias de traducción desarrollen métodos para procesar y extraer el texto que se incluye en ilustraciones, gráficos y archivos PDF, que sean compatibles con formatos de intercambio libres como TMX<sup>92</sup>, que incluyan buenos revisores ortográficos para todas las lenguas consideradas importantes, que muestren el contexto de las coincidencias de traducción recuperadas, la posibilidad de recuperar coincidencias parciales de todos los recursos disponibles cuando la memoria no ofrece una coincidencia exacta, la posibilidad de ver todo el texto original incluso cuando se traduce segmento a segmento, la previsualización de todo el texto meta y compatibilidad con todos los idiomas.

Respecto a la preferencia de programas de traducción asistida con una sola aplicación modularizada o un paquete de aplicaciones donde cada una está destinada a la realización de diferentes tareas, el 75% de los traductores prefiere aplicaciones únicas, frente al 25% que selecciona el paquete de aplicaciones.

Uno de los comentarios de los usuarios respecto a cuál sería el programa de traducción asistida ideal refleja según la autora las ideas aportadas en el experimento: ««[An ideal TM system would be] a feature-rich, single small-scale tool that is highly-adaptable to individual translators' needs, [...] that places preference on new features that enable a translator to get started quickly on any format [and] to access additional reference materials more directly (CD dictionaries, major online dictionaries, bilingual parallel texts).»» (Lagoudaki 2006: 31).

#### 4.3.2.12 Corpas Pastor y Seghiri (2009)

Corpas Pastor y Seghiri (2009) explican qué búsquedas pueden realizarse en un corpus de documentos sobre seguros de viaje como recurso documental en traducción, en la búsqueda de terminología, fraseología, conceptos y discurso, tanto en la traducción directa como inversa de extractos de pólizas de seguros de viaje. Se seleccionan dos

---

<sup>92</sup> TMX (Translation Memory Exchange) es un estándar abierto basado en XML, que se utiliza para el intercambio libre de memorias de traducción creadas con programas TAO o de localización. Ha sido desarrollado por OSCAR, (Open Standards for Container/Content Allowing Re-use), un grupo de LISA (Localization Industry Standards Association).

fragmentos originales, uno en inglés y otro en español que deberán traducirse a español e inglés respectivamente.

En primer lugar, se refieren a la búsqueda de terminología y fraseología. El término inglés *travel insurance policy* podría equivaler a *seguro turístico* que ha tenido una gran tradición en el sistema legal español, sin embargo a partir de una directiva del Consejo de la UE que se transpuso a la legislación española se empezó a utilizar *travel assistance* que se tradujo al español con el calco *asistencia en viaje* y que ha sustituido al *seguro turístico*, a pesar de que la asistencia en viaje es solo una parte de un seguro turístico que cubre otros aspectos como cancelación del viaje o asistencia médica. En el corpus *seguro turístico* aparece en 15 concordancias y *seguro de asistencia en viaje* en 26, pero además *asistencia en viaje* produce 107 resultados.

Otro ejemplo es la búsqueda de un equivalente en inglés para *Contrato de Seguro*, ya que el traductor puede pensar en dos posibilidades: *assurance contract* o *insurance contract*. La consulta de *contract* en el corpus recupera *contract of insurance*, que además se emplea en la expresión *This is your contract of insurance*, que puede utilizarse en la traducción.

También se explica la búsqueda de un equivalente en español para la expresión inglesa *cover, conditions and exclusions*. La consulta en el corpus de *condiciones, coberturas y exclusiones* genera una sola concordancia. La consulta de *exclusiones* recupera *límites y exclusiones* y también *garantías, límites y exclusiones*. En el caso de la búsqueda de un equivalente en inglés de *condiciones generales*, la consulta en el corpus de *conditions*, obtiene como resultado *General Terms and Conditions*.

El corpus también puede emplearse para obtener información conceptual. En el ejemplo propuesto se busca un equivalente en español para el término inglés *insured person*. En la legislación española existen tres figuras que pueden investigarse en el corpus: *asegurado*, *tomador* y *beneficiario*. El *asegurado* es el sujeto del contrato, tanto en su persona (en el caso de un seguro de vida o de pensiones) o en una propiedad suya (como un seguro de hogar), el *tomador* es el que asume el seguro y paga las cuotas y no necesariamente tiene que ser el *beneficiario* que es la persona que recibirá la asistencia que cubre el seguro. Por la información encontrada en el corpus, *insured person* debería traducirse por *Asegurado* con inicial mayúscula, en los contextos *persona asegurada* aparece en 20 ocasiones, mientras que hay 5.692 y 646 concordancias de *Asegurado* y *Asegurados* respectivamente.

Por último, el corpus puede servir para obtener información sobre convenciones textuales, por ejemplo para comparar fórmulas de apertura de los contratos, como el equivalente en español de *Important* en inglés. La fórmula de apertura en español que se encuentra en el corpus es *Muy importante*. Otro problema es la búsqueda de un equivalente en inglés de *Artículo preliminar* que aparece en el inicio de los contratos españoles y contiene información sobre legislación relevante, sin embargo en inglés esta información se incluye bajo *Law applicable* y suele encontrarse al final de los contratos.

#### 4.3.2.13 Philip (2009)

Este autor trata de demostrar la utilidad de los corpus comparables generales o de referencia como recurso para el establecimiento de equivalencias en traducción. Indica que los traductores normalmente tienen intuiciones de las posibles equivalencias y que en ocasiones parten de sinónimos, cuyos significados y funciones pueden distinguirse de manera más clara en corpus que en diccionarios y glosarios.

Pone como ejemplo *go red*, cuyos posibles equivalentes pueden encontrarse inicialmente en diccionarios monolingües, bilingües y tesauros. *To go red* puede corresponder a *to become red*, *to blush*, *to flush*, *to redden*, *to turn red* y sus posibles equivalentes son *diventare rosso*, *arrossare*, *arrossarsi*, *arrossire*, *arrossirsi*, *farsi rosso* (*in viso/faccia*) y *far salire il sangue*. Las búsquedas de estos posibles equivalentes en corpus generales o de referencia en inglés (*Bank of English*) e italiano (*CORIS*) permitirán identificar las diferencias en los significados, las formas reflexivas y no reflexivas, transitivas e intransitivas, así como patrones colocacionales, funciones extralingüísticas, preferencias semánticas, contextos de uso, etc., información que según este autor es determinante para seleccionar un equivalente adecuado. Así se observa gracias a los corpus que *arrossare* corresponde a *redden*, mientras que la forma reflexiva *arrossari* sigue los mismos patrones de uso que *become red*, también se accede a construcciones fraseológicas y terminológicas como *have the grace to flush* y *degnare di arrossire* y otras expresiones equivalentes en su significado como *go red as a beetroot* y *diventare rosso come un peperone* (tener vergüenza), o *go red as a lobster* y *diventare rosso come un gambero* (tener una quemadura solar).

### 4.3.3 Estudios cualitativos sobre estrategias de búsqueda en corpus monolingües

En los apartados anteriores hemos visto que un corpus bilingüe tanto paralelo como comparable puede ser un recurso muy valioso para el traductor. Sin embargo, como sugieren también los autores, no siempre es posible disponer de un corpus bilingüe, ni paralelo, ni comparable, dada la dificultad en su compilación.

En este apartado abordamos estudios en los que se muestra la utilidad de un corpus monolingüe, sobre todo en LM, para la resolución de problemas de traducción, principalmente en lo que respecta a la selección de equivalentes de traducción.

#### 4.3.3.1 Pearson (1996a; 1996c)

Pearson (1996c) efectúa un estudio con estudiantes de traducción de la Dublin City University que están realizando un curso de terminología sobre cómo elaborar glosarios. Esta autora señala una serie de categorías de información que los estudiantes pueden consultar sobre un determinado término en un texto. Por ejemplo, los términos en los textos normalmente aparecen precedidos o seguidos por un término más general o un hiperónimo, conectados con expresiones del tipo *called, known as, is a*, por ejemplo *The Ripper is a virus*; o al contrario un término puede aparecer definido por sus hipónimos: *There are many different types of retailer, e.g. independent stores, supermarkets, department stores, etc.* Así se puede buscar a través de relaciones de descriptor-características diferenciadoras.

También existen relaciones de cuasisinonimia que pueden resultar útiles para identificar términos a partir de los cuasisinónimos conocidos por el estudiante y que suelen aparecer junto a expresiones como *i.e, e.g, known as, called, etc.*, como en el ejemplo *very recently stochastic halftoning, also called FM-screening techniques, have become available.*

Otra relación es la de objetivo o función de un término, y que se introducen en los textos con expresiones como *used for, used to, has, have, is, are, involve, produce*, como en el ejemplo *'Packaging' is defined as all products made of any material to be used for the containment, protection, handling, delivery and presentation of goods from raw materials to processed goods, from the producer to the user or the consumer.*

Según Pearson (1996c: 210-215), a partir de estas pistas, los estudiantes de traducción desarrollan métodos para resolver tres tipos de problemas principalmente.

El primer problema es identificar el significado de un término. Algunos términos específicos de una cultura que aparecen con frecuencia en textos financieros, políticos o económicos son problemáticos para el traductor porque no suelen encontrarse en diccionarios generales y especializados. Para poder identificar un equivalente en LM, el traductor debe conocer primero el significado en LO. En este caso, el traductor puede contactar con un especialista (que si es hablante de la LM difícilmente tendrá conocimientos de la LO) o con un hablante nativo de la LO (que al no ser especialista no tendrá conocimientos del campo de especialidad). El propio texto original puede dar pistas del significado de los términos y también es posible consultar otros documentos de la misma temática.

La autora pone el ejemplo de búsqueda del significado de *National Investment Savings* y *Special Savings Accounts* en los textos del CD-ROM FactFinder. Se consulta el término cuyo significado se desconoce para identificar aquellos artículos en los que se menciona. A partir de ahí, se analizan dichos textos con una herramienta de análisis de corpus llamada *MicroConcord*. *National Investment Savings* aparece en un texto definido de forma explícita, se trata de un texto escrito por un especialista para un público no especialista; en el caso del término *Special Savings Accounts* la definición no se identifica tan fácilmente, en ocasiones las definiciones se introducen de forma indirecta, en esta ocasión *Special Savings Accounts* aparece relacionado con otros términos como *DIRT*, *special investment accounts*, *(SIAs)*, *1992 Finance Bill*.

El segundo problema es encontrar un equivalente en LM. Cuando se ha encontrado un equivalente en un diccionario bilingüe, hay que verificarlo, por ejemplo en textos paralelos, del mismo campo y con la misma función comunicativa. Así, se toman los términos *neuron* (en francés *neurone*) que se analizan en el texto francés «Des neurones électroniques pour rendre les ordinateurs intelligents», publicado en *Science et vie* en abril de 1992 y el texto en inglés «How Neural Networks Learn from Experience», publicado en *Scientific America* en septiembre de 1992. Se buscan concordancias de *neuron* y *neurone* y se eliminan los resultados de *artificial neuron* y *neurone électronique* porque se consideran términos independientes. Por las definiciones y explicaciones de ambos términos en los textos, se determina que son equivalentes. Otro ejemplo propuesto es *salves d'impulsion electriques* donde los estudiantes, con ayuda del diccionario, traducen *salve* por *salvo*,

mientras que, con la ayuda de un texto paralelo, encuentran el equivalente *spikes of electrical activity*.

El tercer problema es identificar colocaciones y términos relacionados apropiados, una información que difícilmente se encuentra en los diccionarios y para la que los corpus se convierten en una fuente valiosa, ya que para conseguir una buena traducción técnica es imprescindible emplear la fraseología correcta.

En Pearson (1996a: 88-94) se proponen otros métodos para obtener información sobre una palabra en un corpus. Los recursos empleados para el análisis son la *Microsoft Encyclopaedia* en CD-ROM y las ediciones del *Financial Times* de 1992-1995, también en CD-ROM. Para la consulta de los textos se emplea el programa de análisis textual *MicroConcord*.

En una primera fase, los estudiantes de traducción tienen que analizar un texto manualmente en busca de un término determinado y tienen que identificar oraciones en las que se describe su significado, términos relacionados e información fraseológica. Entonces, los estudiantes tienen que comparar la información obtenida del texto con la que proporciona el diccionario especializado y comprueban que la información proporcionada por el texto es mucho más completa, especialmente en lo que respecta a términos relacionados e información fraseológica.

En una segunda fase, tienen que realizar el mismo proceso con ayuda de un programa de análisis textual. Se pone como ejemplo la búsqueda del término *coal*. En las concordancias resultantes podemos identificar fácilmente tipos de carbón fijándonos en las palabras que aparecen en la primera posición a la izquierda de *coal*, así encontramos los hipónimos *bituminous coal*, *subbituminous coal*, *anthracite coal*, etc. Si *coal* aparece seguido de un sustantivo o sustantivos, accedemos a términos compuestos relacionados con el carbón, por ejemplo *coal liquefaction techniques*, *coal production*, *coal combustion*, *coal utilization processes*. También puede accederse a definiciones o explicaciones de *coal*, como por ejemplo *coal: solid fuel of plant origin*; *coal and other fossil fuels*, así como a colocaciones: *produced*, *burned*, *transformed*, *cleaned*, *consumed*.

El corpus también puede emplearse para acceder a la forma desarrollada de un acrónimo, abreviatura o sigla buscándolas en el corpus directamente tal y como se haría en un diccionario, por ejemplo se puede acceder a la forma desarrollada de *TIPP* (*Taxe sur l'Importation des Produits Pétroliers*) o *CCT* (*clean coal technologies*).

Otra estrategia es la verificación de la definición proporcionada por el diccionario en un contexto de uso real, ya que el significado proporcionado por el diccionario puede haber variado con el paso de los años. Por ejemplo, la palabra *sleaze* en el diccionario *Oxford* se define como algo sórdido o de baja moral, en relación con el comercio sexual. La búsqueda de *sleaze* en el corpus de *Financial Times* revela un uso diferente al sugerido por el diccionario. Por ejemplo, *sleaze* aparece como sinónimo de corrupción en el gobierno británico, pero no tiene el componente sexual con el que contaba la palabra en su origen.

#### 4.3.3.2 Bowker (1998a; 1998b; 2000; 2003a)

Bowker (1998a) presenta un estudio en el que muestra la utilidad de un corpus monolingüe en LM para la búsqueda de terminología puntual en contextos con una temática que evoluciona rápidamente, como por ejemplo, la informática.

Efectúa una prueba con 5 estudiantes de traducción de cuarto año de la Dublin City University que cursan la asignatura de traducción especializada francés-inglés, pero no tienen conocimientos de informática. Los estudiantes conocen la herramienta *WordSmith Tools* pero nunca la han utilizado para la búsqueda de terminología en traducción. Traducen un extracto de 100 palabras de un texto real publicado sobre microprocesadores. Se pide a los estudiantes que realicen búsquedas de los siguientes términos: *carte mère*, *carte d'extension*, *bus de données*, *fréquence d'horloge* y *connecteurs pour barrettes de mémoire vive*.

El experimento consta de dos partes. En una primera fase, los estudiantes deben traducir el texto entregado en una hora y cuarto y tienen que realizar la búsqueda de la terminología con ayuda de cualquier recurso convencional que encuentren en la biblioteca: diccionarios generales y especializados, bilingües y monolingües, enciclopedias, libros, revistas especializadas y también Internet.

En una segunda fase, dos semanas después, los estudiantes deben traducir el mismo texto en 45 minutos (una hora y cuarto, menos la media hora que hubieran invertido en construir el corpus del experimento). La búsqueda puede realizarse con *WordSmith Tools* en un corpus monolingüe especializado en LM (inglés) de un millón de palabras que ha sido diseñado expresamente para este estudio. El corpus es comparable respecto al texto origen desde el punto de vista del contenido, la función y la fecha de publicación. Al final de cada prueba se pregunta a los estudiantes sobre sus impresiones acerca de los recursos utilizados y, en el caso del corpus, se analizan también las estrategias de búsqueda.

En los resultados, Bowker (1998a: 24) muestra una serie de tablas con los aciertos y los errores de cada estudiante en la propuesta de denominaciones en LM para los términos indicados, dependiendo de si la búsqueda se ha producido con un recurso convencional o con el corpus. El número total de errores es mayor con las búsquedas en recursos convencionales (22 fallos), que con las búsquedas en el corpus (7 errores). También se observa que en la mayoría de los casos, cuando se produce un error con la ayuda de un recurso convencional, se corrige cuando se busca en el corpus.

Con el fin de comprender la razón de estos resultados se pregunta a los estudiantes sobre sus actitudes y sus estrategias de búsqueda. Respecto a la búsqueda en diccionarios, los estudiantes tienen dificultades para acceder a términos compuestos completos, por lo que buscan en el diccionario bilingüe cada uno de los componentes del término compuesto y combinan los equivalentes encontrados siguiendo las reglas gramaticales de la LM. El problema es que el término propuesto no tiene por qué ser correcto. Es el caso de los ejemplos erróneos: *carte mère* (*carte* es *card* y *mère* es *mother*, por tanto se propone el equivalente erróneo *mother card*). Lo mismo sucede con *carte d'extension* (*extension card*), *bus de dones* (*information bus*), *fréquence d'horloge* (*clock frequency*).

Respecto al uso de textos paralelos (que en este experimento se conciben como documentos impresos), la mayoría de estudiantes reconocen no utilizarlos por falta de tiempo. El principal problema señalado es que los términos que se desea encontrar no aparecen destacados, ni identificados en el índice, por lo que la localización de los mismos pasa por la lectura de largos pasajes de texto. Asimismo, aunque se encuentre el término deseado en el texto paralelo, seguramente aparezcan únicamente una o dos ocurrencias del mismo, por lo que es difícil averiguar si el término seleccionado es de uso generalizado o restringido. En cuanto a Internet, un estudiante explica que los tiempos de espera y el número de documentos recuperados son excesivos.

En cambio, según los estudiantes el corpus electrónico es más actual y completo que los diccionarios, más fácil de consultar que los textos paralelos y ofrece resultados más interesantes. El corpus ofrece tanto términos simples como complejos en sus contextos, y permite ver numerosos ejemplos de uso al mismo tiempo, lo que facilita la selección de la palabra más adecuada y generalizada.

Las estrategias de búsqueda observadas por Bowker en el experimento son las siguientes. Los estudiantes suelen comenzar la búsqueda de la denominación o equivalente en LM en un diccionario bilingüe general, tratando de encontrar un equivalente o un equivalente parcial. Este equivalente o equivalente parcial se convierte en una



denominación candidata, una intuición que se comprueba en el corpus. Se introduce una denominación y se buscan sus colocaciones en el corpus.

En el ejemplo de *carte mère*, si accedemos a las colocaciones del término *card*, encontramos: *expansion card*, *fax/modem card*, *graphics card*, *network card*, *sound card*, *tuner card*, *video card*, *WinTV card*, etc. Si el estudiante intuye que alguno de los términos propuestos podría ser el adecuado, lo verificaría en contexto. Si no, efectúa otra búsqueda, en este caso accede a colocaciones del término *mother*. Entre ellas, aparece el término *motherboard*, que probablemente sea el adecuado, por lo que comprueba su uso en las concordancias, además puede acceder a información sobre sintaxis y fraseología.

Si existen varios candidatos, la búsqueda de concordancias puede ayudar al traductor a seleccionar uno. Si son términos sinónimos, según este estudio, se puede tener en cuenta la frecuencia para seleccionar el más utilizado. Por ejemplo, en las concordancias, *clock rate* y *clock speed* son sinónimos, pero si se comprueba la frecuencia, *clock speed* aparece 157 veces en el corpus y *clock rate* 49.

Asimismo, el corpus es útil para descartar intuiciones erróneas. Por ejemplo, con la ayuda de recursos convencionales, *barrette* (en *barrettes de mémoire vive*) se traduce por *strip* en la mayoría de los casos. Sin embargo, *strip* no aparece en el corpus. Según señala Bowker, con la ayuda del corpus sí que se encuentran formas mejores de expresar esta idea, aunque en este caso no todos llegan al término adecuado.

El principal inconveniente de buscar en un corpus monolingüe en LM es que el traductor tiene que partir de una información útil que pueda consultar en el corpus. Por ejemplo, uno de los estudiantes no encuentra el equivalente de *mémoire vive* en el corpus porque la búsqueda de *memory* y *live* no recupera resultados relevantes. El motivo es que la denominación adecuada en inglés son unas siglas, *RAM*.

La conclusión de Bowker en este estudio es que, aunque el corpus no sea la solución definitiva a los problemas terminológicos a los que se enfrentan los traductores, sí que es un buen recurso complementario a los recursos tradicionales, por lo que merece la pena continuar profundizando en las búsquedas que ofrece este tipo de herramienta.

En Bowker (1998b) se efectúa un estudio similar en el que los sujetos son 14 estudiantes de traducción de último año de la Dublin City University, que han cursado tres cursos de traducción y seis cursos de la LO (francés), que son nativos de la LM (inglés), con conocimientos de informática, pero no expertos en el tema concreto del texto (tecnología de escáneres ópticos), con conocimientos en el uso de herramientas de análisis de corpus pero no para búsquedas en traducción.

Se recopilan los siguientes recursos tradicionales: diccionarios bilingües generales (*Robert-Collins, Oxford-Hachette, y Harrap's*), diccionario monolingüe general (*Oxford English dictionary*) y especializados (cinco diccionarios y glosarios monolingües de informática), una enciclopedia de electrónica, un manual de usuario y un artículo de escáneres ópticos, una monografía sobre edición de textos. Además, se compila un corpus monolingüe en LM (inglés) sobre escáneres ópticos. Este corpus de 1,5 millones de palabras se obtiene de un corpus de informática más amplio en el que se han buscado textos que contengan palabras clave de la temática de los escáneres ópticos. Este corpus se consulta con *WordSmith Tools*.

Se realizan dos traducciones sobre escáneres ópticos de 182 y 301 palabras francés-inglés, 7 de los estudiantes traducen el primer texto con ayuda de los recursos tradicionales y el segundo con ayuda de los diccionarios bilingües y el corpus monolingüe; los otros 7 traducen el primer texto con ayuda de los diccionarios bilingües y el corpus monolingüe y el segundo con ayuda de los recursos tradicionales. Disponen de 2 horas para traducir cada texto y tienen que hacer comentarios sobre las búsquedas y la utilidad de los recursos monolingües proporcionados, tanto tradicionales, como el corpus.

Se analizan las traducciones atendiendo a los siguientes criterios: 1) errores de comprensión en lo que respecta a la falta de comprensión temática; 2) errores de producción en cuanto a incorrecta selección de términos, construcciones no idiomáticas, errores gramaticales y registro incorrecto.

Bowker proporciona tablas en las que contabiliza el número de errores cometidos por cada estudiante en cada una de las categorías de errores analizadas y observa que los estudiantes que han utilizado el corpus cometen menos errores.

Incluye varios ejemplos de problemas y estrategias de búsqueda, según se utilice un recurso tradicional o el corpus. El primero es un problema de comprensión del texto origen que ha provocado una mala traducción. *Sensibilité aux nuances* debería haberse traducido por *colour-recognition capability*. Aunque ningún estudiante proporciona la traducción correcta, los que utilizan el corpus se acercan más al concepto expresado en la LO. Los diccionarios especializados tratan el concepto de escáner en general, pero no del reconocimiento de colores; el manual de usuario solo hace mención a los escáneres de blanco y negro y color; la monografía sobre edición de textos y el artículo contienen discusiones sobre el reconocimiento de colores pero son difíciles de localizar porque no aparecen en el índice. En el corpus, la búsqueda de *sensitiv\** permite observar rápidamente

que aparecen palabras como *colour*, *greyscale*, *shade\** y *shading*, y al leer los contextos es posible obtener más conocimientos sobre el tema.

El segundo es un problema de selección correcta de términos en LM. *Vitre* debería haberse traducido por *glass platen* o *scan bed*. Los diccionarios especializados no contienen el término *glass platen*. Los equivalentes propuestos por el diccionario bilingüe son *glass*, *pane of glass* y *window*. *Glass platen* aparece en el manual de usuario pero no en el índice, por lo que es difícil de localizar. Ninguno de los estudiantes que utilizan estos recursos proporciona la solución correcta. En el corpus, *glass platen* aparece en 41 ocasiones y puede accederse a él fácilmente buscando colocaciones de la palabra *glass*. 3 estudiantes que utilizan el corpus dan con la solución.

Otro problema de selección correcta de términos en LM es *scanner à plat*, cuyo equivalente es *flatbed scanner*. Los estudiantes que utilizan recursos tradicionales escogen erróneamente *flat scanner*, aunque el equivalente adecuado aparece en la monografía de edición de textos. Todos los que utilizan el corpus aciertan, e incluso un estudiante comenta que ha dudado entre *flatbed* y *flat-bed*, pero que ha encontrado en la lista de palabras por frecuencia del corpus que *flatbed* aparece 1508 veces, mientras que *flat-bed* 92.

El tercer problema ejemplificado es el de las construcciones idiomáticas. El texto original contiene *photodiodes sensibles à la lumière*. Los diccionarios especializados incluyen *photodiode*, la monografía *light-sensitive elements* pero no en el índice. Los estudiantes que utilizan estos recursos tradicionales proporcionan construcciones gramaticalmente correctas pero no idiomáticas en el discurso especializado. Los estudiantes que consultan el corpus proporcionan una de las dos construcciones idiomáticas que aparecen en él (*light-sensitive photodiodes* o *photosensitive diodes*).

Otro ejemplo de construcción idiomática es *la tête de numérisation du scanner*. Todos los estudiantes que utilizan el corpus proporcionan un equivalente adecuado (*scan head*, *scanning head* o *scanner head*). Tres de los siete estudiantes que utilizan recursos tradicionales emplean la expresión no idiomática *head of the scanner*.

En las reflexiones de los estudiantes respecto de los recursos y las búsquedas, se aportan comentarios sobre la insatisfacción con los diccionarios, principalmente por no encontrar la información que se busca (Bowker 1998b: 16-18):

[...] “I didn’t find the monolingual dictionaries very useful in this situation because they didn’t really provide you with the specialized terms relating to scanners. The only information they provided was a very basic and broad definition of the concept. It never mentioned the types: hand-held, flat-bed, etc.”

[...] “I did not find the specialized dictionaries very helpful for this exercise. Some of the words I looked up were not listed (e.g. flatbed scanner) or the information was insufficient (e.g. scanner).”

[...] “I didn’t find the specialised dictionaries helpful really at all. I looked up ‘scanner’ thinking it would give the names of the different types but it didn’t.”

[...] “I did not find the technical dictionaries to be of great assistance as the terms were either very broad (dictionary of computing terms) or very technical (like the engineering dictionary). I could not find any entry referring to the different types of scanner.”

Dos estudiantes expresan su preferencia por los recursos tradicionales frente a los corpus debido al hábito de uso:

[...] “Maybe it’s force of habit, but I felt I used the dictionaries more.”

[...] “I really didn’t find the corpus tools much use; the dictionaries were much handier, quicker and easier to use. Then again, I may have been a little afraid of the corpus tools!”

Otro admite la utilidad del corpus pese a su reticencia inicial:

[...] “Although I was initially dubious of its possible use, the concordancer was in fact quite useful.” [...]

Cinco estudiantes aluden a la pérdida de tiempo en consultar el corpus, se incluye el comentario de uno de ellos:

[...] “if one was more used to the system, it hopefully wouldn’t be so time-consuming.”

Respecto a las búsquedas en el corpus, un estudiante lo considera útil para la comprensión:

[...] “I think that the concordances were useful for determining the exact sense of a particular word regarding the subject of scanners, considering I don’t know much about them.”

Cuatro estudiantes se refieren a su utilidad para seleccionar la terminología correcta, tres estudiantes mencionan la ayuda para elegir entre dos posibles equivalentes, y otros tres estudiantes para verificar que un término existe o que se utiliza de una forma determinada.

Bowker (2000: 22) trata de obtener estrategias para acceder en un corpus monolingüe en LM a términos desconocidos en LM cuando el término original no puede emplearse como punto de acceso en las búsquedas: «[...] in order to look up a term in a TL dictionary, the translator has to know what term to look for, and if the translator already knows the term, chances are there is no need to look it up. The main challenge then is to devise a method for looking up “unknown” terms in resources where the SL term cannot be used as an access point».

Según esta autora, hasta que se desarrollen herramientas específicamente diseñadas para cubrir estas necesidades de los traductores, estos deberán desarrollar metodologías en

las que apliquen cualidades como la imaginación, la flexibilidad y la paciencia, donde el corpus actúa como mediador para conseguir un fin, pero donde el último responsable de interpretar la información será el traductor.

Participan dos estudiantes de traducción de la Dublin City University, sin conocimientos en la temática del texto origen. Se compila un corpus monolingüe en LM (inglés) sobre la misma temática que el texto origen: sistemas operativos IO2 (informática). Este corpus de 150.000 palabras se compila a partir de la limitación de un corpus de informática más amplio, que ya se empleó en Bowker (1998b). El tiempo para compilar el corpus es de 45 minutos. Este corpus debe consultarse con *WordSmith Tools*.

Se realiza un experimento en el que dos estudiantes realizan la misma traducción de 180 palabras francés-inglés, es un extracto de un artículo publicado en una revista semiespecializada sobre una nueva tecnología de sistemas operativos llamada IO2. Uno de los estudiantes traduce con ayuda de recursos tradicionales (diccionarios bilingües y monolingües especializados y textos paralelos impresos), el otro estudiantes traduce con ayuda de los recursos lexicográficos tradicionales y el corpus monolingüe diseñado para el experimento.

Bowker incluye ejemplos de estrategias de búsqueda que podrían haber utilizado los estudiantes en el corpus durante el experimento.

**Identificación de términos importantes:** para identificar términos importantes en un texto paralelo, el traductor debe leerlo y posiblemente encuentre el equivalente que está buscando. En un corpus de miles de palabras no es posible leer, pero pueden emplearse otras estrategias como generar una lista de palabras frecuentes del corpus con la que el traductor detecta rápidamente las palabras más frecuentes y comunes en un campo. Según señala esta autora cuando el traductor comprende un concepto en LO pero no conoce o no es capaz de recordar la denominación más adecuada para expresarlo en LM, es más fácil seleccionarla de entre una lista de posibilidades que tratar de recordarla de memoria. Por ejemplo, se busca un equivalente del término *pilote* en LO. Se genera una lista de palabras frecuentes en el corpus monolingüe en LM, entre las que se encuentra *driver*. El traductor sabe que el equivalente de *pilote* en otros contextos es *pilot*, una persona que conduce (*drive*). Por medio de una búsqueda de concordancias de *driver*\* puede verificar si es el término adecuado en el contexto.

También es posible generar una lista de palabras clave del corpus, comparando el corpus especializado monolingüe con un corpus general (en este caso, la autora utiliza un corpus de 95 millones de palabras de artículos del periódico *The Guardian*). En la lista de

palabras clave del corpus especializado aparece *driver* en la posición 62, a pesar de que es una palabra polisémica que también puede haberse empleado con frecuencia en los textos generales.

**Investigar una intuición lingüística:** por medio de la búsqueda de concordancias, el traductor puede comprobar una intuición lingüística en el corpus. Por ejemplo, si *écrire* normalmente se traduce por *write*, se consulta *writ\** en el corpus especializado. El comodín permite recuperar distintas formas de un mismo lema. Por los contextos se comprueba que *write* puede ser el equivalente en este contexto, pero también hubiera servido para desmentir una intuición falsa. Por ejemplo, en el caso de *pilote* cuyo equivalente en lenguaje general es *pilot*. Una búsqueda de *pilot\** en el corpus monolingüe especializado no produce resultados, por lo que el traductor puede comprobar que su intuición no era adecuada y probar con otras estrategias de búsqueda.

**Investigar una intuición conceptual:** el ejemplo que pone Bowker (2000: 29) es que el traductor haya comprendido erróneamente el concepto al que se refiere *pilote* y lo haya traducido como *beta version*, refiriéndose a la versión piloto de un programa, una confusión posible porque en inglés existen los *pilot projects* en el sentido de estudios de prueba y en informática las versiones beta son versiones de prueba. Sin embargo, la palabra *beta* no aparece en ninguna ocasión a una distancia de hasta 25 palabras de algunas de las palabras clave definitorias que se han encontrado en el contexto del texto original como *IRTOS*, *write* o *API*. Para investigar el concepto al que se refiere *pilote* el estudiante busca la palabra clave *IRTOS* y en este contexto, lo que se escribe (*write*) no son *beta versions* sino *drivers*.

**Seleccionar entre varias opciones (que pueden ser variantes ortográficas, sinónimos u orden de palabras):** en cuanto a la comprobación de variantes ortográficas, por ejemplo se puede comprobar si un término *sous-système* debería escribirse con o sin guion. En el corpus *subsystem* aparece 124 veces, mientras que *sub-system* solo 2. Otro ejemplo son los acrónimos originales *Irtos* y *Raid* del texto origen, que en el corpus en LM se escriben con mayúscula en todas sus letras: *IRTOS* aparece en 22 ocasiones en el corpus, mientras que *Irtos* no aparece; *RAID* 200 veces y *Raid* solo 2.

En la elección entre sinónimos, o pseudosinónimos, pone el ejemplo del término *mémoire système* que tiene dos equivalentes sinónimos según el diccionario bilingüe *system memory* y *main memory*. En el corpus *system memory* solo aparece 5 veces, mientras que *main memory* aparece 20. Sin embargo, el corpus puede proporcionar más

información aparte de la meramente estadística ya que, según esta autora, la frecuencia de aparición no puede ser elemento definitivo único para seleccionar un sinónimo u otro.

Bowker (2000: 32) señala un criterio a tener en cuenta en el significado, a la hora de decantarse por la selección entre dos términos aparentemente sinónimos. Además del *significado denotativo*, deberán tenerse en cuenta las relaciones sintagmáticas o colocaciones. De hecho, sostiene que la sinonimia plena casi nunca existe, sustituibilidad en cualquier contexto. Por ejemplo, en la selección de un equivalente para *périphérique*, los diccionarios especializados indican que *peripheral* y *device* aluden al mismo concepto. Además, *peripheral* aparece en el corpus 92 veces y *device* 263. Si consultamos los contextos, ambos términos, empleados como sustantivos, sí que parecen intercambiables en todos los contextos pero, utilizados como adjetivos, parece que la tendencia es que *peripheral* se combina con sustantivos animados (*vendors, manufacturers*) mientras que *device* aparece como colocación de sustantivos inanimados (*module, driver*). Otra información que se obtiene de la búsqueda es que, aunque *peripheral* no aparece como adjetivo de *driver* (*peripheral driver*), el sustantivo *peripheral* sí que aparece en el mismo contexto que *driver* (*drivers for peripherals*), mientras que no hay ocurrencias de *drivers for devices*. Este tipo de información, según la autora, no aparece en los diccionarios.

En cuanto al orden de las palabras, el ejemplo proporcionado es *sous-système disque* para el cual hay que proporcionar un equivalente. El traductor tiene como posibles candidatos *subsystem disk* o *disk subsystem*, ya que en este caso no está seguro de si en este término compuesto el adjetivo inglés debería preceder al sustantivo. En el corpus aparecen ambas combinaciones: 12 de *disk subsystem* y una de *subsystem disk*. Pero si se comprueba el contexto de *subsystem disk* vemos que *disk* no actúa como sustantivo del término compuesto, sino como un adjetivo más de otro sustantivo *subsystem disk card*.

**Investigación de uso:** el corpus puede aportar numerosa información sobre el uso de las palabras. Por ejemplo, si una abreviatura en el texto original francés está en singular, aunque el término es plural, como en el caso de *des modules de services intermédiaires nommés ISM*, en el corpus monolingüe especializado el estudiante puede observar que esta abreviatura en LM inglés puede utilizarse en plural, *ISMs*.

**Búsqueda de raíz + comodín:** en algunos campos especializados como en la informática, esta estrategia puede resultar útil, ya que en algunas lenguas como el francés y el inglés, los términos equivalentes suelen compartir la raíz y difieren en los sufijos: *périphérique/peripheral, serveur/server, etc.* Por ejemplo, en la búsqueda de un equivalente para *encryptage*, el estudiante puede introducir en el corpus *encrypt\** y

recupera contextos de *encrypt* y *encryption*, que puede comprobar en contexto para seleccionar el término adecuado. Además, esta estrategia de búsqueda permite al traductor acceder a información sobre patrones de uso diferentes en lenguas distintas, uso de categorías gramaticales diferentes. Así, en este mismo ejemplo, vemos por los contextos que el concepto que envuelve al sustantivo *encryptage* puede expresarse de una forma más natural en inglés utilizando un verbo *encrypt*.

**Colocaciones o agrupaciones de palabras de términos conocidos o de partes de términos:** cuando el traductor desconoce un término puede buscar a partir de términos que sí que conoce y que deberían aparecer en una posición cercana al término que desconoce. Por ejemplo, en el término *systeme d'exploitation*, el traductor puede intuir que *systeme* será equivalente a *system*. Una estrategia sería observar qué palabras coaparecen con *system* en el corpus. *Operating* aparece 241 veces en la primera posición a la izquierda de *system*.

Otra estrategia podría consistir en la generación de una lista agrupaciones de palabras (en este caso dos palabras) que contengan *system*. El primer término de la lista es *operating system*. En ambos casos, pueden comprobarse los contextos de *operating system* para confirmar si es el término adecuado.

**Estructuras sintácticas:** existen estructuras sintácticas en la LO que tienen una estructura predecible en otras lenguas. Por ejemplo, para encontrar el equivalente en inglés de *transaction point à point*, la estructura francesa X à X suele equivaler en inglés a X to X. Si buscamos en el corpus el patrón *-to-*, accedemos a expresiones en inglés como *peer-to-peer*. La búsqueda posterior de *peer-to-peer* en el corpus nos permite comprobar que se combina con *transactions*.

**Abreviaturas:** este método funciona cuando la LM es el inglés. En ocasiones la forma expandida de un término en una lengua se abrevia con sus siglas en inglés. Para acceder a la forma expandida en inglés del término basta con introducir las siglas que encontramos del texto original en un corpus de la temática en inglés. Por ejemplo, en la búsqueda de un equivalente para *API (interface de programmation d'application)* las siglas utilizadas en el texto original francés sugieren que se ha utilizado la abreviatura en inglés, de lo contrario se hubiera empleado *IPA*. Por tanto, para acceder a la forma desarrollada del término en inglés, buscamos la abreviatura *API* en el corpus monolingüe en lengua inglesa, con lo que encuentran contextos de *application programming interface (API)*.



**Nombres de productos:** un término cuyo equivalente se desconoce aparece en una posición cercana al nombre propio de un producto; por ejemplo en informática, al lado de la palabra *Windows*, podemos acceder al equivalente a partir de buscar el nombre del producto en el corpus en LM, ya que probablemente en la LM el equivalente que buscamos también aparecerá como colocación del nombre propio, que no varía entre las dos lenguas.

**Número:** se puede buscar a partir de números que aparecen en una posición cercana al término en LO del cual desconocemos el equivalente en LM, ya que los números también son invariables entre lenguas. En ocasiones las cantidades varían de unas lenguas a otras, por lo que la búsqueda de un número exacto puede no ofrecer resultados, pero es posible buscar números parciales que sí ofrecen resultados. Por ejemplo, en el texto original aparece *1200 ppp*, la búsqueda de *ppp* y *1200* no ofrece resultados en el corpus monolingüe inglés. La introducción de un número con un comodín \*0 ofrece, entre otras, concordancias de *300 dpi scanner*, *4000 dpi scanner*, *400 dpi colour scanner*. Una búsqueda de *dpi* permite descubrir que se refiere a *dots per inch*, equivalente a *point par pouce*.

Bowker (2003a) presenta otro experimento con una metodología muy similar a la de los estudios anteriores. Divide a estudiantes de traducción en dos grupos, unos tienen que efectuar una traducción con la ayuda de recursos tradicionales como los diccionarios y otros la misma traducción apoyándose en un corpus. Se aportan más ejemplos de búsqueda.

En el primer ejemplo, se busca un equivalente en inglés para el término francés *numériseur*. Una consulta en diccionarios, incluidos los más actuales, sugiere *digitizer*. Sin embargo, los estudiantes que utilizan el corpus, emplean *scanner*, que es el término que usan actualmente los expertos en el tema.

En la traducción de una receta del alemán al inglés, con el uso de recursos convencionales, los estudiantes traducen los artículos definidos, por ejemplo emplean *Season the trout* como equivalente de *Die Bachforellenfilets würzen* y *Mix in the Butter* para *Die Butter untermischen*. Sin embargo, los estudiantes que emplean el corpus omiten los artículos definidos y utilizan las frases *Season trout*, *Add butter*.

En la traducción de construcciones pasivas, los estudiantes que consultan diccionarios mantienen las mismas construcciones pasivas del alemán en el inglés, *Die fertigen Röllchen werden auf einer Platte angerichtet*, *The finished roll is arranged on a plate*, mientras que los estudiantes que buscan en el corpus utilizan el imperativo en voz activa *Arrange finished roll on a plate*.

Los corpus también resultan más útiles en la búsqueda de equivalentes de términos compuestos, por ejemplo *la tête de numérisation du scanner* en francés se traduce al inglés por *the head of the scanner* por parte de los estudiantes que buscan en el diccionario, y por *scan head*, el término adecuado en el lenguaje de especialidad, por aquellos que consultan el corpus.

En ocasiones la frecuencia de uso en un corpus puede ayudar a decidirse por un equivalente, como en el caso de *Gewähr* cuyo equivalente propuesto en los diccionarios es *guarantee*, mientras que una búsqueda en el corpus recupera 35 resultados de *guarantee* y 475 de *warranty*; de la misma forma en la traducción de *Gewährleistungszeit*, hay 28 concordancias de *warranty period* y 2 de *guarantee period*, y para *Garantieleistungen* se producen 6 concordancias de *warranty service* y solo una de *guarantee service*.

#### 4.3.3.3 Friedbichler y Friedbichler (2000)

Los autores efectúan un estudio piloto en el que participan 24 estudiantes de tercer y cuarto curso de Innsbruck University (IÜD) en la realización de un curso de traducción médica.

Se les pide que efectúen tareas de traducción por parejas utilizando un diccionario electrónico de medicina, un corpus bilingüe comparable de textos de neurología compilado de Internet en el semestre anterior y también la Web con la ayuda del buscador *Netscape*. Al finalizar las tareas, se pide a los estudiantes que evalúen los recursos en función de su utilidad. En primera posición sitúan a los corpus como la herramienta más útil, después Internet para consultas específicas, pero señalan el tiempo invertido en las búsquedas; también el diccionario electrónico resulta útil especialmente la búsqueda en todo el contenido del diccionario, y respecto a los diccionarios convencionales solo se señala la ventaja de la familiaridad frente a los otros recursos.

Estos autores explican que las herramientas de análisis de corpus están diseñadas para lingüistas, lexicógrafos y profesores de lengua e indican qué características podrían interesar a los traductores. A diferencia de los lingüistas y profesores de lengua que seguramente precisan analizar y evaluar los resultados durante horas, al traductor profesional le interesa encontrar rápidamente la palabra o expresión más adecuada. Por eso, un requisito de estos programas es que sean rápidos. Otra característica deseable sería que el programa permitiera cambiar rápidamente la búsqueda en corpus diferentes, ya que al traductor puede interesarle buscar en más de un recurso. También es necesario que el programa sea intuitivo y fácil de aprender a usar, por lo que la interfaz de búsqueda debe

ser simple en la que se pueda indicar en un mismo cuadro de diálogo tanto la consulta como el corpus seleccionado. Es necesario que se proporcionen funciones de búsqueda variadas (comodines que sustituyan palabras o caracteres de una palabra, posibilidad de distinguir o no de mayúsculas y minúsculas, búsqueda de varias palabras en un mismo contexto, etc.).

#### 4.3.3.4 Stewart (2000)

Desde su experiencia como profesor de traducción, pretende demostrar la utilidad de un corpus general monolingüe en LM en el proceso de traducción inversa de textos generales, folletos turísticos del italiano al inglés. Como resultados proporciona ejemplos de búsqueda en el corpus monolingüe *BNC* para el acceso a equivalentes en inglés de términos originales en italiano a partir de colocaciones y comprobación de nombres propios.

El primer ejemplo es *strada panoramica* en italiano, que el traductor podría traducir literalmente al inglés como *panoramic road*. En el diccionario bilingüe, el equivalente *panoramic* incluye una colocación *panoramic view*, pero el traductor necesitaría saber si *road* o algún sinónimo puede combinarse con *panoramic*. En el corpus *BNC*, la búsqueda de una lista de colocaciones de *panoramic* muestra que casi nunca se combina con *road* o sinónimos. Sí que se detectan 75 resultados de *view* y 37 de *views* en plural.

En el segundo ejemplo, se precisa la traducción de *Gran Giro de la Città*, que puede traducirse literalmente como *big tour of the city*. El diccionario no ayuda al traductor a averiguar si el equivalente idóneo en este contexto para *gran* es *big*, *great*, *grand* o *large*. La búsqueda de una lista de colocaciones de *tour* y *tours* en el *BNC* muestra que se combina con cualquiera de los posibles adjetivos, por lo que el traductor consulta las concordancias del corpus en las que aparecen dichas colocaciones. Por ejemplo, concordancias de *grand* con *tour(s)* en un contexto de 5 palabras de distancia. El 30% de los resultados resultan irrelevantes y, de los contextos relevantes, la mayoría se encuentran en un contexto de tours de música, teatro o deporte y no parecen apropiados en itinerarios turísticos, que es el tema del texto original. En el 50% de los casos sí que se emplea *grand* con *tour(s)*, con iniciales mayúsculas, lo que le confiere un estatus de nombre propio, como en la asociación histórica al *Grand Tour of Europe*, además todos estos resultados aparecen en contextos turísticos. En el caso de *great* y *tour(s)* se muestra que la combinatoria no es la que se busca, al igual que *large* y *tour(s)*.

En el tercer ejemplo se muestra la búsqueda de nombres propios, en este caso de la artista *Palma il Giovane* que puede traducirse literalmente como *Palma the Young*, *Palma the Younger*, *Pliny the Younger* o puede dejarse en italiano. La búsqueda de *Palma* en el corpus *BNC* recupera dos concordancias de *Palma Giovane* y una de *Palma il Giovane*. Un ejemplo similar es el de la traducción del título de película *Via col vento* que podría traducirse como *Gone with the wind*. La mayoría de traductores podrían estar seguros de que *with the wind* es correcto pero no recordar el resto del título *off*, *away*, *went*, *gone*. La búsqueda de la siguiente secuencia de palabras con un comodín al principio \* *with the wind*, recupera 20 concordancias del título *Gone with the wind*.

#### 4.3.3.5 Bowker y Pearson (2002)

Bowker y Pearson (2002: 198) consideran que el corpus monolingüe también puede emplearse como recurso de traducción. Según estas autoras este recurso puede emplearse como alternativa al corpus paralelo, ya que los corpus monolingües son más fáciles de compilar que los paralelos. Aunque estas autoras indican que la consulta en corpus monolingües requiere estrategias de búsqueda más ambiciosas para extraer la información útil para la traducción; se precisa creatividad, flexibilidad y paciencia.

Para ilustrar sus estrategias de búsqueda, Bowker y Pearson emplean un corpus monolingüe en inglés de *Computer Select* en CD-ROM, con un total de 29.589 tokens de artículos sobre virus informáticos.

La primera estrategia que proponen consiste en verificar una intuición. El traductor a veces cree que conoce el equivalente adecuado pero quiere verificarlo en otro recurso. Por ejemplo, el término francés *cheval de Troie* puede equivaler en inglés a *Trojan horse*, las concordancias en el corpus sobre virus informáticos demuestran que la intuición es correcta.

El término *virus dans la nature* puede equivaler a *natural virus* o *virus found in the nature*. La búsqueda de *virus* ofrece 600 resultados, lo que complica su lectura. La consulta de *natur\** proporciona 5 resultados, de los cuales en ninguno se emplea en un contexto para describir a un virus. En este caso la intuición ha fallado. Sin embargo, desmentir una intuición siempre es un paso hacia la solución adecuada. El problema es encontrar el término adecuado en un corpus monolingüe cuando no sabemos realmente lo que estamos buscando.

Otra estrategia a la que se refieren las autoras es la búsqueda de equivalentes desconocidos. Siguiendo el ejemplo de *virus dans la nature*, pueden emplearse técnicas más creativas para localizar equivalentes. Por ejemplo, puede buscarse el patrón *in the*, que aparece 77 veces, pero puede combinarse con la búsqueda en contexto de *virus\**, que ofrece un total de 19 concordancias, también puede agilizarse la búsqueda ordenando las concordancias por la primera palabra a la derecha de la palabra de búsqueda. Encontramos, *virus in the wild* entre otros. También es posible generar una lista de agrupaciones de palabras que contenga *in the*, en la que aparece *in the wild*.

Otro ejemplo es la búsqueda de un equivalente para *les virus furtifs et semi-furtifs*. La búsqueda de *furtive* no ofrece resultados. La consulta de *semi\** incluye en sus resultados *semi-stealth virus* que parece ser el equivalente adecuado. La búsqueda de *stealth* ofrece 8 concordancias de *stealth virus*.

El conocimiento de los patrones estructurales de una lengua puede ser útil, por ejemplo en la búsqueda de un equivalente para *réseau poste à poste*. El equivalente de *réseau* es *network*. La consulta de *network* en el corpus ofrece 80 concordancias. La consulta de *post\** no tiene resultados. La búsqueda de *-to-* a cinco posiciones de distancia de *network* incluye *easy-to-create*, *easy-to-install*, *PC-to-Mac*, *up-to-date*, *host-to-host*, *peer-to-peer* y *point-to-point*. *Peer-to-peer* parece el equivalente adecuado. La búsqueda de *peer-to-peer network* en un contexto más amplio indica que es el equivalente adecuado.

Un corpus monolingüe también puede servir para elegir entre distintas opciones (Bowker y Pearson 2002: 202-203), por ejemplo, para decidir si utilizar un término con o sin guion (*anti-virus*, *antivirus*). Una lista de palabras frecuentes del corpus revela que *anti-virus* aparece en 21 ocasiones, mientras que *antivirus* se repite 209 veces.

Esta estrategia también puede emplearse para investigar la sinonimia. Por ejemplo, el diccionario bilingüe indica que *se propager* puede traducirse por *propagate* o *spread*, pero no se indica cuál es el más adecuado en el contexto de los virus informáticos. En el corpus *spread* aparece 46 veces, mientras que *propagate* solo 4. Aunque la estadística no puede ser el único factor para decidirse por un sinónimo u otro, puede ser de ayuda.

Los corpus monolingües también pueden servir para investigar el uso. El traductor puede haber averiguado que el equivalente de *logiciel antivirus* es *antivirus software*, pero desconocer cuáles son sus colocaciones. Si se ordenan las concordancias por la primera posición a la derecha, puede identificarse que los verbos que se combinan con *antivirus software* son *protect*, *detect*, *scan*, *check*. *Detect* y *scan* aparecen tanto en presente como en pasado y *protect* solo en presente. A la derecha de los verbos encontramos qué tipo de

cosas se ven influenciadas por los antivirus: el verbo *detect* se combina con *virus*, *scan* con *file*, *protect* con *computer*, *you* y *organization*.

También mencionan estrategias para investigar el estilo. En el caso de la oración *Si l'on ne supprime pas le virus, le virus effacera les fichiers*, el pronombre *on* en francés es impersonal. En inglés el impersonal puede construirse con *one*, *you* o una construcción pasiva. De las 69 concordancias de *one*, en la mayoría no se emplea como pronombre personal, sino como numeral. La consulta de *you*\* revela que *you* se emplea en estos textos como construcción impersonal. Asimismo, también puede observarse que las formas contraídas *can't*, *don't*, *haven't* y *you're* son muy frecuentes en este tipo de textos, por lo que pueden emplearse en la traducción.

El corpus monolingüe puede servir para buscar contextos explicativos (Bowker y Pearson 2002: 206-207). Un traductor no puede recibir formación en todas las áreas de especialidad que necesita traducir: derecho, medicina, informática. Un corpus puede ayudarlo a conocer conceptos determinados y ser más preciso conceptualmente en la traducción. La búsqueda de *virus*\* ofrece 600 concordancias, por lo que es difícil leerlas todas en la búsqueda de contextos definatorios o explicativos. La solución según estas autoras es buscar contextos ricos en conocimiento. Hasta que se desarrollen herramientas que recuperen automáticamente este tipo de contextos, el traductor puede emplear estrategias para encontrarlos, por ejemplo la búsqueda de patrones léxicos específicos: *is an*, *is the*, *is one*, *kind of*, *sort of*, *type of*; o patrones que aparecen cuando un término se define por sus partes *has a*, *part of*, *made up of*, *consists of*, o por su función *used to*, *used for*, *employed as*.

Bowker y Pearson (2002: 188-189) también se refieren a la búsqueda en un corpus etiquetado como guía para escribir, ya que puede consultarse a partir de clases de palabras en lugar de palabras concretas. Por ejemplo, podemos saber qué adjetivos describen a *printer* si se busca *printer* en combinación con la categoría gramatical adjetivo en una posición inmediatamente a la izquierda. Encontramos adjetivos como *big printer*, *cheap printer*, *colour inkjet printer*, *highly usable printer*, *large printer*, *lightweight printer*, *monochrome printer*, *new useful printer*, *personal printer*, *portable printer*, *quick printer*, *robust printer*, *serious printer*, *solid*, *plucky*, *little printer*, *sporty little printer*.

Otro ejemplo es generar una lista de categorías gramaticales del corpus, en la que se observa que los verbos en tercera persona del singular son muy frecuentes. Si se hace una consulta en el corpus con dicha categoría gramatical, encontramos los verbos *delivers*, *offers*, *comes in/with*, *includes*, *provides*, *uses*, *supports*. En la búsqueda de *performance*

precedido de la categoría gramatical adjetivo, encontramos los adjetivos *fast, good, impressive, solid, strong, below-average, slow, sluggish, unspectacular*.

#### 4.3.3.6 Coffey (2002)

Coffey (2002) presenta un estudio en el que muestra la utilidad de un corpus monolingüe en LO para el traductor como recurso de búsqueda. Este autor reconoce, sin embargo, que como ayuda en traducción pueden resultar más útiles otros corpus como los comparables, paralelos y monolingües en LM.

El corpus monolingüe en LO es una fuente de información lingüística y enciclopédica, que puede proporcionar información sobre determinados aspectos del texto original, sobre todo en casos en los que el traductor quiere comprobar si un determinado aspecto de la lengua que se refleja en el texto original es típico de dicha lengua o del tipo textual. Esta información puede ser especialmente útil si la lengua del texto original no es la lengua materna del traductor.

#### 4.3.3.7 Wilkinson (2005a; 2005b; 2006a; 2006b; 2007; 2010)

Wilkinson (2005b) compila un corpus monolingüe en LM y no etiquetado para la traducción de textos turísticos del finlandés al inglés. El corpus es de 670.000 palabras y contiene folletos turísticos en inglés de Reino Unido, Estados Unidos y Canadá, que se extraen de Internet, la mayoría en formato PDF, y que se convierten a texto plano.

Para este autor la mayoría de corpus compilados disponibles en línea, por ejemplo el *BNC*, aunque seguramente resulten de utilidad para la compilación de diccionarios y la investigación lingüística, no son demasiado útiles para los traductores, bien porque son demasiado generales, o bien porque están desfasados, muchos están compuestos además por textos del lenguaje oral o histórico, y no pueden tomarse como referente al traducir lenguaje escrito actual.

Otro problema que señala es que algunos corpus no se encuentran disponibles al público y, si lo están, el precio por su uso es muy elevado, hay que pagar una suscripción o comprar el CD-ROM (Wilkinson 2006a). Por otro lado, este autor indica que no existen corpus especializados, ni gratuitos, ni comerciales, por lo que los traductores deberían ser capaces de compilar sus propios corpus<sup>93</sup>.

---

<sup>93</sup> En Wilkinson (2010) incluye un tutorial sobre cómo utilizar el buscador de WebAsCorpus.org para compilar un corpus especializado, mediante la inclusión y exclusión de palabras de búsqueda y, una vez

No obstante, reconoce que, a pesar de las ventajas, compilar corpus puede ser una tarea pesada, que seguramente no sea productiva para su uso en un único encargo de traducción, o en traducciones especializadas cortas, aunque por otro lado indica que la compilación de un corpus en LM puede mejorar la productividad y calidad de las traducciones, sobre todo si siempre se traduce un mismo tipo de textos (Wilkinson 2006a).

Wilkinson (2006b) indica, además, que los traductores profesionales no están familiarizados con las herramientas de gestión de corpus, aunque el uso de estas herramientas permitan la consulta de textos en LM para acceder a la terminología y la fraseología de una forma más rápida y sistemática que en la consulta de textos impresos o en línea. Este autor supone que la razón por la que los traductores no utilizan los corpus es porque este tipo de recursos no han sido pensados para el traductor (aunque cada vez las interfaces son más intuitivas), pero sobre todo porque es una herramienta que no se incorpora en las clases de traducción (solo se tratan los corpus en cursos de herramientas TAO o lingüística de corpus), por lo que los traductores no la integran como un recurso de búsqueda más.

Durante el curso 2004-2005 utiliza el corpus turístico que ha compilado en sus clases de traducción en el centro Savonlinna School of Translation Studies (Universidad de Joensuu), donde enseña a los estudiantes a utilizar *WordSmith Tools* y algunas estrategias de explotación de corpus, y les asigna tareas de traducción de textos turísticos del finlandés al inglés. A partir de esta experiencia, el autor proporciona ejemplos de cómo los estudiantes supuestamente han tenido que utilizar el corpus para la búsqueda de terminología en sus traducciones (Wilkinson 2005b).

Se refiere a la búsqueda de colocaciones especialmente a partir de la generación de concordancias KWIC. Por ejemplo, como en el siguiente caso en el que se pretende encontrar adjetivos que son colocaciones de sustantivos. Al traducir una frase en la que aparece el término *rapids*, en el corpus pueden encontrarse una serie de adjetivos que se pueden emplear en combinación con dicho sustantivo: *boiling rapids*, *churning rapids*, *dangerous rapids*, *frothy rapids*, *gurgling rapids*, etc.

Se propone otro ejemplo de búsqueda de colocaciones para el acceso a un equivalente en inglés de los adjetivos *hoidettu* o *kunnostettu*, que se emplean para caracterizar a los caminos utilizados en el esquí de fondo (*cross-country ski trails*). Según



este autor, la búsqueda en recursos tradicionales sugiere la posibilidad de los adjetivos *conditioned*, *maintained*, *restored* o *reconditioned*. Sin embargo, de las más de 1000 concordancias de la palabra *trail* en el corpus, ninguna aparece en combinación con estos adjetivos. Sí que se producen 40 concordancias en las que el adjetivo *groomed* está situado inmediatamente a la izquierda de *trail*. Si se busca el adjetivo *groomed*, aparecen 128 concordancias, entre las que se encuentran: *groomed bicycle and walking trails*, *groomed classic and skating trails*, *groomed cross-country ski trails*, *groomed fairways*, *groomed off-road trails*, *groomed runs*, *groomed slopes* y *groomed wilderness trails*.

Wilkinson (2005b) también sugiere que una forma de aligerar la búsqueda (ya que es difícil identificar 40 apariciones del adjetivo *groomed* de entre un total de 1000 concordancias) es la generación de listas de colocaciones (por ejemplo, en este caso las palabras que aparecen con mayor frecuencia 5 posiciones a la izquierda de *trail*) o agrupaciones de palabras (las agrupaciones más frecuentes de 3 palabras que contienen el término *trail*). Con cualquiera de estas dos búsquedas, habríamos accedido en el corpus de una manera más rápida al adjetivo *groomed*.

Wilkinson (2005b) también propone el uso del corpus como medio de verificación de términos que se han encontrado previamente en otros recursos, como diccionarios o Internet, o que proceden de la propia intuición del traductor. Por ejemplo, en la búsqueda de un equivalente para *koiravaljakkoajelu*, los estudiantes acceden en los recursos tradicionales a los equivalentes *dog sled*, *dog sledge* y *dog sleigh* (cada una de ellas con variantes escritas también con guion o como una única palabra). Según este autor, el corpus puede servir en estos casos para decantarse por una u otra opción. Por ejemplo, una búsqueda del patrón *dog\** en el corpus, nos lleva a 22 concordancias de *dog sled*, 27 de *dogsled* y 6 de *dog-sled*. También hay 68 resultados de *dogsledding* (escrito también en dos palabras). Finalmente, las concordancias muestran que *dog sled* aparece como colocación de los términos *adventure*, *excursion*, *ride*, *trip* y *tour*. No se producen resultados de *dog sledge* y *dog sleigh*, ni de sus variantes.

Wilkinson (2005b) señala que en ocasiones los resultados se encuentran por casualidad. En la búsqueda de palabras contextualmente relevantes puedes conseguir información de forma accidental, por ejemplo como en el caso del adjetivo *groomed* para calificar a *trail*. En este sentido, una búsqueda en el corpus puede proporcionar *pistas*, que pueden llevar al traductor a profundizar en la búsqueda de una palabra, buscándola en un contexto más amplio, o a buscar nuevos patrones de búsqueda (Wilkinson 2007: 112). Los

resultados pueden ser útiles para la traducción actual, o también para otras traducciones, por ejemplo la búsqueda de *dog\** da como resultado *dog sled*, pero también *dog musher*.

Según este autor (Wilkinson 2005b), un corpus también puede servir al estudiante para encontrar fraseología (*language chunks*), que ayuda a conseguir naturalidad en su discurso. El ejemplo que pone es que en el término *ruska-aika* (periodo del otoño en el que las hojas cambian de color y el paisaje es muy bonito), la palabra *ruska* tiene información implícita que será necesario explicitar en inglés. Si el traductor busca en el corpus la palabra *autumn* o *fall* puede acceder a palabras y locuciones que le pueden servir para describir esta época del año, como *fall foliage season*, *brilliant foliage in fall*, *stunning fall foliage*.

En un estudio posterior, Wilkinson (2005a) explica que la mayoría de las búsquedas a las que se alude anteriormente se producen cuando el usuario tiene una idea sobre aquello que espera encontrar en el corpus. Sin embargo, la experiencia de este autor es que a algunos estudiantes les sucede que no saben cómo encontrar lo que están buscando, o incluso que no saben lo que tienen que buscar: «[...] “I don't know how to find what I'm looking for” and even “I don't know what to look for”». En un artículo posterior, Wilkinson (2007: 108) se refiere a estas dos situaciones como búsqueda de *known unknowns* y de *unknown unknowns*.

Therefore it is necessary to develop other strategies for discovering “unknowns”, or at least “lesser knowns” and “partially knowns”. As Varantola (2002, p.180) points out, search strategies must sometimes be elaborate, and if no adequate search string or term springs to mind, translators need to think of indirect ways of finding what they are looking for.

(Wilkinson 2007: 108)

En Wilkinson (2005a) se denomina a este tipo de búsqueda *fuzzy searching*, búsquedas creativas para localizar términos desconocidos (*unknown terms*), un tipo de búsqueda a la que ya se habían referido antes otras autoras como Bowker (2000: 37) y Bowker y Pearson (2002: 200), que también hemos tratado en este apartado.

El autor propone dos ejemplos de búsqueda propios a modo de Thinking-Aloud Protocol, pero que supuestamente están basados en las discusiones de estudiantes en clase respecto a las estrategias de búsqueda empleadas.

En el primer ejemplo, el estudiante de traducción quiere encontrar un equivalente en inglés del término finlandés *omatoimisilla*. En el diccionario bilingüe, el único equivalente proporcionado es *independent*. El estudiante decide comprobar este equivalente en el corpus. La búsqueda de *independent* en el corpus no ofrece los resultados esperados. En el texto origen, *omatoiminen* se emplea en oposición (relación de antonimia) a *opastettu*

(*guided*), por lo que se introducen en el corpus consultas como: *guided and* y *guided or*, para comprobar con qué términos se contrasta el adjetivo *guided*. Se encuentran muchos resultados de *self-guided*, algunos de *unguided*, un resultado de *independent* y un resultado de *individual trips*. El estudiante decide entonces buscar *self-guided* y *unguided*. Se encuentran 56 resultados, de los cuales solo 5 corresponden a *unguided* y todos en textos canadienses. *Self-guided* aparece tanto con guion (en 40 casos) como sin él. Por tanto, el estudiante decide finalmente seleccionar *self-guided* como equivalente.

En el segundo ejemplo, el estudiante pretende encontrar un equivalente para *mönkijäsafareita*, un término compuesto que contiene la palabra *mönkijä*, que el estudiante no encuentra en el diccionario bilingüe ni en ningún glosario. Como es una palabra que debería aparecer en combinación con *safaris*, el estudiante decide comprobar qué colocaciones del término *safari(s)* aparecen en el corpus. Obtiene más de 100 resultados, entre los que se encuentran *photo safaris* y *wildlife safaris*, bastantes de *quad bike* y *quad safaris*, pero solo en folletos turísticos de inglés británico, también encuentra *ATV safari(s)*. Decide buscar *quad* y encuentra muchos resultados de *quad bikes* y *quad biking*, en folletos turísticos de inglés británico; también hay muchas concordancias de *quad* en textos canadienses, pero utilizado como adjetivo que precede a *chairs* y *chairlifts*. Decide probar con la búsqueda de *ATV*. Obtiene 60 resultados, de los que solo uno es de un texto en inglés británico, pero se utiliza con frecuencia en textos de Canadá y Estados Unidos. Se puede ver que *ATV* es una abreviatura de *All Terrain Vehicle*. Entonces, decide comprobar *ATV* y *quad bike* en la enciclopedia en línea *Wikipedia*. *ATV* es un término general que designa a vehículos todoterreno y los que son de cuatro ruedas normalmente se llaman *quad bike*. Al buscar *quad bike* en Internet, obtiene resultados de Australia, Nueva Zelanda y Reino Unido y, si restringe la búsqueda a páginas de Canadá «site:ca» y Estados Unidos «site:us», también encuentra resultados, pero pocos. Decide no utilizar *quad bike* porque el folleto que está traduciendo va dirigido a un público internacional. Restringe la búsqueda de *All Terrain Vehicle* a páginas del Reino Unido «site:.uk» y obtiene más de 10.000 resultados, por lo que decide que este término es más conocido en distintos países. Además, los vehículos que aparecen en la imagen de *Wikipedia* son iguales que los de la imagen del folleto que se está traduciendo, por lo que selecciona *ATV safaris*.

Finalmente, en este artículo, Wilkinson (2005a) se refiere a la búsqueda avanzada, que define como una búsqueda de concordancias con palabras relevantes del contexto, que

pueden aparecer o no a una distancia determinada de la palabra de búsqueda. En Wilkinson (2007: 113-120) proporciona los siguientes tres ejemplos.

En el primer ejemplo, el traductor quiere encontrar un equivalente en inglés del verbo *pilkkiä* o del sustantivo *pilkkiminen* en finlandés, que significan «pescar a través de un agujero en el hielo». El diccionario bilingüe sugiere *jig*, *jigger* o *jigging*. En el diccionario monolingüe y en Internet, *jigging* se refiere al acto de introducir un cebo en el agua, pero no contiene el matiz de que se haga en invierno ni en el hielo. Una traducción explicativa como «fishing with a jig through a hole cut in the ice» resultaría poco factible, especialmente si la idea se repite con frecuencia en el texto. Para acceder al término adecuado, el estudiante de traducción puede buscar, en el corpus de turismo, la palabra *fishing* a cinco posiciones de distancia a derecha o izquierda de *winter*. Se encuentran bastantes resultados del término *ice fishing*. La búsqueda de *ice fishing* y *ice-fishing* proporciona más resultados que pueden verse en un contexto más amplio. La búsqueda en otros recursos de la definición de *ice fishing*, como por ejemplo en la *Wikipedia*, demuestra que ese es el equivalente adecuado.

En el segundo ejemplo, se busca un equivalente en inglés del término en finlandés *hevosrekiäjelu*, que aparece en la siguiente frase: *Koiravaljakko- ja hevosrekiäjelut tilauksesta* (viajes en trineos de perros y trineos de caballos con reserva previa). En búsquedas de un estudio anterior, veíamos que los equivalentes de *Koiravaljakko* podían ser *dogsled*, *dog-sled* o *dog sled*, pero el traductor desconoce si puede utilizarse también *horse sled*. La búsqueda de *horse\** en el corpus de turismo genera demasiadas concordancias para comprobar, pero si se ordenan por la palabra central de búsqueda, se puede ver fácilmente que no hay resultados de *horsesled*, *horsesledge* o *horsesleigh* (con o sin guiones), y si se ordena por la primera palabra a la derecha, también se identifica que no hay resultados de *horse sled*, *horse sledge* o *horse sleigh*, aunque sí que se encuentran resultados de estos términos en Internet, especialmente de *horse sleigh rides*. Entonces, el traductor busca de nuevo en el corpus *horse\** con *sle\** en contexto y encuentra numerosos resultados de *horse-drawn sleigh ride*, término que también aparece con mucha frecuencia en las páginas de Internet.

El tercer tipo de búsqueda tiene como objetivo distinguir entre el uso de distintos sinónimos. El traductor genera una lista de palabras ordenadas alfabéticamente en el corpus; puede ver que *sled* aparece 85 veces en el corpus y en diferentes textos, *sleds* en 21 ocasiones, *sleigh* y *sleighs* 34 y *sledge* y *sledges* 2. El autor señala que solo la estadística no es determinante para seleccionar un sinónimo u otro. Sin embargo, una búsqueda de

listas de concordancias de estos sinónimos sí que puede ayudarnos a detectar rápidamente con qué palabras se combinan en contexto y, por tanto, a seleccionar el más adecuado. Por ejemplo, la lista de concordancias de *sled* o *sleds* muestra que su colocación más frecuente es *dog*, que aparece tanto a derecha como a izquierda y, sin embargo, *horse* no aparece como colocación. En el caso de *sleigh* o *sleighs* aparecen colocaciones de *horse*, pero no de *dog*.

#### 4.3.3.8 Montero Martínez y Faber (2008)

Montero Martínez y Faber (2008) aportan ejemplos propios de cómo puede utilizarse un corpus ad hoc monolingüe en LM consultado con la ayuda de *WordSmith Tools* para la producción textual en la LM en una traducción inversa español-inglés de un artículo especializado sobre oncología.

Para la búsqueda de equivalentes de términos simples en LM, estas autoras generan listas de palabras ordenadas alfabéticamente, tanto del texto origen como del corpus en LM. En este campo la comparación de ambas listas es útil para encontrar terminología equivalente ya que existen muchos términos en ambas lenguas que comparten la misma raíz o incluso calcos o cognados, como *adyuvant* y *adyuvante*, *adrenalectomy* y *adrenalectomía*; también es útil para encontrar abreviaturas y acrónimos.

Para la búsqueda de equivalentes de términos compuestos emplean concordancias, por ejemplo para acceder al término compuesto *sarcoma indiferenciado embrionario hepático* o *sarcoma embrionario indiferenciado hepático* de los que el traductor puede intuir la traducción al inglés de cada uno de los componentes del término compuesto, pero no el orden en el que aparecen en inglés. La búsqueda del núcleo del término compuesto *sarcoma* en las concordancias del corpus puede ayudar a encontrar un equivalente, que puede agilizarse con la función de agrupaciones de palabras o clusters que recupera palabras que se combinan frecuentemente en el corpus, bien por una dependencia gramatical (*depend on*, *depend de*) o colocacional como en el caso del siguiente término compuesto. En este ejemplo al buscar agrupaciones de cuatro palabras con una frecuencia mínima de dos apariciones en el corpus que contengan el término *sarcoma*, recuperamos la unidad *hepatic undifferentiated embryonal sarcoma*.

También es posible hacer una búsqueda por aproximación, por ejemplo utilizando una raíz léxica y un asterisco, como en el caso de *resección completa del tumor*, cuyo equivalente se busca en el corpus consultando *resec\** como hipótesis, pensando que el

equivalente en inglés podría tener esta raíz. Entre las concordancias encontramos *complete surgical resection*.

Los corpus también pueden emplearse para resolver una duda sobre el uso, por ejemplo la corrección sintáctica de un equivalente, si el traductor no está seguro de si el equivalente de *neoplasia maligna de hígado* es *malignant neoplasm of the liver* o *malignant neoplasm of liver*. Esta incógnita puede solventarse con la opción de colocaciones (*collocates*) o patrones (*patterns*). En la búsqueda de colocaciones de *liver*, vemos que la preposición *of* aparece siete veces en la primera posición a la izquierda y en 18 casos en la segunda posición a la izquierda, además *the* aparece también en 18 ocasiones en la primera posición a la izquierda, por lo que esto parece indicar que hay 7 ejemplos de *of liver* y 18 de *of the liver*. Esta misma información se puede deducir más fácilmente en la búsqueda de patrones de *liver*, donde se observa la construcción de la unidad *mesenchymal hamartoma of the liver* y *hamartoma of the liver*.

#### 4.3.3.9 Sánchez-Gijón (2009)

Esta autora proporciona ejemplos en los que un corpus monolingüe en la LM puede resultar útil a los estudiantes de traducción para resolver problemas lingüísticos en los que comprenden un concepto de la LO pero desconocen cómo expresarlo en la LM, o no saben que existe un término o una expresión concreta para hacerlo.

[...] it is the corpus in the target language that most helps students to identify and solve many of the linguistic problems related to knowledge shortcomings - they understand the underlying concept of a specific source language term, but they do not know how to express it in the target language, or even do not realise that it is indeed a fixed term and not just a casual expression.

(Sánchez-Gijón 2009: 120)

Por ejemplo, en la búsqueda de un equivalente en español para *room temperature* y sus hipónimos *current room temperature* y *desired room temperature*, el traductor consulta en el corpus en español sobre aires acondicionados *temperatura* y accede a contextos con el mismo significado que el texto original en los que se utiliza *temperatura ambiente*, también encuentra *temperatura ambiente actual* y *temperatura ambiente deseada*.

Otro ejemplo es la búsqueda de un equivalente para *personal injury*, que el traductor comprende y para el cual puede intuir una traducción. Consulta *persona\** en el corpus y obtiene resultados de *lesiones personales* que además comprueba que se combina con los verbos *evitar* y *producir*.

Otra utilidad del corpus es conocer la forma en la que el emisor se dirige al lector en un determinado tipo de textos, por ejemplo en lo que respecta a las formas verbales. En este sentido, la generación de una lista de palabras del corpus puede proporcionar al traductor las formas verbales más utilizadas, como en este caso *debe, ser, deben, puede, están y está*. La consulta de *deb\** genera muchos resultados de la estructura deber+infinitivo en contextos en los que el emisor da instrucciones y el receptor no aparece explícitamente. Para la misma función también se utiliza la estructura de verbo modal (deber, poder) + ser + participio. Por último, en la lista se encuentra *consulte*, una búsqueda posterior de *consult\** revela que el imperativo se utiliza en este tipo de contextos cuando se hace referencia al lector de forma explícita.

#### 4.3.4 Resumen de estrategias de búsqueda en corpus detectadas en los estudios

En los apartados anteriores hemos visto que distintos **tipos de corpus**, tanto **monolingües**, como **bilingües comparables y paralelos** pueden resultar útiles con el uso de determinadas estrategias de búsqueda para el acceso a la terminología que precisan los traductores y que en muchos casos no encuentran en los diccionarios, ni en otros recursos.

No obstante, respecto a la **metodología** empleada en los estudios cabe decir que, aunque algunos tratan de vislumbrar el uso de corpus en la traducción profesional, los sujetos examinados son en su mayoría estudiantes de traducción. Asimismo, algunos estudios se basan en la experiencia individual del investigador como docente de traducción o como traductor, o se indica que se han utilizado estudiantes en el estudio, pero no se especifica cuántos ni se aporta mayor información al respecto.

Hemos visto **estrategias de búsqueda** en corpus bilingües (paralelos y comparables) y también en corpus monolingües.

Algunos muestran la utilidad de los corpus bilingües, tanto paralelos como comparables, para la selección de equivalentes en traducción (Danielsson y Ridings 1996; Malmkjær 1998; Peters et al. 1996). Los corpus paralelos resultan de interés sobre todo en aquellos casos en los que la selección de una denominación natural en la LM no existe, o no puede utilizarse porque la traducción está condicionada por la situación bilingüe en la que se produce por cuestiones culturales (Pearson 2003). Malmkjær (1998: 539) alude a la

enorme utilidad de los corpus paralelos para comprobar cómo se expresan ideas, conceptos o información en diferentes idiomas<sup>94</sup>.

No obstante, los autores sostienen que la manera en que se compilan los corpus paralelos (normalmente un texto traducido por cada texto original) implica limitaciones en la relevancia de la información que proporcionan al traductor, ya que los equivalentes se basan en las decisiones de traducción adoptadas por un único individuo, que no tienen por qué coincidir con las de otros traductores. Malmkjær sugiere que la solución sería crear corpus paralelos en los que se proporcionara más de una traducción para cada texto original, aunque reconoce el problema de que en la mayoría de géneros los textos no se traducen más que una vez.

Asimismo, como señalan diversos autores, como Peters y Picchi (1998), las limitaciones de los corpus paralelos en la selección de equivalentes de traducción adecuados en LM reside en que estos se extraen de textos traducidos y, en este sentido, no ofrecen mayores garantías que el diccionario bilingüe salvo que los términos se encuentran rodeados de un contexto e información de uso. Los equivalentes proporcionados por los textos traducidos se encuentran de alguna manera vinculados a la LO, por lo que no comparten los rasgos lingüísticos del vocabulario de los textos originales.

Por este motivo, como hemos visto, los usuarios reclaman el uso combinado de corpus paralelos con los que puede accederse más rápidamente a los equivalentes de un texto (permitiendo restricciones tanto en la lengua origen como en la lengua meta), pero sin dejar de lado los corpus comparables donde realmente puede comprobarse el comportamiento natural de los términos en una lengua.

Otros estudios se centran en la relevancia que pueden tener para los traductores los corpus monolingües. Sin tratar de cuestionar la importancia y la utilidad de los corpus bilingües, Bowker (1998b) por ejemplo apunta la dificultad de compilar corpus paralelos, ya que a la complejidad de encontrar textos bilingües, se añade el tiempo que se tarda en alinearlos; además, muchos traductores no repiten con tanta frecuencia la temática de los textos que traducen como para contar con corpus paralelos amplios (y menos sobre cualquier ámbito).

Por esa razón, han proliferado también los estudios sobre el uso de corpus especializados monolingües, por un lado, porque son más fáciles de compilar que los

---

<sup>94</sup> Proporciona algunos ejemplos de búsqueda en la traducción de un texto literario (Malmkjær 2003).



corpus paralelos y, por otro, porque también pueden convertirse en una fuente de gran valor para acceder a la terminología de los textos que se traducen. En palabras de Bowker (2000: 21), el uso de un corpus monolingüe en LM se asemeja en gran medida a las búsquedas que los traductores han venido haciendo en lo que se ha conocido como textos paralelos, entendidos como documentos independientes en la LM que comparten la misma función comunicativa que el texto origen.

Hemos visto que la consulta en corpus puede resultar útil para acceder a información relevante para el traductor tanto en LO como en LM. Un corpus en LO puede emplearse para tareas de comprensión (Coffey 2002), de manera general para la búsqueda de conocimientos temáticos, para documentarse sobre un tema antes de ponerse a traducir (por ejemplo, sobre el *missile defence program*) (Roberts y Bossé-Andrieu 2006), y específicamente para la búsqueda del significado o la explicación de un término. En este caso, la búsqueda más habitual cuando se desconoce el significado de un término o expresión es consultarla en el corpus; además en esta estrategia el traductor suele combinar la consulta de un término con expresiones de contenido semántico (como por ejemplo *is a*) ya que pueden ayudar al traductor a recuperar contextos ricos en conocimiento como podrían ser contextos definatorios y de relaciones conceptuales entre términos.

En la búsqueda de información en LM, los estudios sugieren que un corpus en LM es útil para la búsqueda de equivalentes desconocidos (Bowker 2000; Bowker y Pearson 2002: 200; Sánchez-Gijón 2009: 120; Wilkinson 2005a, 2007: 108). En este sentido podemos distinguir entre distintas estrategias de búsquedas, algunas de tipo inductivo y otras de tipo deductivo (véase por ejemplo, Sánchez-Gijón 2004a: 212-215).

En las búsquedas de tipo inductivo, el traductor suele partir de una intuición, hipótesis o idea previa sobre el equivalente que desconoce. Esta hipótesis puede generarse a partir de los propios conocimientos que posee el traductor sobre la LM, o puede tratarse de un equivalente encontrado en un recurso como por ejemplo un diccionario que el traductor ha consultado previamente<sup>95</sup>. Dicha hipótesis se comprueba en el corpus, lo que lleva a su selección o a su rechazo. En ocasiones la hipótesis no es una palabra completa

---

<sup>95</sup> Wright y Budin (1997: 848) consideran que incluso si el diccionario nos resuelve un problema terminológico, es preferible comprobar la información obtenida en contexto: «it is highly desirable to reconfirm information found in glossaries and dictionary resources by finding *in vivo* terms in context rather than relying solely on existing dictionary or glossary resources».

sino que se intuye cuál podría ser la raíz de la palabra, que también puede comprobarse en el corpus combinándola con un comodín.

Asimismo, en la búsqueda de equivalentes de términos compuestos, el traductor puede introducir la consulta a partir del equivalente de uno de los componentes del término compuesto o una estructura sintáctica típica que sí que conozca o que intuya para acceder al término compuesto completo.

En la búsqueda deductiva no se cuenta con una propuesta de equivalente, sino que el traductor se aproxima a él en el corpus a partir de la consulta de información que ha sacado del contexto y que está relacionada con el equivalente que desea encontrar.

En este sentido, la consulta de un término relacionado (combinado con una expresión semántica si fuera necesario) puede favorecer la recuperación de sinónimos, antónimos, hipónimos, hiperónimos, etc. en los contextos de un corpus.

La búsqueda a partir de palabras que pueden aparecer en un contexto inmediato al término que se pretende encontrar también puede ser útil. Determinadas palabras invariables en LO y LM como números, nombres propios, abreviaturas o siglas pueden favorecer la recuperación de términos que se encuentran en una posición cercana contextualmente a estas formas invariables. La búsqueda de palabras clave definitorias y colocaciones típicas de una palabra pueden ser de utilidad (véase por ejemplo Bowker 2000).

También se mencionan estrategias de búsqueda a partir de categorías gramaticales en corpus anotados. En este sentido, se percibe un interés por otro tipo de anotaciones más complejas que podrían ser de utilidad en las búsquedas del traductor, como anotaciones semánticas y también discursivas (Bowker y Pearson 2002: 188-189).

Los acrónimos, abreviaturas o siglas implican problemas al traductor<sup>96</sup> que puede resolver con ayuda de un corpus. El principal escollo que plantean es la necesidad de conocer la forma desarrollada o ampliada y un equivalente en LM (Pearson 1996a: 87).

Por último, una vez localizado un candidato a equivalente a partir de las estrategias de búsqueda anteriores, la verificación del mismo es una fase que está presente en casi toda búsqueda en traducción.

---

<sup>96</sup> Bonet Heras (2004: 43) señala que el traductor técnico debe prestar atención a las siglas, ya que los textos especializados se han convertido en una «sopa de siglas», es decir, que su uso ha aumentado de manera espectacular. Sin embargo, según este autor los diccionarios de siglas y abreviaturas que se editan son un fracaso ya que se encuentran desfasados desde el mismo momento de la publicación.

Los equivalentes en LM que encuentra el traductor en un corpus se encuentran rodeados de un contexto que puede ayudar a deducir el significado o aportar una explicación que ayude al traductor a averiguar si el equivalente encontrado es el adecuado. Los autores recomiendan que la última comprobación de un equivalente se efectúe en contexto.

Hemos encontrado la verificación de un posible término equivalente, la distinción entre dos posibles equivalentes (sinónimos o variantes ortográficas) y la verificación del orden de los componentes de un término compuesto.

En los corpus, además de los equivalentes en LM, el traductor puede acceder a la fraseología o las colocaciones de dichos términos, una información muy valiosa que favorece que el traductor se exprese con naturalidad en las traducciones. Además proporcionan información adicional que en muchos casos no se encuentra en los diccionarios como convenciones textuales (por ejemplo la manera de referirse al interlocutor en un folleto turístico o las fórmulas de apertura de un contrato), cuestiones de estilo o de uso (Corpas Pastor y Seghiri 2009).

Asimismo, algunos autores abogan por un enfoque LCW (Lexicografía-Corpus-Web) en el que se combinen el uso de diccionarios monolingües y bilingües para acceder a la terminología básica y corpus e Internet para búsquedas en cualquier lengua que proporcionan información más allá de los diccionarios, especialmente sobre colocaciones y uso en contexto (Friedbichler y Friedbichler 2000).

Finalizamos este apartado con la reflexión de que las herramientas de análisis de corpus (programas de concordancias) con las que pueden consultarse corpus monolingües y también bilingües no han sido diseñadas específicamente para traductores, como sería deseable (Bowker 2000: 22). De hecho, aunque numerosos estudios tratan de demostrar la utilidad de los corpus para la búsqueda en traducción, los corpus no se han implantado con suficiente fuerza en la traducción profesional (Aston 2009: X) porque su uso puede no resultar rentable a corto plazo dadas las pérdidas de tiempo que todavía supone su consulta y sobre todo su compilación, por lo que la investigación debería encaminarse a mejorar la eficacia de los corpus desde el punto de vista de las búsquedas y la compilación, temas que hasta la fecha se encuentran lejos de ser solucionados.

#### **4.4 Estudios sobre estrategias de búsqueda de los traductores en Internet**

Hemos visto que Internet es una fuente inagotable de textos de las más diversas temáticas que, además, se actualiza constantemente. En este sentido, el problema no es encontrar información en Internet, sino que sea fiable y relevante a la búsqueda efectuada, lo que puede requerir una gran inversión de tiempo y provocar frustración.

Austermühl (2001: 52) señala que para que el traductor no se pierda en el ciberespacio, tendrá que desarrollar una serie de estrategias de búsqueda, de lo contrario encontrar la respuesta a una consulta puede ser tan complicado como buscar una aguja en un pajar.

Se ha demostrado que los traductores utilizan Internet como un gran corpus y las herramientas de búsqueda en Internet, en su mayoría los buscadores (especialmente *Google*) como una herramienta de análisis de corpus (Jääskeläinen y Mauranen 2004; Wheatley 2006), por encima de otros recursos como los corpus e incluso los diccionarios.

Autores como Ferraresi (2009) y otros que veremos en este apartado señalan que los traductores utilizan *Google* para acceder a definiciones o para buscar expresiones en la LM, y que aplican una serie de estrategias de búsqueda basándose en las funciones de los buscadores, como la búsqueda de expresiones exactas entrecomilladas o la restricción de dominios.

En esta línea, en el marco del proyecto MeLLANGE<sup>97</sup> (Wheatley 2006), se realiza un estudio en el que se encuesta a estudiantes y profesionales de la traducción sobre el uso de Internet y los corpus mientras traducen. De 1015 respuestas obtenidas (74% de traductores profesionales y 26% de estudiantes), el 94,4% utilizan *Google* durante el proceso de traducción para realizar búsquedas terminológicas, solo el 43,9% compila corpus y, por tanto, utiliza los corpus, y solo el 20,2% utiliza programas de gestión de corpus. De los que no utilizan corpus en el estudio de Wheatley (2006), el 41% nunca ha oído hablar de ellos, el 21% no tiene tiempo para compilarlos, el 20,2% no sabe utilizar programas de gestión de corpus, el 7,5% no les encuentra ventajas respecto a *Google*, el 6,7% no les encuentra ventajas respecto a las memorias de traducción y el 3,6% indica otros motivos. Jääskeläinen y Mauranen (2004: 50) reportan resultados muy similares en

---

<sup>97</sup> Acceso a los resultados completos del estudio en <http://www.mellange.eila.jussieu.fr/Mellange-Results-1.pdf>.

un cuestionario en el que preguntan a traductores finlandeses sobre el uso de herramientas electrónicas para la búsqueda de terminología. Los resultados demuestran que, mientras que el uso de diccionarios electrónicos y de Internet es generalizado entre los encuestados, los corpus son herramientas casi desconocidas.

Por tanto, los traductores dan prioridad a los buscadores de Internet para realizar sus búsquedas terminológicas. Esto se debe seguramente a que los buscadores presentan la ventaja de que los usuarios, entre ellos los traductores, están más familiarizados con el uso de los buscadores que con otras herramientas. Tal y como sugiere el estudio efectuado por Wheatley (2006), la mayoría de traductores desconocen el uso de las herramientas de gestión de corpus. Además, no sienten la necesidad de invertir el escaso tiempo del que disponen en aprender a utilizar nuevas herramientas de búsqueda.

De hecho, aunque hace unos años algunos traductores todavía tenían reparos en afirmar que utilizaban *Google* como recurso para la documentación y búsqueda terminológica, actualmente admiten su uso sin problema e incluso sitúan a las herramientas disponibles en Internet como algunos de sus principales recursos de búsqueda (Alcina et al. 2005: 221; Budin y Wright 1997: 845; Corpas Pastor 2004: 225; Ferraresi 2009; Friedbichler y Friedbichler 2000: 112; Montero Martínez y Faber 2008: 165-166; Roberts y Bossé-Andrieu 2006: 203; Sánchez-Gijón 2003b: 1). Esta tendencia probablemente aumentará en los próximos años (Maniez 2007: 166).

Internet se ha convertido en uno de los recursos más utilizado por los traductores, incluso por encima del resto de recursos de consulta (Palomares Perraut y Pinto 2000: 116), no solo para la búsqueda documental en general<sup>98</sup> (Gonzalo García 2004: 279), sino también específicamente para la búsqueda terminológica (Budin y Wright 1997: 845; Champe 1997: 508; Gómez 1999; Jääskeläinen y Mauranen 2004: 50; Wheatley 2006). Champe (1997: 508) señala lo siguiente.

---

<sup>98</sup> La competencia documental en traducción no solo implica la búsqueda de información temática, sino que también engloba, como hemos visto en apartados anteriores, encontrar soluciones para los problemas terminológicos específicos que plantea una traducción. Así lo explica Gonzalo García (2004: 277) en su definición de competencia documental: «[...] el traductor que actúa de mediador en el proceso comunicativo entre especialistas, tendrá que adquirir un nivel suficiente de competencia sobre el tema, conocer las reglas de confección de los textos especializados en ambas lenguas (concisión, precisión y sistematicidad) y la diversidad de tipos o géneros textuales (artículo divulgativo, informe técnico, patentes, normas, prospectos, manuales de instrucciones, etc.), y, dentro de cada ámbito, distinguir las unidades terminológicas y fraseológicas específicas».

After terms have been extracted from the text of the manual, there are many ways to obtain information about them. It is increasingly possible to find terms in English on the World Wide Web, using a combination of search engines and directory queries. Information retrieved by such a search can provide explanations, definitions, descriptions of function and, what is particularly important when looking for or coining TL equivalents, synonyms in the SL.

(Champe 1997: 508)

Por un lado, ya hemos visto que la investigación aboga por el estudio de los usuarios para la elaboración de herramientas centradas en el usuario y, por otro lado, el traductor es un usuario potencial de Internet. De hecho, se observa un cambio en la investigación en Internet de un *system-oriented approach* a un *user-oriented approach*, que se plasma en el creciente interés por indagar el uso de Internet por parte de sus usuarios (Pinto y Sales 2007: 533).

Sin embargo, según algunos autores, esto contrasta con la todavía patente escasez de estudios empíricos específicos sobre las estrategias de búsqueda de los traductores en Internet (Enríquez Raído 2011; Pinto y Sales 2007: 536). En el marco de investigaciones más generales sobre el proceso de traducción o la competencia traductora, Internet se ha mencionado en mayor o menor medida dentro de los recursos de referencia del traductor, aunque sin profundizarse en las búsquedas (a modo de ejemplo véanse los trabajos descritos en el capítulo 2).

Algunos estudios, que sí se centran en las búsquedas de los traductores en los recursos, hemos visto que analizan el uso de diccionarios. También hemos recogido estudios sobre el uso de corpus en traducción, algunos de los cuales incluyen ejemplos de búsquedas en Internet, aunque no como principal objeto de estudio (sirvan de ejemplo Kübler 2003; Varantola 2006; Vargas Sierra 2002).

Las estrategias de búsqueda de traductores en Internet que presentamos en este apartado se han obtenido del análisis de trabajos de distinto tipo. Nos referiremos a estudios experimentales específicos sobre el comportamiento en la búsqueda de información de los traductores en Internet (*information behaviour*) de Alanen (1996), Palomares Perraut y Pinto (2000), Pinto y Sales (2007), Enríquez Raído (2011). Asimismo, algunos traductores profesionales comparten su propia experiencia sobre las estrategias de búsqueda que utilizan en Internet (Gambín 2004; Gómez 1999; Sevilla Muñoz y Sevilla Muñoz 2005a, 2005b; Walter 2006). También hemos extraído una serie de estrategias de búsqueda de tutoriales específicos de búsqueda en Internet adaptados para traductores o mediadores lingüísticos (Bergeron y Van Steenberg 2000; Riediger 2009a; Riediger et al. 2006). Algunos manuales de traducción y terminología para traductores presentan capítulos dedicados a las búsquedas en Internet (Austermühl 2001; Budin y Wright 1997;

Montero Martínez y Faber 2008). Los principales hallazgos de estos trabajos los resumimos a continuación por orden cronológico.

#### **4.4.1 Alanen (1996)**

Alanen (1996) distribuye un cuestionario sobre el uso de Internet a una serie de traductores profesionales. Los resultados muestran que los servicios de Internet más utilizados por los traductores son, por este orden, el correo electrónico, la Web en general y las listas de correo. También mencionan el ftp, los grupos de noticias y el IRC.

En otra pregunta, los traductores indican que Internet favorece el acceso a otras fuentes de información valiosas como son las personas, los traductores se refieren a compañeros, otros profesionales o especialistas, amigos y familiares, aficionados a algún tema, y también los propios clientes.

La información que los traductores buscan en Internet son términos, conceptos, modismos e información contextual, ortográfica y temática. La mitad de los participantes alegan utilizar Internet cuando los recursos tradicionales de búsqueda fallan. De estos, unos pocos como último recurso. La otra mitad sitúan a Internet como su principal recurso de búsqueda. Uno de los traductores con un año de experiencia se plantea incluso cómo los traductores podían trabajar antes sin la existencia de la red.

Finalmente, entre las ventajas de Internet se menciona la rapidez y el fácil acceso a multitud de recursos actualizados, el contacto con los clientes y la transferencia de archivos. Entre las desventajas incluyen la inversión de tiempo, el ruido consistente en mensajes o publicidad no deseada, el excesivo volumen de información difícil de analizar, las distracciones o las adicciones a la red, el funcionamiento lento o el bloqueo de los servidores, problemas de compatibilidad y la escasa fiabilidad.

#### **4.4.2 Bergeron y Van Steenberg (2000)**

Bergeron y Van Steenberg (2000) presentan un tutorial de búsqueda en Internet para traductores donde se centran principalmente en estrategias de búsqueda en buscadores. Según estos autores, los directorios pueden resultar útiles al traductor solo en el caso de que el traductor pretenda documentarse sobre un tema de manera general, pero no para las búsquedas terminológicas.

Para la selección de las palabras clave que se introducirán en el buscador, se recomienda comenzar la búsqueda con palabras más generales para después tratar de restringir o redefinir la búsqueda<sup>99</sup>. Por ejemplo, al traductor puede interesarle limitar los resultados a páginas de una lengua determinada. Asimismo, si una consulta no ha proporcionado suficientes resultados, esta puede ampliarse con el uso de sinónimos.

Algunas estrategias específicas son la búsqueda de información general en glosarios monolingües, por medio de la introducción de la palabra clave que identifica el tema del que se pretende encontrar información junto con palabras clave que significan glosario en la lengua que se precise (*glossaire, glosario, glosario, glosar, glossarium, o variantes lexique, dictionnaire, encyclopédie, etc.*). Por ejemplo, *+assurance\* +glossaire* para recuperar glosarios de seguros en francés.

Otra estrategia denominada por estos autores como *reverse search* consiste en acceder a páginas web relevantes partiendo de una página que el traductor considera fiable. Con el operador *link:* se puede saber qué páginas se encuentran vinculadas a la página que se introduce a continuación del operador.

En la búsqueda de definiciones en Internet, estos autores proponen introducir la palabra cuyo significado se desconoce a la vez que expresiones que suelen introducir definiciones, con la función de entrecomillado, por ejemplo «*la demutualisation est*».

En la búsqueda de equivalentes indican que, una vez localizado un equivalente en un diccionario, se compruebe el contexto y uso introduciéndolo en un buscador. En el caso de que no haya sido posible encontrar un equivalente previamente, la estrategia planteada es tratar de adivinar una posible equivalencia e introducirla en un buscador como palabra clave de búsqueda para verificarla. Por ejemplo, en una traducción del francés al inglés sobre el poeta *Simonide* en el que aparece también la expresión *Art de la Mémoire* el traductor necesita conocer sus denominaciones en inglés. Supone que el nombre del poeta en inglés tendrá la misma raíz que en francés pero distinta terminación. Para la otra expresión probablemente se emplee la traducción literal *The Art of Memory* en inglés, por lo que introduce la consulta *+art of memory +simon\**. En los primeros resultados el traductor accede a páginas en las que se explica que *Simonides* escribió *The Art of Memory*.

---

<sup>99</sup> Los manuales de búsqueda en Internet recomiendan que en las palabras clave de búsqueda no se incluyan palabras vacías de significado como artículos, pronombres, conjunciones o preposiciones, ya que los buscadores no las tendrán en cuenta (Bergman 2004: 15).



En ocasiones la estrategia de búsqueda de una raíz mediante la introducción de palabras truncadas<sup>100</sup> es útil para detectar equivalentes en determinados pares de lenguas semejantes o en determinadas disciplinas donde la terminología es similar de unas lenguas a otras, por ejemplo para conocer el equivalente holandés del término inglés *pilocytic astrocytoma*, la consulta *+pilocyt\* +astrocyt\** recupera *pilocytair astrocytoom*.

Otra estrategia para el acceso a equivalentes es el uso de la función de idiomas del buscador. En este caso se introduce el término en LO y se restringe la búsqueda a páginas de la LM con el fin de recuperar páginas en la LM que contengan el término en LO con una definición o un equivalente. Por ejemplo, el traductor introduce *+“goldilocks economy”* con el filtro de lenguas francés; también puede añadirse el nombre de la LM a la consulta, *+“white knight” +français*. Asimismo, la introducción de palabras clave únicas a una lengua (la preposición *avec* en francés) puede estimular la recuperación de páginas de dicha lengua, por ejemplo *+“leverage buy out” +avec*. De la misma forma, puede resultar útil introducir palabras clave en lengua meta asociadas con el tema de la traducción, *+“white knight +opa* (opa son las siglas de *offre public d’achat*).

Para algunos términos como nombres de plantas o animales funciona la estrategia de búsqueda en latín como puente entre dos lenguas. Para ello, se introduce el término en LO (junto con la palabra *latin*) para acceder a páginas en las que se incluya la equivalencia latina. Posteriormente se introduce el término en latín y se restringe la búsqueda a páginas de la LM. Por ejemplo, en la traducción inglés-español de *coalfish* se introduce la consulta *+coalfish +latin*. En la primera página de resultados se accede a la expresión latina equivalente *Pollachius virens*. A continuación se consulta *+“pollachius virens”* con la restricción de idiomas español y en el primer resultado aparece *carbonero*. Este término puede verificarse con una nueva consulta en el buscador.

Para confirmar un término, su uso u ortografía pueden introducirse las distintas opciones en el buscador y comprobar el número de resultados, aunque los datos estadísticos no son siempre concluyentes. Por ejemplo, pueden comprobarse las expresiones *en toute circonstance* y *en toutes circonstances*.

---

<sup>100</sup> Aunque la búsqueda a partir de las raíces de las palabras podría resultar una estrategia de búsqueda en Internet útil para el traductor, el análisis de técnicas de búsqueda del apartado anterior y la consulta de manuales de Internet sugieren que la posibilidad de truncamiento (introducción de una raíz seguida de un asterisco para abarcar en una consulta todas las variantes o terminaciones existentes) no es posible en la mayoría de buscadores. Así, por ejemplo, para la búsqueda de palabras en singular y plural recomiendan introducir las dos variantes (bird OR birds) (Bergman 2004: 17).

La búsqueda de imágenes puede ayudar a comprender el significado de un término o a comprobar si el término recuperado es el adecuado. Por ejemplo, al traductor puede interesarle observar imágenes de distintos tractores para comprender cómo funcionan sus partes, para ello introduce la consulta *+tractor +caterpillar*.

#### 4.4.3 Palomares Perraut y Pinto (2000)

Palomares Perraut y Pinto (2000) tratan de averiguar las necesidades, hábitos y usos documentales del traductor profesional. Para ello, distribuyen un cuestionario a través de listas de correo de traducción en inglés, francés y español. La encuesta consta de 19 preguntas de respuesta cerrada (solo dos son parcialmente abiertas). Se analizan 96 cuestionarios.

En cuanto al perfil de los traductores encuestados, un 84,4% son traductores *freelance* que trabajan en empresa privada y tienen mucha experiencia (casi la mitad más de 10 años). En cuanto a la formación académica, el 68,4% posee estudios en Traducción. Las lenguas de trabajo son principalmente el inglés como LO (85%), el 57,3% trabaja también con el español y el resto de idiomas por orden descendente son italiano, francés, portugués, alemán, ruso, catalán, árabe, gallego y noruego.

Respecto a la tipología de textos traducidos (teniendo en cuenta que un mismo traductor puede traducir distintos tipos de textos) destacan las traducciones técnicas (electrónica, telecomunicaciones, tecnología de telefonía móvil), después las traducciones comerciales, seguidas de las legales; la categoría otros ocupa el siguiente lugar y, al ser abierta, puede incluir muchos tipos de texto que no habían previsto los investigadores; a continuación aparecen los textos literarios y los periodísticos y también se mencionan textos realizados con traducción automática.

Se pregunta sobre la importancia de la documentación en el proceso traductor, el tiempo invertido y los tipos de recursos empleados. Para el 93,4% la documentación es una necesidad, más de la mitad dedica el 25% del tiempo a ello y se realiza sobre todo en la fase de transferencia del significado a la LM. Los autores observan un descenso de la documentación en los traductores con más años de experiencia.

Las necesidades documentales principales que se señalan son las terminológicas, también se mencionan la necesidad de información conceptual o temática, problemas gramaticales, de estilo y fraseológicos. Estas necesidades existen independientemente de la experiencia del traductor.

En cuanto a los hábitos para documentarse, el primer lugar de consulta es el archivo personal del traductor, y casi en la misma medida Internet, después expertos o colegas, también bibliotecas y clientes. En la pregunta sobre una segunda fuente de información se mencionan los diccionarios y no se observa preferencias por el formato electrónico o papel.

En cuanto a la pregunta específica sobre el uso de Internet, el 99% lo utiliza como fuente de información y más del 50% afirma que recurre a Internet siempre que traduce. Las aplicaciones más utilizadas son el correo electrónico, seguido de la Web (tanto para la comunicación como para la búsqueda de información), después aparecen las listas de correo a modo de foros para preguntar a otros traductores por la traducción de términos, los chats aparecen muy por debajo.

La satisfacción con los recursos es alta (74%), en esta pregunta el 27% añade la razón de su satisfacción, la mayoría se refiere a Internet, como en las siguientes afirmaciones: *all information needed can be found in the Internet; especialmente na Internet consegue-se praticamente cualquier tipo de informação; the Internet is an infinite resource for any kind of information needed, Internet is the most generous about information*; aunque para los traductores con lenguas minoritarias Internet no parece tan útil: *because of my language pair. It's quite difficult to find something p e in Slovenian.*

#### 4.4.4 Austermühl (2001)

Este autor identifica tres tipos de búsqueda en Internet, la búsqueda institucional (*institutional search*), la búsqueda temática (*thematic search*) y la búsqueda de palabras (*word search*).

La *búsqueda institucional* se efectúa sobre todo en periódicos y organizaciones internacionales (ONU, UNESCO, UNICEF, OTAN, UE, Banco Mundial), como fuentes de textos paralelos. En estas páginas el traductor debe descubrir los modos de encontrar los contenidos de forma sistemática, accediendo por ejemplo a las funciones *Search*, *Documents*, *Databases*, *Glossaries*, etc. Las teclas CTRL+F pueden emplearse también una vez identificado el texto que le interesa al traductor para acceder a la terminología de forma más directa sin necesidad de leer todo el texto.

La *búsqueda temática*, a partir de áreas de conocimiento, se realiza en directorios como *Yahoo!*, que organizan las páginas web en categorías y subcategorías, a las que se accede a través de la estructura temática, aunque también pueden realizarse búsquedas de palabras clave en las categorías. Esta es una búsqueda que puede resultar útil para encontrar información general temática, especialmente en la fase de recepción o

comprensión, pero también en la de producción para comprobar si determinados términos o frases pueden emplearse en la LM. Por ejemplo, en la búsqueda de información sobre biotecnología, donde también puede interesar si en inglés se utiliza *biotech* o *biotechnology*.

La *búsqueda de palabras* según este autor se efectúa en buscadores, los cuales comparan secuencias de caracteres introducidas por el traductor con su base de datos, sin realizar diferenciación semántica. En este caso la consulta se realiza a partir de palabras clave, que pueden estar combinadas o no con distintos tipos de operadores y filtros de restricción.

Los operadores booleanos AND (con su alternativa +) pueden emplearse para buscar más de una palabra +*dictionary* +*Spanish*; OR para incluir sinónimos o equivalentes en la búsqueda *dictionary OR glossary, dictionary OR diccionario OR dictionnaire*; NOT para excluir palabras.

Los operadores de proximidad, como las comillas para buscar una frase exacta, ADJ y FOLLOWED BY para la búsqueda de frases y NEAR para la combinación de palabras a cierta distancia. NEAR, FOLLOWED BY y ADJ resultan útiles al traductor para localizar colocaciones determinando el rango colocacional de una combinación lingüística. Por ejemplo, en la confirmación de terminología y fraseología se puede comprobar la existencia de una expresión buscándola entre comillas “to implement a strategy”. También en la traducción del inglés al alemán de la expresión *to dispel doubts*, normalmente los traductores no tendrán problemas en traducir uno de los componentes de la colocación, en este caso el equivalente de *doubt* es *Zweifel*, pero no resulta tan obvio saber si se puede hablar de *Zweifel zerstören, Zweifel auflösen, Zweifel vertreiben, Zweifel zerstreuen*. Para solucionarlo puede emplearse el operador NEAR combinando ambos componentes de las posibles expresiones en alemán. El operador NEAR es útil porque no siempre los componentes de una colocación aparecen contiguos sino que pueden encontrarse a cierta distancia. Una vez identificadas las páginas con la posible solución, *Zweifel zerstreuen*, el traductor puede comprobar rápidamente la validez de la colocación buscando con CTRL+F.

Los filtros de formato, como por ejemplo PDF o Word permiten acceder a un tipo concreto de textos, lo que también resulta útil al traductor para crear su propio corpus. La restricción de dominio puede limitar la búsqueda a páginas de un país y el truncamiento de palabras tiene en cuenta las variantes gramaticales de un mismo término.

Además de la búsqueda propiamente en Internet, este autor menciona la existencia de otras herramientas de búsqueda en línea útiles para el traductor, como catálogos de bibliotecas, tiendas de libros, enciclopedias virtuales, diccionarios, glosarios, bases de datos multilingües, periódicos y revistas.

#### 4.4.5 Gambín (2004)

Gambín (2004) también aporta una serie de estrategias para acceder a la terminología en una traducción. Gracias a Internet, el traductor puede acceder a textos paralelos escritos originalmente, en este sentido la estrategia es buscar textos en la LM en los que aparezcan los términos en la LO para los cuales se precisa un equivalente. Internet también facilita el acceso a especialistas en la materia a los que el traductor puede consultar en caso necesario.

Asimismo, presenta un resumen de las opciones de búsqueda de *Google* que más útiles pueden resultar al traductor y aporta ejemplos. El signo + puede emplearse en la búsqueda de acrónimos, abreviaturas o siglas que pueden confundirse con palabras vacías de significado y que el buscador ignoraría, como AS (Associate in Science). El signo – se emplea para reducir el ruido de la búsqueda excluyendo aquellas palabras que recuperen páginas que el traductor no desea. Las comillas pueden servir para verificar expresiones, por ejemplo “*associate in science (AS)*” y también para recuperar definiciones cuando no se encuentran en el diccionario, “*trans facts are defined*”.

La restricción por dominios se emplea para verificar terminología en la página de un cliente, por ejemplo para comprobar si la empresa Intel utiliza el término *placa base* o *placa madre* se efectúan dos consultas *site:intel.com “placa base”* y *site:intel.com “placa madre”*. Otra función es la de comprobar las variantes regionales de un término en una lengua, por ejemplo en español *computadora* aparece en 11.500 páginas con dominio .es de España frente a las 69.300 de dominio .mx de Méjico; *ordenador* se encuentra en 137.000 páginas con dominio de España y en 4.100 con dominio de Méjico. Finalmente, esta restricción actúa como filtro para recuperar páginas que previsiblemente han sido escritas originalmente en una lengua, no son traducciones. Por último, el operador OR puede ser útil para comprobar un término en distintas páginas de prestigio, por ejemplo *placa base* en las páginas de las empresas Microsoft, Intel y Hewlett Packard con la consulta “*placa base site:hp.es OR site:microsoft.com OR site:intel.com*”.

Asimismo, alude a la consulta en foros de traductores como *ProZ*, *aquarius.net* o *gotranslators.com*, aunque advierte de la importancia de validar siempre las respuestas ya

que no siempre son fiables y en muchos de estos foros los traductores responden a las preguntas con el fin de obtener puntos con los que mostrar información sobre su experiencia a clientes potenciales.

#### 4.4.6 Gile (2004)

De los resultados proporcionados por Gile (2004), encontramos la adquisición de información *ad hoc* en Internet, que ha reemplazado casi por completo los textos impresos en la Université Lyon 2, aunque para esta autora la confianza de los textos de Internet es excesiva especialmente en la búsqueda de colocaciones.

Por ejemplo, en un texto sobre psicología cognitiva, donde aparece la siguiente afirmación sobre cómo estudiar el funcionamiento cognitivo humano: *An obvious but naïve answer is that one studies the physiological mechanisms that underlie the behaviour*, una estudiante traduce *studies* por *explorer* (explorar) porque la colocación *explorer le cerveau* la ha encontrado en muchas páginas de Internet en un contexto parecido al de la traducción. Por el contrario, las estudiantes no utilizan palabras que no hayan encontrado en Internet, por ejemplo una estudiante pensó en utilizar *examiner leur cerveau* como traducción de *inspect their brain* pero la descartó porque no encontró colocaciones en Internet con la consulta *examiner + cerveau*.

Según la autora, el relativamente alto número de casos de este tipo plantea un riesgo de que los estudiantes tomen Internet como un macrodiccionario y un sustituto de su propio análisis.

En un texto sobre clonación, la oración *Cloning is the production of a duplicate organism without the process of reproduction* se traduce por *Le clonage est la duplication d'un organisme en dehors du processus naturel de la reproduction*, donde se añade al texto meta el adjetivo *naturel*, que no aparece en el texto origen. Esta decisión se justifica porque la estudiante quiso aportar claridad al discurso, ya que se estaba hablando de dos tipos de reproducción y dicha distinción se efectuaba en páginas de Internet que había consultado como [www.cne-ethique.fr](http://www.cne-ethique.fr) y [www.inrmarseille.com](http://www.inrmarseille.com).

Los ejemplos encontrados en los informes también permiten al profesor advertir a los estudiantes sobre el riesgo de creer ciegamente en las respuestas que encontramos en los diccionarios, ya que palabras que tienen un significado en el lenguaje general pueden utilizarse y adquirir un significado específico en un lenguaje especializado. Es el caso de la oración *What is important is that the theory be accurate in predicting a subject's actions under a certain condition*, que aparece en un texto sobre psicología cognitiva. Una

estudiante evita el uso de *prédire* en la traducción de *predicting* (aunque es adecuado en este contexto) y selecciona *prévoir*, que es sinónimo de anticipación en un sentido muy amplio. Explica su elección porque el diccionario general francés *Petit Robert* indica que *prédire* se emplea en el campo de la parapsicología. Sin embargo, estudiantes que consultaron textos especializados de Internet observaron que *prédire* podía emplearse en el contexto de la traducción.

No obstante, hay que tener cuidado en este punto porque la experiencia de la autora como docente le ha demostrado que en ocasiones los estudiantes seleccionan una unidad léxica para el texto meta en función de su parecido con la del texto origen, sin realizar un análisis sistemático de su significado real.

En otras ocasiones la elección entre dos términos se produce en función de cuestiones estadísticas o cuantitativas (de resultados recuperados por *Google*) que pueden asociarse erróneamente a criterios cualitativos de validez del término o de su uso más extendido.

Por ejemplo en la traducción de un texto para un público francés, las estudiantes sabían que las búsquedas tenían que realizarse en páginas de Francia, en lugar de páginas de Canadá, Suiza o Bélgica. Sin embargo, lo que no tuvieron en cuenta a la hora de basar sus decisiones en criterios cuantitativos fue que el número de páginas canadienses indizadas en el buscador era mucho mayor que francesas, suizas o belgas y que el número más elevado de resultados a favor de uno u otro término podía determinarlo que las páginas canadienses eran más numerosas.

#### **4.4.7 Walter (2006)**

Walter (2006) presenta algunas estrategias de búsqueda que pueden emplearse en *Google* y que resultan útiles desde el punto de vista del traductor profesional para el acceso a equivalentes en LM. Se basa principalmente en la generación y confirmación de hipótesis a partir de los conocimientos lingüísticos previos, así como la búsqueda en paralelo del término original y de una hipótesis de equivalente o parte de él combinado con el filtro de restricción de idiomas.

En la primera estrategia propuesta, el traductor busca un equivalente en inglés para el término en alemán *filtergängig*, un adjetivo que denota aquello que puede pasar por un filtro. El traductor genera una hipótesis, *filter-passing* y, para confirmarla, introduce ambos términos en un buscador con el objetivo de acceder a algún texto paralelo. Entre los resultados encuentra un texto inglés-alemán en el que verifica el equivalente.

En otra estrategia el traductor busca un equivalente en inglés británico para *Kreuzdiagnosen und Sterndiagnosen* y supone que están relacionados con el ICD (International Classification of Diseases). Con la consulta *ICD Kreuzdiagnosen site:.de* verifica que *Kreuzdiagnosen* es superior a *Sterndiagnosen* y que ambos pertenecen al ICD. Con la hipótesis de que *Stern* equivale a *asterisk* se introduce la consulta *ICD "asterisk codes" site:.uk* con lo que se accede a dos páginas fiables que confirman el equivalente y que contienen también *dagger codes*, el equivalente de *Kreuzdiagnosen*.

En el tercer ejemplo se busca un equivalente en alemán para la expresión en inglés *feed and discharge chutes*. El traductor conoce que *feed and discharge* normalmente se traducen por *Aufgabe* o *Austrag* y *chute* por *Schurren* o *Rutschen*. Así, realiza consultas en páginas en alemán *site:.de* de todos los posibles términos compuestos *Aufgabeschurre*, *Aufgaberutsche*, *Austragsrutsche* y *Austragsschurre*. En los resultados se encuentra una página con imágenes de *Aufgabeschurre* y otra en la que aparece *Austragsschurre*.

En la cuarta estrategia, el traductor busca un equivalente en inglés del término alemán *Rückbildungsphase* en el contexto de un electrocardiograma. Una búsqueda en la enciclopedia *Wikipedia* de la entrada de electrocardiograma en inglés (*ECG*) y en alemán (*EKG*) permite acceder a los textos paralelos en los que se obtiene el equivalente *repolarization phase*.

Los recursos en línea que este traductor profesional considera útiles son las enciclopedias monolingües disponibles en distintas lenguas, como por ejemplo *Wikipedia*, glosarios o diccionarios bilingües o multilingües, textos paralelos existentes en sitios web de empresas multinacionales o instituciones públicas, documentos monolingües especializados, como por ejemplo artículos, accesibles a partir de *Google Scholar*, portales y foros de traductores e intérpretes, como el de *ProZ*.

#### 4.4.8 Maniez (2007)

Maniez (2007) concibe Internet como un enorme corpus en bruto que se actualiza constantemente y a los buscadores como programas de análisis de corpus. En este sentido, muestra cómo el traductor especializado puede acceder a terminología y fraseología en línea. Se centra en la medicina, un tema que según el autor está muy extendido en la Web, y explica cómo comprobar la validez de grupos nominales cuando el traductor tiene dudas en cuestiones sintácticas.

Uno de los ejemplos propuestos es la búsqueda de términos poco comunes o patrones sintácticos. En este caso, se pretende acceder a colocaciones del adjetivo *dose-dependent*.



La consulta de dicho adjetivo en un buscador recupera 2 millones de resultados (frente a los 7 del *British National Corpus*). En la mayoría de ellos aparece combinado con el sustantivo *effect* o su plural *effects*. Si se quisiera excluir este término de los resultados, ya que no es la colocación que el traductor está buscando, puede emplearse el operador menos: *dose-dependent –effect –effects*, con lo que se accede a otras colocaciones como *inhibition, response, stimulation, etc.*

También es posible comprobar la frecuencia de uso de los términos y buscar equivalentes. En la búsqueda de un equivalente en francés del término en inglés *control the infection* el traductor compara la frecuencia de aparición mediante una búsqueda en *Google* del término original (que aparece en 73.000 páginas en inglés) y de la traducción literal al francés *contrôler l'infection* (que aparece en 730 páginas). Incluso teniendo en cuenta que el porcentaje de páginas en inglés en la red es mayor que las de francés, los resultados llevan al traductor a pensar que existe un equivalente en francés más adecuado.

El uso del comodín ayuda cuando el traductor desconoce el equivalente exacto de un término compuesto pero conoce o intuye la traducción de alguno de sus componentes, en este caso se introducen las consultas *pour \* l'infection* y *en \* l'infection*, que recuperan expresiones más frecuentes, *combattre / éliminer l'infection*.

En otro ejemplo, en la búsqueda de un equivalente en francés para el término en inglés *gluten-sensitive enteropathy*, la consulta de *entéropathie \* gluten* recupera más de 350 resultados entre los que se incluyen los equivalentes *entéropathie au gluten, entéropathie d'intolérance au gluten, entéropathie induite par le gluten, entéropathie de sensibilité au gluten*.

El comodín también puede ser útil cuando el traductor conoce un término que está asociado al equivalente que desconoce. Por ejemplo, en la búsqueda del equivalente desconocido de *petite circulation* (que no encuentra en el *Grand Dictionnaire Terminologique*), el traductor conoce el equivalente del término asociado *grande circulation (systemic circulation)*. La consulta “*systemic and \* circulation*” permite acceder en los primeros 20 resultados a la expresión *systemic and pulmonary circulation*. La entrada de *pulmonary circulation* en el *Grand Dictionnaire Terminologique* contiene el sinónimo *lesser circulation*, que podría ser el equivalente de *petite circulation*.

Otra función señalada es la de corpus diacrónico con la función de restricción por fecha o intervalo de fechas, por ejemplo en *Google*. Una función que el autor prevé para el futuro es la de corrección gramatical con la que los buscadores serán capaces de detectar que la secuencia de palabras consultada (por ejemplo, *blood rich in carbon dioxide*) tiene

una frecuencia escasa en la red y proponer otras más recurrentes y probablemente correctas (*carbon dioxide-rich blood*).

#### 4.4.9 Pinto y Sales (2007)

Este es un estudio sobre el comportamiento en la búsqueda de información en Internet de estudiantes de traducción. En primer lugar, las investigadoras recurren a la metodología del panel de expertos, en el que participan 8 profesores de traducción expertos en diversas áreas (dos de interpretación, uno de traducción general, uno de traducción científico-técnica, uno de traducción jurídica y administrativa, uno de traducción audiovisual y uno de traducción literaria). La idea es recopilar por medio de sugerencias por correo electrónico las necesidades de documentación e información de los futuros traductores para incorporarlas en las planificaciones docentes. Los resultados de este estudio previo aportan ideas<sup>101</sup> a las autoras para desarrollar un segundo instrumento de recogida de datos, un cuestionario.

Los participantes que responden al cuestionario son 193 estudiantes de traducción e interpretación de la Universitat Jaume I, de segundo, tercero o cuarto año (se descarta a los de primero porque todavía no han cursado la asignatura de Documentación).

El cuestionario consta de 25 preguntas. Las primeras 20 son de respuesta cerrada, donde tienen que valorarse la importancia y el éxito alcanzado en una serie de elementos, con una escala de Likert<sup>102</sup> de 1 a 4: habilidades de organización y planificación, resolución de problemas, capacidad de análisis y síntesis, conocimientos informáticos, habilidades en la gestión de información y toma de decisiones.

Los resultados muestran que, de las cinco habilidades que se miden, para los alumnos de primer ciclo las más importantes son la toma de decisiones, seguida de la capacidad de análisis y síntesis, habilidades de organización y planificación, habilidades en la gestión de información y, por último, conocimientos informáticos.

Los estudiantes de segundo ciclo también dan más importancia a la toma de decisiones, pero después a los conocimientos informáticos, habilidades en la gestión de información, capacidad de análisis y síntesis y habilidades de organización y planificación.

---

<sup>101</sup> También se tienen en cuenta otras aportaciones como las del estudio de la ANECA, las normas INFOLIT y la propuesta ALFINTRA (Pinto 2005).

<sup>102</sup> Los encuestados marcan en una escala en qué medida están de acuerdo o en desacuerdo con una pregunta o una declaración.

Asimismo, las habilidades en las que se requiere mejora según los estudiantes son la gestión de información, la organización y planificación y los conocimientos informáticos.

Las últimas 5 preguntas del cuestionario son de respuesta abierta y tratan de indagar sobre las estrategias de búsqueda utilizadas por los estudiantes. Los estudiantes tienen que explicar cómo averiguarían qué obras de Federico García Lorca se han traducido al chino; si alguien ha obtenido un doctorado en España en traducción audiovisual; un sinónimo de *shaggy*; si alguna universidad española ofrece la combinación de idiomas español-búlgaro en los estudios de traducción e interpretación; y quién escribió la novela *Leila.exe*.

Los resultados muestran que algunos estudiantes no sabrían resolver estas situaciones, la mayoría realizarían una búsqueda en *Google* sin refinar (solo algunos estudiantes de los últimos cursos utilizan algunas funciones de búsqueda más avanzadas, y utilizan otros recursos tanto electrónicos como personales, aunque las autoras explican que los estudiantes no reflejan sus estrategias de búsqueda). En la segunda situación algunos se refieren a las bases de datos ISBN y TESEO; en la cuarta mencionan portales universitarios, páginas de departamentos de traducción e interpretación y profesores que conocen; en la quinta mencionan de nuevo la base de datos ISBN y sitios web de librerías, como el de la FNAC.

Las sugerencias de mejora de las autoras es enseñar a los estudiantes de primer ciclo estrategias para la búsqueda y recuperación de información en recursos electrónicos (desde una perspectiva evaluadora) y repertorios de recursos; en segundos ciclos se recomienda ahondar en la gestión de la información para la optimización de los recursos, actualización de los repertorios de recursos y estrategias avanzadas para la búsqueda y recuperación de información en recursos electrónicos (también desde una perspectiva evaluadora).

#### **4.4.10 Montero Martínez y Faber (2008)**

Montero Martínez y Faber (2008) aportan ejemplos de estrategias de búsqueda en buscadores de Internet, principalmente *Google*, que pueden resultar, según estas autoras, indispensables para el traductor en la consulta de términos, expresiones y contextos.

Según indican, una buena estrategia de búsqueda en Internet combina la utilización de *palabras clave u operandos y operadores* para restringir los resultados. En este sentido proponen tres tipos de operadores: *los operadores básicos*, el signo más, que suele aplicarlo por defecto el buscador, por ejemplo en la búsqueda de *industria glosario*; el signo menos, colocado delante de las palabras para excluirlas de la búsqueda, como por ejemplo en *fabric -textile -cloths*; el asterisco para truncar una palabra, por ejemplo *glos\**

recuperará *glosario, glossary, glossaire, glosarios, glosa, etc.*; las comillas, que según las autoras resultan muy útiles para buscar expresiones compuestas con el mismo orden que se han introducido, como en “masonry bonded hollow wall”; *los operadores lógicos o expresiones booleanas*, como AND por ejemplo en *texture AND rock AND metamorphic*, OR (*texture OR fabric*), AND NOT (*fabric AND NOT textile*) y NEAR en el que los términos deberán aparecer separados por un número de palabras (*catarata NEAR ojo*).

Los *operadores filtro* sirven para limitar las búsquedas sobre todo en lo que respecta a la calidad de los textos recuperados. Según Montero Martínez y Faber, uno de los más provechosos para el traductor es *site:*, que puede utilizarse seguido de un dominio relacionado con un campo de actividad o con el código de un país.

Una estrategia de búsqueda propuesta en este caso es la consulta del término compuesto *espigón de encauzamiento* en páginas que el traductor pretende que sean fiables y para ello restringe la búsqueda a páginas con dominio .edu con la intención de que se recuperen páginas de universidades, instituciones estatales, científicas y otros organismos, por lo que introduce la consulta *site:.edu “espigón de encauzamiento”* en el buscador *Google*.

También en *Google*, se puede hacer la búsqueda de “*espigón de encauzamiento*” con la opción de búsqueda de imágenes para recuperar fotografías que permitan al traductor conocer el tipo de espigón al que se refiere.

Otra estrategia sería emplear la consulta “*sediment trap*” *glosario ingeniería site:.es* para acceder a glosarios de ingeniería en los que aparezca un equivalente de *sediment trap* en español.

Estas autoras también hablan de la utilización de una estrategia de *aproximación onomasiológica* para encontrar el término en LM en los casos en los que no se encuentra tan fácilmente un término en un glosario o en un diccionario. En esta ocasión la estrategia sería aproximarse a la información poco a poco, primero buscando una definición del término en LO que ayude a buscar el equivalente en LM. Por ejemplo, podría consultarse en *Google “sediment trap” glossary* para acceder a la definición *a device for trapping or otherwise preventing silt or sediment from entering a stream*; con la búsqueda de una imagen también se podría entender mejor el concepto y, una vez identificado el concepto en LO, se puede buscar un glosario en español mediante la consulta “*sediment trap*” *glosario OR vocabulario OR diccionario*, o también accediendo a recursos visuales como diccionarios de ingeniería con dibujos o gráficos.

#### 4.4.11 Enríquez Raído (2011)

Enríquez Raído (2011) presenta un estudio sobre el comportamiento en la búsqueda de información (*information behaviour*) de un grupo de estudiantes de traducción. Examina las tareas de búsqueda en la Web, que consisten en una o más sesiones de búsqueda en Internet para resolver un problema de traducción.

Los participantes del estudio son cuatro estudiantes de traducción del Centre for Translation and Interpreting Studies de la University of Auckland, en el marco de un curso de traducción científico-técnica de 12 semanas de duración en un semestre. Dos de las estudiantes son de Nueva Zelanda (con L1 inglés y L2 español), una de Taiwán (con L1 inglés, L2 mandarín y L3 español) y una de Rusia (con L1 ruso, L2 inglés y L3 español).

Se emplean distintas metodologías, por un lado, el estudio de casos mediante la observación directa y la monitorización y, por otro, la investigación por encuestas, consistente en dos cuestionarios y una entrevista. Posteriormente, se procede a la triangulación de los resultados obtenidos con cada método.

En cuanto al estudio de casos, se comparan las tareas de búsqueda en Internet en la traducción del español al inglés de dos textos científicos (de 200 palabras de media). Según la autora se observa a las estudiantes en su entorno natural, la clase de traducción y se graban sus actuaciones con el programa *BB FlashBack*, que está instalado en los ordenadores que utilizan en clase y puede grabar toda la actividad que aparece en pantalla, incluyendo las búsquedas de los estudiantes en Internet, así como los movimientos y clics del ratón, atajos de teclado utilizados, etc.

Los dos cuestionarios se administran en línea, uno de ellos está destinado a obtener información sobre el perfil de los usuarios (sexo, edad, calificaciones, lenguas de trabajo, conocimientos y experiencia en traducción y búsquedas en Internet); en el otro cuestionario se pide a los participantes que hagan un informe sobre sus conocimientos temáticos del texto original y la descripción de las tareas de búsqueda efectuadas en Internet para resolver problemas de traducción. La entrevista es semiestructurada y se emplea para ampliar información.

Antes del estudio principal, se efectúa un estudio piloto en el que se evalúan todas las herramientas que van a utilizarse, menos la entrevista. Las pruebas las completan un estudiante de doctorado en traducción, que realiza traducciones profesionalmente de forma esporádica y una docente de traducción con 15 años de experiencia.

En cuanto a los resultados, respecto al uso de recursos en línea y electrónicos para la traducción, de los sujetos con más experiencia (el estudiante de doctorado y la docente que participan en el estudio piloto) uno de ellos menciona la enciclopedia en línea *Wikipedia*, el diccionario monolingüe en línea de *Merriam-Webster*, el diccionario bilingüe *PC Bibliothek*, el diccionario multilingüe *Leo.org*, el programa de traducción asistida *Trados*, el programa de gestión terminológica *MultiTerm*, archivos electrónicos, programas de buscar y reemplazar y de análisis de corpus; el otro participante se refiere a los diccionarios en línea de *WordReference*, diccionarios impresos sin especificar, los glosarios de *ProZ* y textos paralelos de Internet. De las cuatro estudiantes de traducción que participan en el estudio principal, algunas aluden a recursos que han utilizado en el contexto académico (diccionarios de *WordReference*, el diccionario en línea *Dictionary.com*, diccionarios monolingües de *Oxford*, *Collins* y *DRAE*, así como personas). Un factor determinante en la elección de los recursos parece ser el acceso fácil y rápido, este es el argumento que esgrime una de las estudiantes para elegir los diccionarios de *WordReference* en lugar de *Voyager*, accesible a través del catálogo de la biblioteca de la University of Auckland.

En otra pregunta sobre los instrumentos de traducción empleados, se refieren a los siguientes: uno de los sujetos del estudio piloto, textos paralelos, diccionarios, bases de datos, expertos, sistemas de traducción asistida, programas de localización; el otro participante del estudio piloto, memorias de traducción, recursos en línea o no, buenos procesos de postedición; de las estudiantes que participan en el estudio principal, una de ellas menciona diccionarios monolingües y bilingües, Internet, otros textos relacionados, consulta a otros traductores, hablantes de las dos lenguas que intervienen en el proceso de traducción; otra estudiante se refiere a Internet, diccionarios, páginas web de traducción, expertos; otra participante incluye diccionarios en papel, programas de traducción, recursos en línea; la última alude a Internet para obtener documentos del campo, consulta a personas o autoridades.

Respecto a las estrategias de búsqueda utilizadas para resolver problemas de comprensión, se pide a los sujetos que prioricen las siguientes estrategias (consulta del diccionario bilingüe, tratar de inferir el significado por el contexto, consulta del diccionario monolingüe, otros). Los resultados se muestran en la Tabla 27<sup>103</sup>.

---

<sup>103</sup> Los nombres de los sujetos proporcionados en la tabla son inventados por la autora.

	Bilingual dictionary	Contextual meaning	Monolingual dictionary	Other
Bob	3	1	2	<i>Search for expression in context (parallel texts)</i>
Daniel	1	3	2	
Laura	2	1	3	
Maria	3	1	2	
Anna	1	3	2	<i>Look on the internet and try to see other usage of the expression to better understand the expression</i>
Martha	2	1	3	<i>I have a lot of hispanohablante [Spanish-speaking] friends and I often consult with them to try and understand a phrase or term or double check I understood it in the first place</i>

**Tabla 27. Estrategias para resolver problemas de comprensión (Enríquez Raído 2011: 258).**

En cuanto a las estrategias de búsqueda utilizadas en la resolución de problemas de formulación, los participantes deben establecer el orden que seguirían (consulta del diccionario bilingüe, intentar expresar por uno mismo la idea del texto origen de diversas formas en la LM, consulta del diccionario monolingüe en LM, otras estrategias).

	Bilingual dictionary	Own renditions	Monolingual dictionary	Other
Bob	1	2	3	<i>Choice I would be to confirm my tentative solutions on the Internet (e.g. in newspaper texts etc.)</i>
Daniel	3	1	3	<i>Use a tool like the synonym list in MS Word (which can be accessed by right-clicking on the words involved in the phrase; for me, this process can sometimes lead to an inspiration</i>
Laura	2	3	1	
Maria	2	1	3	
Anna	1	2	3	
Martha	2	3	1	<i>I am a big fan of dichos [sayings] so I would try to find something equivalent in English even if it meant digressing from the original ie to retain the sentiment rather than a word for word translation</i>

**Tabla 28. Estrategias para resolver problemas de formulación (Enríquez Raído 2011: 261).**

Una de las estudiante que opta por el diccionario bilingüe como primera opción en la formulación señala que el principal motivo es la rapidez: «[...] ...so if I had to translate something, if I want to know what the word was, like its original meaning and try to guess what the equivalent would be in English I would use that [refiriéndose al diccionario monolingüe de la RAE]. But if I need to just get the translation of that word fast, then I just use *WordReference*» (Enríquez Raído 2011: 262).

Sobre los buscadores empleados para las búsquedas, todos los participantes indican que *Google* es el que más utilizan y tres de ellos señalan que *Yahoo!* es el segundo más utilizado, otros buscadores mencionados son *Baidu* (buscador chino), *Yandex* y *Rambles* (buscadores rusos) y los metabuscadores *Metacrawler* y *Dogpile*.

En cuanto a las búsquedas con operadores, los participantes mencionan los siguientes operadores y funciones. AND o + para forzar que una palabra aparezca en los documentos recuperados o para asegurarse de una ortografía determinada, las comillas para buscar frases fijas, NOT o - para excluir palabras de la búsqueda, NEAR para buscar palabras clave que tienen que aparecer cercanas, ADJ cuando tienen que aparecer justo al lado, define: para buscar definiciones de una palabra, site: para buscar en determinados dominios o páginas de un determinado país, también el asterisco para buscar palabras incompletas.

Una de las tareas que esta investigadora encarga a los participantes de su experimento está enfocada a observar cómo efectúan búsquedas temáticas para documentarse sobre la migración de las ballenas grises en el Océano Pacífico y la asistencia sanitaria por maternidad. Se analiza la complejidad de la búsqueda (búsqueda simple o avanzada), el tamaño (número de términos) y la eficacia (obtención de resultados relevantes).

En lo que respecta a la complejidad, los traductores más avanzados efectúan búsquedas avanzadas con una o más palabras claves combinadas con operadores, incluyendo site: y define:. Una de las estudiantes realiza tanto una búsqueda simple sin operadores, como una búsqueda completa de una frase entre comillas. Otra estudiante utiliza el operador + delante de todas las palabras de búsqueda, aunque no era necesario en *Google*.

En cuanto al tamaño, la media de términos empleados en una búsqueda es 4,5 excluyendo artículos, preposiciones y signos de puntuación, excepto en los casos de búsqueda de frases entrecomilladas.



Las cuatro estudiantes admitieron que no sabían utilizar la búsqueda avanzada. Según la autora, la inexperiencia de las estudiantes de traducción en el uso de la búsqueda avanzada parece coincidir con otros usuarios de Internet tal y como se afirma en Battelle (2006: 273)<sup>104</sup>.

En cuanto a la selección de los resultados, algunas estudiantes basan la elección de un término frente a otro en criterios estadísticos (número de resultados que aparecen en el buscador), una conclusión a la que según (Enríquez Raído 2011: 292) también llega Gile (2004).

También las estudiantes afirman fijarse para la selección de resultados en el título de las páginas web (que aparece en la primera línea de cada resultado), el pequeño resumen o descripción donde aparecen las palabras de la búsqueda destacadas y la URL de la página web.

Enríquez Raído (2011: 297-298) se refiere a criterios de fiabilidad de las páginas web. Según su estudio, los participantes no tienen en cuenta el número de enlaces externos e internos de la página, sino más bien la calidad lingüística de la página, la corrección en cuestiones gramaticales y errores, sobre todo ortográficos, etc., también el prestigio de la organización en un campo, comentarios positivos o negativos de la página en otros sitios web. Dos estudiantes se muestran menos críticas respecto a la fiabilidad de las páginas de Internet, por ejemplo una de ellas considera *Wikipedia* y las páginas comerciales fuentes fiables de información, aunque admite por otro lado que las páginas de instituciones son más fiables que las comerciales y que la autoría es un criterio importante a la hora de evaluar una página. Otra estudiante también indica que confía en la información de los foros, como *ProZ* y *WordReference*: «[...] I dunno, it's just psychological. If I get to talk to someone directly and they respond directly to my question, I feel like it's more reassuring, rather than just going for a random answer. (laughter)» (Enríquez Raído 2011: 302).

---

<sup>104</sup> Otros estudios como el de Spink y Jansen (2005: 77), que examinan las consultas introducidas por los usuarios en *AlltheWeb.com*, *Altavista* y *Excite*, demuestran que el uso de operadores es del 10% siendo las comillas el más utilizado, según los datos proporcionados por *Google* esta cifra se rebaja al 5% en lo que respecta al uso de la búsqueda avanzada. Datos disponibles en *Google Web Search Help*: <http://www.Google.com/support/websearch/bin/answer.py?hl=en&answer=136861>

Baeza-Yates y Ribeiro-Neto (1999: 389-390) aportan datos similares: el 80% de las búsquedas introducidas en buscadores de Internet no utilizan operadores booleanos, el 25% utilizan una sola palabra de búsqueda y la media es de dos o tres palabras de búsqueda, el 80% de los usuarios no modifica la búsqueda y el 85% examinan y acceden solo a los primeros resultados.

En cuanto a la clasificación de la información que los traductores encuentran en Internet, los marcadores son el procedimiento utilizado por todos los participantes menos uno (una estudiante copia y pega la información relevante en un documento de Word), algunos clasifican por carpetas la lista de favoritos. El traductor más avanzado del experimento también descarga los archivos de Internet y los guarda en carpetas específicas en su disco duro, lo que sugiere que compila corpus electrónicos. Ninguno de los participantes se refiere a un nuevo tipo de clasificación e intercambio de enlaces, las *folksonomies*, una alternativa a las ontologías de la Web Semántica. Ejemplo de ello, según la autora, son las aplicaciones *Flickr* (un programa de gestión e intercambio de fotografías) y *del.icio.us* (una página en la que se comparten marcadores).

En el experimento se analizan también las necesidades de búsqueda, se identifican un total de 81, de las cuales más de la mitad corresponden a estudiantes que no tienen confianza con las lenguas de la traducción.

Se observan búsquedas para obtener información fraseológica, la confirmación de colocaciones, por ejemplo la colocación *armas predilectas*, problemática para todos los sujetos, también el término especializado *transgénicos*, las siglas *OMG* (organismos modificados genéticamente) que algunos estudiantes desconocen y no se explicita en el texto y otros desconocen el equivalente en inglés, otras necesidades de información son temáticas. Todos los problemas mencionados están relacionados con problemas de comprensión o de significado y el desconocimiento de un equivalente (Enríquez Raído 2011: 344).

Enríquez Raído (2011: 352-353) presenta unas tablas donde indica las necesidades de información y los objetivos de búsqueda a los cuales asigna unos códigos que presentamos aquí entre paréntesis, en los que identifica *armas predilectas* (búsqueda de una definición, búsqueda de un equivalente), *transgénicos* (definición + equivalente, uso), *guía roja y verde* (significado contextual, equivalente), *OMG* (resolución del acrónimo, abreviatura o sigla; resolución del acrónimo, abreviatura o sigla + equivalente), *cultivo* (equivalente), *conjunto* (significado contextual), *alcance* (equivalente, significado contextual), *ganadería* (equivalente), *malas hierbas* (léxico-sintáctico, equivalente, significado contextual, equivalente + significado contextual), *supone incremento* (significado contextual, equivalentes), *imprevisible* (definición), *agentes tóxicos* (definición, equivalente, uso), *pasa por* (significado contextual), *reforzar* (equivalente), *resistencia* (definición, equivalente, ortografía), *constituir* (definición, equivalente).

Respecto a la elección de recursos el 74,6% de las búsquedas se efectúan en direcciones conocidas, en este caso de los diccionarios *WordReference* y el *DRAE*, el 18,6% en buscadores (en la casilla de búsqueda de google.com o google.co.nz y una de ellas en la búsqueda avanzada de *Google*), una búsqueda directa en la base de datos de una página (en *Wikipedia* o en el sitio web de Greenpeace) y también búsqueda por navegación o *navigational query* (en *WordReference*). Los traductores con menos experiencia suelen empezar la búsqueda por aquello que conocen, diccionarios en línea principalmente, mientras que el traductor con más experiencia, prefiere empezar las búsquedas por lo desconocido<sup>105</sup>, con consultas en motores de búsqueda.

Todos los estudiantes buscan en el diccionario como la primera fuente de información, independientemente de la necesidad, por lo que en la mayoría de los casos hacen un mal uso de ellos, una de las estudiantes busca la colocación *armas predilectas* en un diccionario monolingüe (*DRAE*) y otros cuatro en un diccionario bilingüe (*WordReference*). Lo mismo sucede con las siglas *OMG*, a excepción de una de las estudiantes que busca las siglas en google.co.nz. Una de las estrategias de búsqueda en el diccionario es dividir expresiones complejas en cada uno de sus componentes para buscarlos en el diccionario, como en el caso de *supone incremento* que buscan dos de las estudiantes.

#### 4.4.12 Resumen de estrategias de búsqueda en Internet detectadas en los estudios

A partir del análisis sobre estrategias de búsqueda de los traductores en Internet, podemos concluir que Internet es uno de los recursos más utilizados por los traductores y que se encuentra presente en la mayoría de las estrategias de búsqueda.

Las herramientas de búsqueda más utilizadas por los traductores para la realización de búsquedas en Internet son los buscadores, sobre todo *Google*, por encima de otras herramientas como los directorios. Asimismo, se menciona la búsqueda en recursos específicos en línea: enciclopedias, sobre todo *Wikipedia*, diccionarios y glosarios. En

---

<sup>105</sup> Enríquez (2011: 271) refiriéndose a Battelle (2006) indica que mientras que una búsqueda por navegación se emplea para encontrar *lo que se conoce*, una búsqueda por palabras clave persigue encontrar *lo que no se conoce*, en palabras de Battelle (2006: 32) «[...] Web blindness: a sense that we know there's a stuff we might want to find, but have no idea how to find it. So we search in the hope it will somehow find us».

algunos estudios también se alude a la consulta en foros, sobre todo *ProZ*, aunque no se profundiza en esta búsqueda (véase Enríquez Raído 2011; Walter 2006).

Internet resulta útil al traductor tanto para una rápida documentación general sobre un tema, como para la posterior búsqueda lingüística, para el acceso a terminología novedosa que no se encuentra en otros recursos que no evolucionan tan rápidamente, como los diccionarios, e incluso los corpus.

Dado que el principal uso de Internet es el de un gran corpus, las estrategias de búsqueda se asemejan en gran medida a las vistas en los corpus, aunque difieren ligeramente debido a la naturaleza de los buscadores de Internet que, como hemos visto, no están diseñados para la realización de búsquedas lingüísticas.

En este sentido, las búsquedas en Internet no son tan directas como en los corpus, en la mayoría de los casos requieren de un mayor número de acciones de búsqueda por lo que podríamos denominar el método de ensayo y error. El traductor prueba con una consulta y en función de los resultados obtenidos redefine la búsqueda.

De este análisis hemos observado que los traductores no efectúan una sola consulta cuando tratan de encontrar algo en Internet, sino que la búsqueda consta de varias consultas sucesivas, que suelen comenzar con consultas más sencillas. El inconveniente principal que plantea este tipo de búsqueda es la presencia de ruido en los resultados. Sin embargo, puede resultar eficaz para aproximarse de una manera general a lo que se está buscando y aplicar después otros tipos de búsqueda.

Posteriormente, el traductor puede redefinir la búsqueda con consultas más avanzadas que pueden implicar la consulta combinada de varios términos, de frases exactas, la exclusión de determinadas palabras, la aplicación de filtros como la restricción de los resultados según criterios como el idioma o el formato de los textos, etc. Hemos visto que en el caso de los estudiantes de traducción, las búsquedas más avanzadas no siempre se producen o no proporcionan los resultados deseados (Enríquez Raído 2011).

En ocasiones Internet puede ser útil para obtener el significado de determinados términos en LO que no se incluyen en el diccionario. Por ejemplo, el uso de comillas combinado con expresiones semánticas del tipo «X is» o «X es», «X means» o «X significa», etc. facilita la recuperación de contextos definitorios (Bergeron y Van Steenberg 2000). Los traductores también utilizan la opción de búsqueda *define* de *Google* (Enríquez Raído 2011).

Asimismo, la información más demandada en los buscadores de Internet son equivalentes de traducción, así como fraseología y colocaciones en LM. La principal estrategia de búsqueda detectada en Internet es de tipo inductivo, como con los corpus, el traductor suele partir de una intuición o hipótesis que verifica en Internet, lo que produce una selección o un rechazo del equivalente. En la búsqueda deductiva, cuando el traductor no cuenta con una hipótesis inicial de traducción, las estrategias también son similares a las de los corpus y se efectúan a partir de información del contexto. En este sentido, los traductores realizan búsquedas en los buscadores a partir de la introducción de palabras clave relacionadas.

En las búsquedas de acrónimos, abreviaturas o siglas, los operadores de inclusión de palabras sirven al traductor para que el sistema de búsqueda tenga en cuenta determinadas palabras que por defecto ignoraría. Algunas siglas son iguales morfológicamente a palabras vacías de significado, por ejemplo las siglas AS (*Associate in Science*) (Gambín 2004).

En la búsqueda de variaciones de una palabra (en singular y plural, *bird* OR *birds*), el operador OR recupera documentos que contienen cualquiera de las palabras introducidas; en otros recursos, como corpus y diccionarios, el traductor utiliza la técnica del truncamiento (*bird\**) para acceder a variaciones de una palabra (Bergeron y Van Steenberg 2000). Sin embargo, la mayoría de buscadores no permiten el uso de comodines para truncar palabras, sino para sustituir palabras completas. Asimismo, dada la enorme cantidad de resultados que pueden incluirse en una búsqueda en Internet, el uso de OR en las consultas evita que, además de la búsqueda de una palabra en singular y en plural, se incluyan otras que puedan provocar demasiado ruido en los resultados de la consulta. Por ejemplo, cuando se consulta *cit\**, el buscador recuperará los términos *city* y *cities*, pero también otras palabras, como *citrus* o *citizen*, que pueden no estar relacionadas con la búsqueda. En el caso de que la consulta genere demasiado ruido, el traductor emplea el operador NOT para excluir palabras de la consulta.

En la búsqueda de términos compuestos y expresiones donde se desconoce el orden exacto de combinación de los términos, los operadores de *proximidad* o *posicionales* permiten recuperar los documentos en función de la proximidad de los términos que aparecen en la ecuación de búsqueda.

En la verificación de términos compuestos o expresiones, el uso de las comillas es una de las opciones más utilizadas. Esta estrategia suele combinarse con la restricción de la consulta a un determinado dominio (por ejemplo un país determinado) o página de Internet (una página oficial, con prestigio, o la página web del cliente).

La búsqueda restringida a páginas oficiales o fiables se utiliza cuando un equivalente aparece en un número elevado de páginas de Internet, pero no es seguro que sea el término correcto en LM. No se trata de una duda absurda, ya que se ha demostrado que las estadísticas sobre frecuencia de aparición de una palabra en las páginas de la red no son motivo suficiente para asegurar que el término esté normalizado.

Gómez (1999) recomienda la herramienta *site search*. Un ejemplo sería restringir la búsqueda de una expresión a páginas de la Unión Europea, *site:europa.eu.int*. La Unión Europea es una organización que suele publicar la mayoría de sus documentos en todas las lenguas de los países miembros. Por tanto, su página web resultará útil cuando queramos consultar versiones de documentos en varios idiomas.

Una estrategia muy utilizada en Internet para la búsqueda de equivalencias es la llamada búsqueda inversa (Sevilla Muñoz y Sevilla Muñoz 2005a, 2005b) o búsqueda mediante la opción de lenguas (Bergeron y Van Steenberg 2000). Consiste en introducir en un buscador el término en la LO y seleccionar la opción de búsqueda de páginas en la LM. El sistema recuperará, entre otros, páginas de la LM en las que aparecerá el término en la LO junto con una explicación o un equivalente, también páginas bilingües o multilingües en las que aparezca el equivalente en contexto, o incluso diccionarios o glosarios que ofrezcan el equivalente en LM. Como con todas las estrategias de búsqueda, la evaluación del equivalente en páginas originales es imprescindible.

Otra estrategia que hemos observado en la búsqueda de equivalentes de nombres de plantas o animales, es la búsqueda de equivalentes a partir del latín como puente entre dos lenguas. Las plantas y los animales suelen tener una denominación científica en latín, que es válida para todas las lenguas. La estrategia de búsqueda consiste en encontrar la palabra en latín a partir de la denominación en LO para después buscar la denominación latina en páginas de la LM. Por ejemplo, si se busca un equivalente en español para un término en inglés podría introducirse en un buscador [término en inglés +latín]. Posteriormente, el término latino encontrado se buscaría en páginas en español (Bergeron y Van Steenberg 2000).

La búsqueda por imágenes y archivos multimedia es de gran utilidad para el traductor y esta es una de las ventajas que aporta Internet frente a otros recursos de búsqueda como los corpus, que no incorporan imágenes. Tal y como afirman Bergeron y Steenberg (2000: 28) «Sometimes a picture can be worth quite a few translated words». Los autores recomiendan recurrir a la búsqueda de imágenes en línea para comprender el

significado de los términos o, en la fase de verificación, para comprobar si dos términos son o no equivalentes.

Los buscadores también se utilizan con frecuencia en traducción para la búsqueda de terminología directamente en recursos léxicos en línea. Budin y Wright (1997: 847) explican que el primer lugar en el que el traductor puede buscar una definición o un equivalente es en un diccionario electrónico utilizando un buscador: «The first place to search for definitions and even equivalents for a given set of terms can involve searching online terminological resources. There are a number of specialized search engines, indexes and portals that provide either direct access to online dictionaries and glossaries or bibliographical information on hardcopy lexicographical and terminological collections».

Existen distintos métodos de consulta en Internet para la localización de diccionarios en línea o para la búsqueda de términos y definiciones en diccionarios y páginas de la Web. Los autores no recomiendan guardar una lista interminable de glosarios, sino solo aquellos que realmente sean útiles para el traductor, debido a la facilidad de encontrar glosarios en línea una vez se conocen las estrategias de búsqueda necesarias para ello.

Hay que saber escoger las palabras clave que introduciremos en el motor de búsqueda. Por ejemplo, para la búsqueda de diccionarios en general, utilizaremos la ecuación de búsqueda [+área temática +diccionario\* +lengua 1 +lengua 2]. Variaciones de esta fórmula nos devolverán diccionarios del área temática que indiquemos, monolingües, bilingües o multilingües, en función del número de lenguas que introduzcamos (una, dos, o más). Un ejemplo de búsqueda directa de un término en un diccionario monolingüe es la siguiente [+término +diccionario\* +lengua 1] y en un diccionario bilingüe o multilingüe [+término +diccionario\* +lengua 1 +lengua 2 +lengua 3]. Cuando la consulta con la palabra clave *diccionario* o *dictionary* no recupera resultados satisfactorios, se recomienda el uso de sinónimos como *léxico*, *glosario*, *enciclopedia*, etc.

## 4.5 *Recapitulación*

En este capítulo hemos tratado una serie de investigaciones sobre las estrategias de búsqueda de los traductores, principalmente en diccionarios, corpus e Internet. A continuación reflexionaremos, en primer lugar, sobre los métodos de análisis que se han utilizado en los distintos estudios con el fin de decidir cuál de ellos nos resultaría más interesante para nuestro estudio sobre las estrategias de búsqueda onomasiológica de los traductores. En segundo lugar, presentaremos una clasificación preliminar de estrategias de búsqueda que hemos elaborado a partir de la síntesis de las estrategias de búsqueda propuestas en la bibliografía reseñada y que nos servirá como punto de partida de nuestro análisis sobre las estrategias de búsqueda onomasiológica.

### 4.5.1 *Métodos de análisis de las estrategias de búsqueda*

Los métodos que se han utilizado en otras investigaciones para el estudio de las estrategias de búsqueda de los traductores han sido diversos. Cada metodología tiene unas características y cada uno de los instrumentos de medida utilizados aporta una serie de ventajas y desventajas en función de diferentes parámetros: si utilizan una metodología cualitativa o cuantitativa, el modo de acceso a los sujetos, el tiempo invertido por el investigador, la facilidad de análisis de los datos y la relevancia y representatividad de los resultados. La reflexión sobre las ventajas aportadas por cada uno de los métodos nos permitirá decidir cuál de ellos podría resultar más adecuado para nuestra investigación sobre estrategias de búsqueda onomasiológica en traducción.

Primero, enumeraremos los métodos que han utilizado las investigaciones que hemos resumido en este capítulo e indicaremos qué tipo de información se ha intentado obtener hasta el momento con cada uno de ellos. Segundo, describiremos las características de cada uno de estos métodos y, finalmente, indicaremos cuál consideramos que se adapta mejor a nuestro estudio.

Algunos estudios han adoptado un enfoque más cuantitativo y otros cualitativo. Dentro de los **estudios cuantitativos**, el instrumento más utilizado ha sido la **encuesta o cuestionario**. Con el método de la encuesta se ha obtenido sobre todo información sobre las tendencias en el uso de diccionarios<sup>106</sup> de los traductores y otros grupos de sujetos (nativos o estudiantes de segundas lenguas). Véanse a modo ilustrativo los estudios de

---

<sup>106</sup> Véase Wiegand (1998: 583 y siguientes).



Corpas Pastor et al. (2001), De la Rosa Martín (2003), Sánchez Ramos (2004b), en los que participan estudiantes de traducción. Por ejemplo, hemos visto estudios que tratan de cuantificar el uso del diccionario monolingüe frente al bilingüe. También se ha empleado este método para calcular el tipo de información más consultada en los artículos de un diccionario (etimología, definiciones, equivalentes, etc.). Asimismo, estos estudios comparan la frecuencia de uso del diccionario frente a otros recursos.

De los **estudios cualitativos**, algunos autores tratan de describir las estrategias de búsqueda de los traductores basándose en su **propia experiencia** como traductores o como docentes en traducción (por ejemplo, Roberts y Bossé-Andrieu 2006; Sánchez-Gijón 2009; Tarp 2002; Varantola 2006). Otros emplean **métodos experimentales** con el uso de **técnicas de observación, anotaciones en plantillas** (Atkins y Varantola 1997), **Thinking-Aloud Protocols** (TAP) y **monitorización**, entre otros. Los TAP han sido el método de investigación más utilizado tradicionalmente en los estudios de traducción<sup>107</sup>, con el fin de conocer las estrategias utilizadas para la resolución de problemas en el proceso de traducción (véanse a modo de ejemplo las obras de Bell 1991; Jääskeläinen y Tirkkonen-Conditt 1991; Krings 1986b; Lörscher 1991; Lörscher 1992, a las que nos hemos referido en el capítulo 2). Un ejemplo de estudio en el que se utilizan los TAP para analizar las estrategias de búsqueda de los traductores, en este caso en los diccionarios, es el efectuado por Fraser (1999).

La monitorización se utiliza, entre otras, en las investigaciones del grupo PACTE (2002; 2005a; 2009; 2011) para medir la competencia traductora, por ejemplo en el estudio de la subcompetencia instrumental.

Las **entrevistas** suelen utilizarse como método complementario de obtención de información en algunos estudios (Enríquez Raído 2011; Hartmann 1999).

En cuanto a las **particularidades de cada uno de los métodos** empleados, la **encuesta** es una metodología que persigue cuantificar y comparar una determinada información. En la mayoría de estudios que hemos examinado se utilizan cuestionarios escritos con preguntas cerradas y respuestas predeterminadas por el investigador, la mayoría de las veces por medio de preguntas de respuesta múltiple.

Las investigaciones que han seleccionado la encuesta como instrumento de medida observan las ventajas de que permite obtener un gran número de datos (método

---

<sup>107</sup> Para más información sobre el uso de los TAP en los estudios de traducción, véase Orozco (2002).

cuantitativo) en menos tiempo; el acceso a los sujetos es relativamente sencillo, ya que no requiere presencia del investigador, cada informante rellena la encuesta por sí mismo; el análisis de los datos es más rápido (siempre que las respuestas sean predeterminadas), ya que permite el uso de programas de análisis estadísticos, como SPSS (Corpas Pastor et al. 2001: 245); y algunos autores que utilizan encuestas sugieren que el método cuantitativo es más objetivo y, por ende, «más científico».

Sin embargo, otros investigadores indican que el método cuantitativo no tiene por qué ser más objetivo o científico que el cualitativo sino que, por el contrario, en el método cuantitativo el investigador puede acabar imponiendo al sujeto, sin que sea su intención, su propia posición respecto a los aspectos investigados. En cambio, en la técnica cualitativa es el sujeto el que marca con su discurso su perspectiva, lo que puede llevar a descubrir nuevas vías para la comprensión del fenómeno investigado (Olaz 2008: 18).

En los **métodos experimentales** utilizados en traducción los sujetos efectúan una tarea que simula una situación real, por ejemplo la realización de una traducción, y pueden emplearse distintos instrumentos de recogida de la información, por ejemplo los TAP y la monitorización.

Por un lado, los **TAP**<sup>108</sup> son un método en el que los sujetos verbalizan sus procesos mentales mientras efectúan la tarea (método introspectivo), o una vez la tarea ha finalizado (método retrospectivo). Estas verbalizaciones normalmente se graban, se transcriben y se analizan.

Por otro lado, en la **monitorización** se graba a los sujetos mientras efectúan la tarea. Posteriormente, el investigador analiza las grabaciones en función de unos parámetros.

Se suele argumentar que los datos obtenidos en la observación y monitorización son más fiables que los de encuestas y entrevistas, ya que en los experimentos se recaba información sobre lo que los sujetos hacen realmente, no de lo que creen que hacen o de lo que creen que deberían contestar.

Sin embargo, se podría alegar, como ya han hecho otros autores antes que nosotros, que los experimentos en los que los participantes deben traducir un texto tampoco pueden considerarse situaciones reales de traducción ya que el investigador selecciona el texto y los problemas, por tanto, de alguna manera también puede influir en los resultados. En este

---

<sup>108</sup> Para más información sobre la metodología de los TAP, véase por ejemplo Wiegand (1998: 1010-1022).

caso los resultados se ven limitados al texto o tipo de problemas que plantea el investigador en el experimento.

Entendemos **entrevista** como un método de investigación esencialmente cualitativo consistente en una conversación oral. Como en los estudios con traductores que hemos examinado la entrevista se ha empleado como método para obtener información complementaria (en algunos casos), la mayoría no especifica los detalles sobre esta metodología (diseño de las entrevistas, realización de las mismas, análisis de los datos, etc.). En muchos estudios solo se hace una breve alusión a que se han realizado entrevistas.

No obstante, cabe decir en este punto que la entrevista es uno de los instrumentos empíricos más utilizados en las investigaciones sociales «como el camino de acceso al conocimiento de los fenómenos sociales» (Olaz 2008: 26). Según este autor, en la entrevista el investigador suele optar por preguntas principalmente abiertas, más útiles cuando se pretende centrar el tema de estudio y cuando se trata de indagar los conocimientos de un experto sobre un determinado tema, donde al acotar las preguntas, el entrevistado no podría «aportar matices cualitativos fundamentales» (Olaz 2008: 39).

La entrevista también presenta inconvenientes. Primero, la entrevista requiere la búsqueda de informantes, el contacto personal, un esfuerzo tanto por parte del investigador como de los informantes, ya que en muchos casos es preciso el desplazamiento del investigador y es necesario acordar citas en un lugar, espacio y tiempo que venga bien a ambos para hacer la entrevista. Esto implica que en las entrevistas el contacto con los sujetos pueda resultar más difícil.

Otro problema de la entrevista es el análisis de los datos (más complicado por métodos cualitativos que cuantitativos). Aunque, para gestionar el análisis de los datos los programas de análisis cualitativo pueden resultar de gran ayuda.

En la Tabla 29 incluimos un resumen de las ventajas y desventajas de las metodologías mencionadas.

	ENCUESTA	EXPERIMENTO	ENTREVISTA
Método cualitativo/cuantitativo	Método más cuantitativo que cualitativo. Obtención de más datos en menos tiempo.	Método más cualitativo que cuantitativo. Obtención de información más relevante.	Método más cualitativo que cuantitativo. Obtención de información más relevante.
Sujetos	Acceso más sencillo a los sujetos. No requiere presencia.	Acceso más difícil a los sujetos. Mejores sujetos, ya que están preseleccionados (expertos).	Acceso más difícil a los sujetos. Mejores sujetos, ya que están preseleccionados (expertos).
Tiempo dedicado	Requiere menor tiempo dedicado por el investigador en el trabajo de campo. Envío de las encuestas a foros o listas de traducción. Los sujetos rellenan la encuesta por sí mismos.	Requiere mayor tiempo dedicado por el investigador en el trabajo de campo. Búsqueda de sujetos que quieran participar en el experimento. Acuerdo con el sujeto de fecha, lugar y hora para hacer el experimento. Puede requerir presencia y desplazamiento del investigador.	Requiere mayor tiempo dedicado por el investigador en el trabajo de campo. Búsqueda de sujetos que quieran contestar a la entrevista. Acuerdo con el entrevistado de fecha, lugar y hora para hacer la entrevista. Requiere presencia y puede requerir desplazamiento del investigador.
Análisis de los datos	Análisis más rápido y sencillo de los resultados con programas estadísticos (si las respuestas son predeterminadas).	Análisis más lento y complejo de los resultados. Puede solucionarse con el uso de programas de análisis cualitativo	Análisis más lento y complejo de los resultados. Puede solucionarse con el uso de programas de análisis cualitativo.
Relevancia y representatividad de los resultados	¿Datos menos reales o fiables que en los experimentos? ¿El encuestado responde lo que cree que hace y no lo que hace realmente, o lo que cree que debe contestar? El uso de preguntas cerradas puede limitar los resultados.	¿Datos más reales o fiables que en las entrevistas? No se trata de una situación real de traducción (sino de una simulación). El uso de un texto determinado puede limitar los resultados.	¿Datos menos reales o fiables que en los experimentos? ¿El entrevistado responde lo que cree que hace y no lo que hace realmente, o lo que cree que debe contestar? Al utilizarse preguntas abiertas, los resultados no están limitados por el investigador.

**Tabla 29. Resumen de las ventajas e inconvenientes de los instrumentos de recogida de datos.**

En nuestro estudio, que presentamos en el siguiente capítulo, pretendemos profundizar y conocer lo máximo posible las estrategias de búsqueda onomasiológica en el proceso de traducción, y no nos interesa tanto contabilizar el número de veces que se utiliza un tipo de búsqueda o un recurso en comparación con otros. Además, este estudio es una primera aproximación a las búsquedas onomasiológicas en traducción, en el que apenas contamos con ideas preconcebidas. De ahí que precisemos una metodología más cualitativa que cuantitativa. Por ello, descartamos el cuestionario cuantitativo con preguntas cerradas como instrumento de recogida de datos.

De los instrumentos de tipo cualitativo que podemos emplear, consideramos que la realización de un experimento limitado a la traducción de un texto con unos problemas concretos tampoco resulta el método más adecuado, especialmente porque uno de los objetivos principales de nuestra investigación es descubrir cuáles son los problemas o las situaciones que motivan la realización de búsquedas onomasiológicas en traducción. Por ejemplo, sería difícil encontrar un texto de una longitud razonable en el que hubiera muchos ejemplos de situaciones que motivaran que el traductor realizara búsquedas onomasiológicas.

Lo que pretendemos es que los traductores reflexionen acerca de las búsquedas que realizan en su actividad de traducción real. Por tanto, concluimos que la entrevista, consistente en una conversación oral y semidirigida entre el entrevistado y el investigador, en la que se efectúan preguntas de respuesta abierta y el traductor no se limita a marcar una opción de respuesta de una lista, nos permitirá obtener más información, pero sobre todo información más relevante porque podremos profundizar en algunas cuestiones relevantes para la investigación, haciendo más preguntas al entrevistado o insistiendo en una pregunta en función de las respuestas obtenidas.

Asimismo, podremos adoptar una serie de medidas en la entrevista para evocar en el traductor una situación real de traducción. Primero, comenzar la entrevista con una introducción a las búsquedas onomasiológicas para situar y refrescar la memoria al traductor y, antes de formular las preguntas, mostrar ejemplos de fragmentos de texto que podrían suscitar búsquedas onomasiológicas, también para simular o evocar en el entrevistado que recuerde situaciones reales en las que haya tenido que realizar este tipo de búsquedas.

Además, desde la primera toma de contacto con el entrevistado, cuando se le pregunte si estaría dispuesto a participar en la entrevista, se le explique el estudio y qué información se le va a preguntar, hasta el momento de la entrevista, el traductor puede

reflexionar sobre las búsquedas onomasiológicas en el trabajo, y así proporcionar en la entrevista ejemplos reales. Por otro lado, al finalizar la entrevista, puede mantenerse el contacto con el entrevistado para posibles aclaraciones, y también por si este desea aportar más ideas sobre las búsquedas onomasiológicas una vez finalizada la entrevista, al haber reflexionado sobre las búsquedas onomasiológicas durante su actividad diaria real de traducción los días posteriores a la entrevista.

Finalmente, el número de participantes dispuestos a participar en una entrevista no necesariamente será limitado. En el método de la entrevista, el contacto con el entrevistado es directo y el investigador se encuentra presente en el momento de la conversación por lo que el entrevistado percibe una mayor implicación por parte de este. Por este motivo, pensamos que un informante puede estar más predispuesto a participar en una entrevista que en responder a una encuesta que le llega por e-mail. Además, el entrevistado también se esfuerza en aportar más información relevante en las respuestas.

#### **4.5.2 Propuesta preliminar de clasificación de estrategias de búsqueda en traducción**

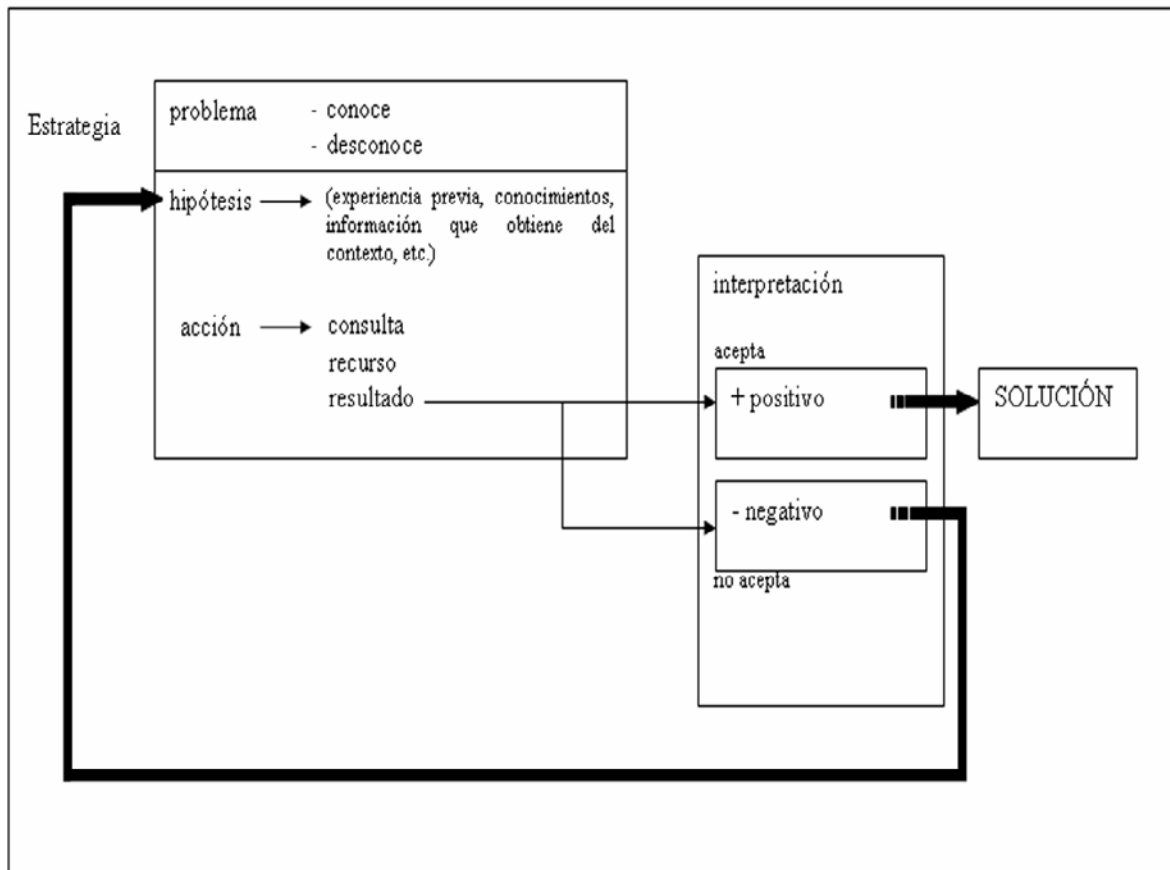
En este capítulo hemos recopilado y resumido una serie de estudios que se basan en un acercamiento a los recursos centrado en los usuarios y que examinan el uso que realizan los traductores, como usuarios, de determinados recursos de referencia, principalmente diccionarios, corpus e Internet.

A partir de los datos aportados en estas investigaciones, hemos analizado y sintetizado en el anexo 2 las búsquedas que se han detectado hasta el momento en traducción. El análisis de las estrategias de búsqueda propuestas en los estudios nos han permitido identificar determinados patrones de búsqueda de los traductores. No obstante, como veremos, los patrones no siempre siguen un modelo fijo. La utilización y combinación de unas estrategias u otras depende de los recursos de los que disponga el traductor, de cada situación de búsqueda y de la propia experiencia del traductor tanto en el tema tratado como en su habilidad para efectuar búsquedas.

A continuación, comenzamos explicando los distintos elementos que hemos identificado en una estrategia de búsqueda (apartado 4.5.2.1). Después, presentamos nuestra clasificación preliminar de estrategias de búsqueda y describimos una a una las estrategias de búsqueda que la componen (apartado 4.5.2.2 y subapartados). Finalmente, aportamos nuestra valoración sobre las estrategias de búsqueda identificadas en los estudios (apartado 4.5.2.3).

#### 4.5.2.1 Elementos que intervienen en una estrategia de búsqueda

En el análisis hemos identificado distintos elementos de una estrategia de búsqueda. En primer lugar, existe un *problema* que es la base de toda búsqueda. El problema consiste en que hay algo que el traductor desconoce. El *desconocimiento* es el que motiva la búsqueda, tal y como se sugiere en Bowker (2000: 22): «[...] if the translator already knows the term, chances are there is no need to look it up». Además, para que haya una búsqueda el traductor también tiene que conocer alguna información, a partir de la cual pueda empezar a buscar. Si el traductor no cuenta con un *conocimiento* que le permita iniciar una búsqueda no hay posibilidad de que la búsqueda se produzca. En resumen, un problema viene determinado por una información conocida y una información desconocida (lo que provoca un problema o una necesidad de búsqueda en el usuario). A continuación, a partir de sus conocimientos y de la información de que dispone, el traductor genera una *hipótesis de solución* que deberá confirmar o rechazar. Esto motiva la puesta en marcha por el traductor de una serie de *acciones de búsqueda* en los recursos, la parte más técnica de estas acciones es la introducción de una consulta en un recurso para obtener un resultado (lo que denominábamos en el capítulo anterior técnica de búsqueda). En la estrategia de búsqueda hay siempre un componente humano, la intención de la búsqueda que determinará tanto la consulta como el recurso seleccionado y los resultados. En una estrategia de búsqueda puede emplearse más de una técnica de búsqueda (combinaciones de consultas, recursos y resultados), tantas como sean necesarias para encontrar aquello que el usuario desconoce. Por último, en una estrategia de búsqueda siempre hay un componente de *evaluación o interpretación* de los resultados, lo que puede llevar a obtener una *solución* en el caso de que el traductor haya encontrado aquello que desconocía, lo que resuelve el problema inicial o, en caso contrario, que propicie nuevas acciones de búsqueda. En la Ilustración 98 incluimos el esquema de nuestro modelo de estrategia de búsqueda.



**Ilustración 98. Esquema de nuestro modelo de estrategia de búsqueda.**

#### 4.5.2.2 Estrategias de búsqueda detectadas en los estudios

En este apartado describimos las estrategias de búsqueda en función de dos parámetros, la necesidad o el problema (que explicamos en el apartado 4.5.2.2.1) y las acciones de búsqueda llevadas a cabo por los traductores para resolver cada problema (apartado 4.5.2.2.2). De las distintas necesidades que puede tener el traductor (temática, estilística, textual, terminológica), vamos a centrarnos en la terminológica, que es el objeto de nuestro estudio. Así, se describen las principales necesidades que han señalado los traductores en el acceso a la terminología y que propician la puesta en marcha de una o más acciones de búsqueda. Además, las estrategias están presentadas por el orden de utilización más habitual de los traductores, según se ha detectado en el análisis.

##### 4.5.2.2.1 Las necesidades o problemas

Para designar las necesidades o problemas utilizaremos la terminología que ya se ha empleado en los estudios de traducción. Las necesidades identificadas son principalmente de dos tipos, de comprensión en la lengua origen (LO) y de expresión en la lengua meta (LM). Dentro de estas necesidades, hemos subdividido necesidades de recepción,



transferencia y producción; y también hemos señalado necesidades de verificación. A continuación, nos referimos a cada una de ellas.

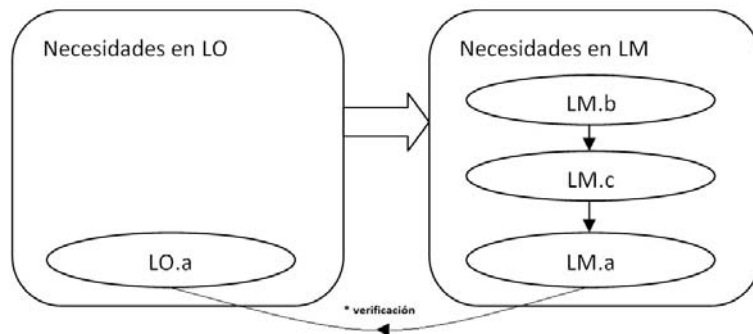
- En la LO, hay una necesidad sobre todo de recepción: que se caracteriza por que el traductor conoce una denominación de la LO pero desconoce su significado.

- En la LM, se produce una necesidad de transferencia: cuando el traductor desconoce una denominación o una expresión (equivalente) en LM para expresar una idea de la LO que conoce.

- En la LM también hay una necesidad de producción: que consiste en que el traductor conoce una denominación, pero precisa saber cómo utilizar dicha denominación o expresión en el contexto de la LM, cómo expresarse en la LM, necesita acceder a información sobre uso, fraseología, colocaciones, etc.

- Además, una vez se ha encontrado una solución en LM, se produce una necesidad de tipo de recepción en LM porque el traductor necesita acceder al significado de la denominación en LM. La necesidad de recepción en LM se produce porque es preciso verificar o comprobar que el significado en LO y en LM coinciden.

Para abreviar, las necesidades las hemos codificado como *LO* si es una necesidad en lengua origen y como *LM* si es una necesidad en lengua meta. Además se han utilizado letras minúsculas. La letra *a* si es una necesidad de recepción, la *b* si es una necesidad de transferencia y la *c* si es una necesidad de producción. La Ilustración 99 muestra la relación entre todas las necesidades o problemas.



**Ilustración 99. Representación de las necesidades o problemas.**

#### 4.5.2.2.2 Las acciones de búsqueda

Para la resolución de cada una de las necesidades o de los problemas se describen acciones de búsqueda efectuadas por el traductor. En las acciones de búsqueda intervienen una o más consultas, instrumentos y resultados, que pueden llevar al traductor a encontrar una solución satisfactoria a su problema. Estas acciones de búsqueda vienen determinadas por la hipótesis de partida que tiene el traductor.

Las acciones de búsqueda también las presentamos a continuación en el orden más común de aparición en una estrategia de búsqueda, aunque ya hemos comentado que el orden o los patrones de búsqueda no son siempre estables. Normalmente, para cada uno de los problemas descritos, si la primera acción de búsqueda no proporciona un resultado satisfactorio, el traductor pasa a la siguiente acción de búsqueda, y así sucesivamente.

Las acciones de búsqueda están codificadas con las letras que designan la necesidad o el problema que ha provocado la puesta en marcha de la acción, y también con números arábigos. Así, para la acción de búsqueda de tipo 1, la 1.1 suele utilizarse como primera opción de búsqueda, antes que la 1.2 y la 1.3. Por ejemplo, ante una necesidad de recepción en LO (LO.a) donde el traductor conoce una denominación en LO pero desconoce su significado, las acciones de tipo LO.a.1 son la consulta de la denominación conocida en un recurso monolingüe para recuperar el significado desconocido. De estas acciones, la LO.a.1.1 es la consulta de la denominación conocida en un diccionario monolingüe en LO para recuperar el campo de definición, la LO.a.1.2 es la consulta de la denominación conocida en un corpus en LO o en Internet para recuperar contextos definitorios en textos del corpus o páginas de Internet en LO, la LO.a.1.3 sería la consulta de la denominación conocida en un diccionario visual o en un buscador de imágenes en línea para obtener una imagen.

##### 4.5.2.2.2.1 Acciones para la resolución de problemas en LO

Los problemas de recepción en LO (LO.a) consisten en que el traductor conoce una denominación en LO pero desconoce la información conceptual o el significado. En la traducción especializada donde el uso de abreviaturas, siglas o acrónimos es muy frecuente, también es común que el traductor necesite conocer el significado o la forma desarrollada de abreviaturas, siglas o acrónimos.

La información conceptual o de significado de una denominación o de una abreviatura, sigla o acrónimo puede obtenerse en recursos monolingües (LO.a.1). Puede consultarse la denominación de la cual se necesita el significado en un diccionario

(LO.a.1.1), normalmente monolingüe, para obtener una definición, aunque algunos estudiantes consultan también el diccionario bilingüe para conocer el significado a partir de un equivalente. El diccionario puede estar o no disponible en línea.

El traductor también puede consultar la denominación en un corpus o en Internet (la búsqueda en Internet en ocasiones también conduce a una enciclopedia en línea como la *Wikipedia*) para acceder a contextos que contengan información sobre el significado (LO.a.1.2). La ventaja de consultar los contextos de un corpus o Internet es que la información está más actualizada que en los diccionarios, con lo que el traductor se asegura de que accede al significado real de un término (algunos diccionarios pueden contener definiciones desfasadas o el significado de un término o expresión puede haber evolucionado con el paso del tiempo).

Una imagen ayuda al traductor a comprender el significado de un término cuando la definición no es suficiente. Las imágenes pueden obtenerse en diccionarios visuales y, en ausencia de un diccionario visual, la búsqueda de imágenes en línea también puede ser útil (LO.a.1.3).

La consulta de un término en una estructura de conceptos<sup>109</sup> o en el diccionario de sinónimos puede ayudar a deducir el significado al observar cómo unos términos se relacionan con otros (LO.a.1.4).

En el caso de que los recursos anteriores fallen, el traductor puede optar por la consulta a personas, como hablantes nativos o especialistas (LO.a.2).

A continuación mostramos un ejemplo de estrategia de búsqueda para la resolución de un problema de recepción en LO, extraído de nuestro análisis. El ejemplo se obtiene de Sánchez-Gijón (2004a: 204-212), esta autora lo identifica como un problema de unidad lingüística conocida e información conceptual desconocida. Dado el término original *peak* en el contexto de un artículo periodístico sobre el fenómeno de las Leónidas, el traductor desconoce el significado de *peak* (problema que hemos codificado como LO.a). El traductor consulta *peak* en un diccionario monolingüe y accede a una definición (LO.a.1.1), aunque interpreta que esta definición no es suficiente porque es demasiado general. La consulta posterior de *peak* en un corpus compuesto por artículos específicos sobre el fenómeno de las Leónidas ayuda al traductor a averiguar en el análisis de las concordancias y las colocaciones (LO.a.1.2) que es una palabra clave típica de este campo

---

<sup>109</sup> Véase por ejemplo Prieto Velasco y López Rodríguez (2009: 206-208).

y que *peak* es el punto máximo que se calcula o se prevé de un acontecimiento que implica un alto número de meteoros por hora.

En la Tabla 30 esquematizamos el ejemplo de estrategia de búsqueda que acabamos de describir siguiendo nuestro modelo de representación<sup>110</sup>. En la primera línea de la tabla incluimos la necesidad que motiva la estrategia de búsqueda. En la primera columna indicamos los distintos elementos que intervienen en la estrategia de búsqueda (sujeto/usuario, problema, hipótesis, acciones de búsqueda e interpretaciones numeradas por orden de aparición y finalmente solución). En la segunda columna describimos la estrategia de búsqueda en función de cada uno de estos elementos.

<b>Necesidad o problema de recepción en LO</b>	
SUJETO/USUARIO:	supuestamente un traductor, basándose en la propia experiencia de la autora Sánchez-Gijón (2004a: 204-212)
PROBLEMA:	LO.a
CONOCE	<i>peak</i> (LO inglés)
DESCONOCE	información sobre el significado, contexto de un artículo periodístico sobre el fenómeno de las Leónidas
HIPÓTESIS (1)	la búsqueda del término original en un diccionario monolingüe permitirá obtener información sobre el significado.
ACCIÓN/ACCIONES (1):	LO.a.1.1
CONSULTA (1)	<i>peak</i>
RECURSO (1)	diccionario monolingüe <i>Collins Concise Dictionary</i>
RESULTADO (1)	ficha de <i>peak</i> con su definición
INTERPRETACIÓN (1)	se refiere a lo más alto. Definición demasiado general.
HIPÓTESIS (2)	la búsqueda del término original en un corpus monolingüe del tema permitirá obtener información sobre el significado especializado en contexto
ACCIÓN/ACCIONES (2):	LO.a.1.2
CONSULTA (2)	<i>peak</i>
RECURSO (2)	corpus monolingüe inglés compuesto por textos periodísticos sobre las Leónidas

<sup>110</sup> Los ejemplos se han extraído del análisis que incluimos en el anexo 2, donde pueden encontrarse más ejemplos de cada una de las estrategias de búsqueda.

RESULTADO (2)	concordancias y colocaciones de <i>peak</i>
INTERPRETACIÓN (2)	<i>peak</i> tiene una frecuencia elevada en el corpus, es una palabra clave del campo de especialidad. Por las colocaciones más frecuentes en una primera posición a la izquierda, las combinaciones <i>shower peak</i> , <i>storm peak</i> y <i>leonid peak</i> son frecuentes. También coaparecen expresiones de tiempo <i>year</i> , <i>nov (november)</i> , <i>hour</i> , <i>years</i> , <i>night</i> , <i>time</i> , <i>hours</i> , <i>when</i> y los verbos <i>expect</i> y <i>predict</i> , lo que sugiere que es un fenómeno que se calcula previamente. La agrupación <i>peak of</i> sugiere que formará parte de un concepto más amplio, en los contextos se encuentra <i>peak of a storm</i> , <i>peak of meteor activity</i> , <i>peak of the leonid activity</i> . A partir de la expresión de tiempo <i>during the peak</i> se observa que esta unidad se calcula por la cantidad de meteoros por hora y aparecen 18 resultados de <i>peak rate</i> .
SOLUCIÓN	<i>peak</i> es el punto máximo que se calcula o se prevé de un acontecimiento que implica un alto número de meteoros por hora.

**Tabla 30. Ejemplo de estrategia de búsqueda para resolver un problema de recepción en LO.**

#### 4.5.2.2.2 Acciones para la resolución de problemas en LM

Las necesidades en LM son las más comunes en traducción, son necesidades de expresión en las que el traductor necesita acceder a denominaciones desconocidas en LM. Hemos identificado tres tipos de necesidades en LM. La necesidad en LM de transferencia (LM.b) consiste en que el traductor conoce un determinado significado de la LO pero precisa acceder a una denominación o equivalente que exprese dicho significado en la LM. La necesidad en LM de producción (LM.c) se produce cuando el traductor necesita acceder a información que desconoce sobre uso, fraseología o colocaciones en LM. La necesidad en LM de recepción (LM.a) aparece porque el traductor precisa acceder al significado de la denominación o equivalente encontrado en LM para verificar o comprobar que el significado en LO y en LM coinciden. Por eso, especificaremos que las necesidades LM.a son necesidades de tipo LM.a.verificación.

En el caso de la necesidad de transferencia en LM (LM.b) de acceder a denominaciones o equivalentes desconocidos en la LM que expresan un significado de la LO, las primeras acciones de búsqueda que un traductor pone en marcha suelen ser, sobre todo en el caso de la búsqueda de términos simples, la búsqueda del término original en un recurso bilingüe (LM.b.1).

Dentro de esta primera opción de búsqueda se observan distintas subopciones: búsqueda del término original completo, simple o compuesto<sup>111</sup>, en el diccionario (bilingüe, o bilingüe visual o de acrónimos) (LM.b.1.1), búsqueda de un componente del término original compuesto en el diccionario bilingüe (LM.b.1.2), búsqueda del término original completo, simple o compuesto, en un corpus paralelo (LM.b.1.3), búsqueda del término original en Internet en un buscador, con filtro de páginas en el idioma de la LM (o no), o con alguna palabra clave como *diccionario* en LM (o no) (LM.b.1.4). Respecto de las abreviaturas, siglas o acrónimos, la consulta suele comenzar en diccionarios de acrónimos pero suele terminar en una consulta en contexto.

También es posible utilizar una búsqueda a partir de una lengua intermedia como puente entre la LO y la LM (por ejemplo, el latín) (LM.b.2). Esta estrategia es muy útil en la búsqueda de equivalentes de nombres de plantas o animales, que tienen una denominación latina internacional. Por ejemplo, para el acceso a la denominación en español del término en inglés *coalfish*, se consulta en un buscador de Internet *coalfish latin* con lo que se accede a páginas en las que aparece el equivalente latino *pollachius virens*. A continuación la consulta de +“*pollachius virens*” con la restricción de idiomas español permite encontrar páginas en las que se accede al equivalente en español *carbonero* que, si se verifica, es la solución adecuada (Bergeron y Van Steenberg 2000).

En algunas ocasiones la búsqueda del término original en un recurso bilingüe es suficiente para que el traductor aporte una solución de equivalente en LM, aunque en la mayoría de los casos se necesitan nuevas acciones de búsqueda. En este sentido a la búsqueda del término original en un recurso bilingüe suele sucederle estrategias de búsqueda de verificación (aceptación o rechazo del equivalente encontrado en lengua meta). Normalmente el recurso en el que se verifica el equivalente es un corpus comparable, un corpus monolingüe en LM y también Internet. A continuación aportamos ejemplos del análisis.

En el ejemplo presentado por Pearson (1996c), si un estudiante de traducción conoce el término original simple en inglés *neuron* y desconoce un equivalente en LM francés, puede encontrarlo buscando el término original en un diccionario bilingüe y accediendo a

---

<sup>111</sup> La búsqueda de términos compuestos completos en el diccionario bilingüe no siempre da resultados (véase el ejemplo de *carte mère* de Bowker 1998a). Algunos diccionarios no facilitan la consulta de términos compuestos, lo que provoca la búsqueda de un componente del término compuesto en el diccionario bilingüe o la búsqueda del término compuesto completo en un corpus paralelo o en Internet.

los campos de equivalencias, donde encuentra *neurone* (LM.b.1.1). El siguiente problema consiste en que hay que verificar en un contexto monolingüe el equivalente encontrado, por lo que consulta *neuron* y *neurone* en un corpus comparable (LM.a.verificación). Al interpretar las concordancias acepta el equivalente *neurone*. Hemos esquematizado el ejemplo en la Tabla 31.

<b>Necesidad de acceder a denominaciones o equivalentes desconocidos en la LM</b>	
SUJETO/USUARIO:	supuestamente un estudiante de traducción, según Pearson (1996c)
PROBLEMA:	LM.b
CONOCE	<i>neuron</i> (LO inglés)
DESCONOCE	denominación equivalente (LM francés)
HIPÓTESIS (1)	puede accederse a un equivalente a partir del término original en un diccionario bilingüe
ACCIÓN/ACCIONES (1):	LM.b.1.1
CONSULTA (1)	<i>neuron</i>
RECURSO (1)	diccionario bilingüe
RESULTADO (1)	entrada de <i>neuron</i> con información de equivalencias
INTERPRETACIÓN (1)	<i>neuron</i> equivale a <i>neurone</i>
HIPÓTESIS (2)	hay que verificar en recurso monolingüe el equivalente que se ha encontrado en un recurso bilingüe
ACCIÓN/ACCIONES (2a):	LM.a.verificación
CONSULTA (2a)	<i>neuron</i>
RECURSO (2a)	el texto en inglés «How Neural Networks Learn from Experience», publicado en <i>Scientific America</i> en septiembre de 1992
RESULTADO (2a)	concordancias en las que aparece <i>neuron</i>
INTERPRETACIÓN (2a)	no se tienen en cuenta los resultados de <i>artificial neuron</i> porque es un término independiente
ACCIÓN/ACCIONES (2b):	LM.a.verificación
CONSULTA (2b)	<i>neurone</i>
RECURSO (2b)	texto francés «Des neurones électroniques pour rendre les ordinateurs intelligents», publicado en <i>Science et vie</i> en abril de 1992
RESULTADO (2b)	concordancias en las que aparece <i>neurone</i>

INTERPRETACIÓN (2b)	no se tienen en cuenta los resultados de <i>neurone électronique</i> porque es un término independiente. Por las definiciones y explicaciones de ambos términos en los textos se determina que son equivalentes.
SOLUCIÓN	<i>neurone</i>

**Tabla 31. Ejemplo de estrategia de búsqueda de un término original simple en un diccionario bilingüe + verificación del equivalente encontrado en un corpus comparable.**

Creemos relevante destacar que el análisis de los estudios sugiere que el diccionario bilingüe no suele proporcionar casi nunca una solución de equivalencia definitiva, sino que más bien aporta al traductor pistas o ideas que debe verificar o que le sirven para efectuar nuevas búsquedas, normalmente en recursos monolingües en LM en contexto. Asimismo, en ocasiones el diccionario bilingüe no produce resultado alguno.

Cuando la búsqueda del término original en un recurso bilingüe no aporta una solución, la siguiente estrategia más utilizada es la búsqueda a partir de una hipótesis basada en la intuición (LM.b.3). Esta estrategia consiste en que el traductor cree saber cuál podría ser el equivalente que necesita y lo busca en un recurso normalmente monolingüe en LM para confirmar si es correcto.

Dentro de esta estrategia de búsqueda se observan distintas subopciones: búsqueda del término equivalente completo intuido, simple o compuesto (LM.b.3.1), búsqueda de un componente del término equivalente compuesto (LM.b.3.2); en ocasiones frente a un término compuesto, el traductor intuye la traducción de uno de sus componentes, pero no conoce el equivalente del término compuesto completo.

La búsqueda a partir de una hipótesis basada en la intuición puede ser suficiente para que el traductor aporte una solución, aunque también es posible que la hipótesis inicial se descarte y sea necesario generar una nueva hipótesis; o también puede que la hipótesis se confirme pero que el traductor necesite una verificación del equivalente. Por tanto, a esta segunda opción de búsqueda a partir de una o más hipótesis por intuición puede sucederle una estrategia de búsqueda de verificación.

Hemos tomado el siguiente ejemplo de Bowker (2000b: 39) en el que un estudiante de traducción conoce el término original compuesto en francés *système d'exploitation* y desconoce un equivalente en LM inglés. Intuye que uno de los componentes del término equivalente compuesto será *system*. Por tanto, busca *system* en un corpus monolingüe especializado en LM y accede a sus colocaciones (LM.b.3.2). *Operating system* es un término muy frecuente en el corpus y seguramente sea el equivalente que el estudiante



necesita. Para verificarlo (LM.a.verificación), busca *operating system* en las concordancias del mismo corpus, con lo que concluye que es la solución correcta. En la Tabla 32 hemos esquematizado el ejemplo siguiendo nuestro modelo de representación de estrategia de búsqueda.

<b>Necesidad de acceder a denominaciones o equivalentes desconocidos en la LM</b>	
SUJETO/USUARIO:	estudiante de traducción, según Bowker (2000b: 39)
PROBLEMA:	LM.b
CONOCE	<i>système d'exploitation</i> (LO francés)
DESCONOCE	equivalente (LM inglés)
HIPÓTESIS (1)	intuición de que uno de los componentes del término compuesto equivalente en inglés es <i>system</i>
ACCIÓN/ACCIONES (1):	LM.b.3.2
CONSULTA (1)	<i>system</i>
RECURSO (1)	corpus <i>ad hoc</i> monolingüe en inglés de textos sobre sistemas operativos (informática), consultado con <i>WorSmith Tools</i>
RESULTADO (1)	lista de colocaciones de <i>system</i>
INTERPRETACIÓN (1)	<i>operating</i> aparece 241 veces en la primera posición a la izquierda de <i>system</i> . El equivalente seguramente sea <i>operating system</i>
HIPÓTESIS (2)	hay que verificar <i>operating system</i>
ACCIÓN/ACCIONES (2):	LM.a.verificación
CONSULTA (2)	<i>operating system</i>
RECURSO (2)	corpus <i>ad hoc</i> monolingüe en inglés de textos sobre sistemas operativos (informática), consultado con <i>WorSmith Tools</i>
RESULTADO (2)	concordancias de <i>operating system</i>
INTERPRETACIÓN (2)	por el contexto <i>operating system</i> es el equivalente adecuado
SOLUCIÓN	<i>operating system</i>

**Tabla 32. Ejemplo de estrategia de búsqueda a partir de una hipótesis basada en la intuición + verificación (aceptación o rechazo del equivalente en LM).**

Como hemos visto, la hipótesis a partir de una intuición de equivalente se busca normalmente en un recurso monolingüe, por ejemplo en un corpus o Internet para acceder a los contextos, colocaciones o agrupaciones de palabras de dicho término. En la búsqueda de términos compuestos también es común la consulta de combinaciones discontinuas de

formas, cuando se conocen más de una de las formas que componen un término compuesto pero se desconoce en qué orden y con qué otros términos se combinan para formar el término compuesto completo.

Cuando la búsqueda a partir de una hipótesis basada en la intuición no aporta una solución, la siguiente estrategia que suelen recoger los estudios con traductores es la búsqueda a partir de una hipótesis basada en información relacionada del contexto (LM.b.4). En ocasiones cuando el traductor no puede acceder al equivalente en un recurso bilingüe y tampoco tiene una propuesta o intuición de equivalente que confirmar en un recurso monolingüe, una opción es aproximarse al equivalente a partir de determinada información relacionada (que puede obtener del contexto). La búsqueda en LM de elementos que aparecerán en el contexto del equivalente que se necesita resultará útil al traductor para localizar dicho equivalente, que se encontrará en la proximidad de estos elementos.

Los elementos del contexto que pueden buscarse en un corpus en LM para aproximarse a un equivalente son posibles colocaciones porque guardan una relación sintagmática. Puede que el traductor desconozca el equivalente de un sustantivo, pero que sí que conozca con qué verbos se combinará este equivalente en el contexto de la lengua meta (LM.b.4.1); o también pueden guardar una relación semántica con el equivalente que se necesita (LM.b.4.2). La utilización en las consultas de expresiones con contenido semántico permite acceder a contextos con contenido de relaciones semánticas, por ejemplo de sinonimia y antonimia, que facilita la recuperación de términos desconocidos a partir de otros términos conocidos que están relacionados.

En los estudios se han mencionado algunos ejemplos de elementos que pueden buscarse en un recurso para acceder a un contexto donde localizar el equivalente desconocido. Por ejemplo hay elementos que no varían de una lengua a otra, porque son de uso internacional, como determinados acrónimos, abreviaturas o siglas<sup>112</sup>, números o medidas<sup>113</sup> y nombres propios<sup>114</sup>. La consulta de estos elementos invariables facilitará al traductor encontrar equivalentes que aparecerán próximos a dichos elementos en LM.

A veces la búsqueda a partir de una hipótesis basada en información relacionada del contexto es suficiente para que el traductor aporte una solución, aunque también es posible

---

<sup>112</sup> Véase por ejemplo Bowker (2000: 42-43) y Sánchez-Gijón (2004a: 269-270).

<sup>113</sup> Véase por ejemplo Bowker (2000: 45) y Zanettin (2001b).

<sup>114</sup> Véase por ejemplo Bowker (2000: 44) y Zanettin (2001b).

que la hipótesis inicial no lleve a la localización del equivalente y sea necesario generar una nueva hipótesis; o también puede que sí que se localice un equivalente pero que el traductor necesite una verificación de dicho equivalente. Por tanto, a esta opción de búsqueda a partir de una o más hipótesis basadas en información relacionada del contexto pueden sucederle estrategias de búsqueda de verificación.

Por ejemplo, Wilkinson (2005b) observa que un estudiante de traducción desconoce el equivalente en inglés de los adjetivos en finlandés *hoidettu* o *kunnostettu* (LM.b). La primera opción es buscar los adjetivos originales en un diccionario bilingüe (LM.b.1.1), pero los resultados obtenidos no le convencen. Como no es capaz de intuir un posible equivalente, trata de localizarlo a partir de la información que conoce del contexto (LM.b.4). Como estos adjetivos se emplean para caracterizar a los caminos utilizados en el esquí de fondo, *cross-country ski trails*, el equivalente deberá aparecer en un contexto próximo a *trail*, así que busca *trail* en un corpus. En las colocaciones y agrupaciones de palabras de *trail* identifica el adjetivo *groomed*, que se verifica para llegar a la conclusión de que es la solución adecuada (LM.a.verificación). En la Tabla 33 sintetizamos este ejemplo.

<b>Necesidad de acceder a denominaciones o equivalentes desconocidos en la LM</b>	
SUJETO/USUARIO:	supuesto estudiante de traducción, según Wilkinson (2005b)
PROBLEMA:	LM.b
CONOCE	<i>hoidettu</i> o <i>kunnostettu</i> (adjetivos en LO finlandés)
DESCONOCE	equivalente (LM inglés)
HIPÓTESIS (1)	los equivalentes de los términos originales <i>hoidettu</i> o <i>kunnostettu</i> aparecerán en un diccionario bilingüe.
ACCIÓN/ACCIONES (1):	LM.b.1.1
CONSULTA (1)	<i>hoidettu</i> o <i>kunnostettu</i>
RECURSO (1)	diccionario bilingüe
RESULTADO (1)	equivalencias: <i>conditioned</i> , <i>maintained</i> , <i>restored</i> o <i>reconditioned</i>
INTERPRETACIÓN (1)	el traductor no cree que estos adjetivos puedan emplearse para describir a un camino ( <i>trail</i> ) en este contexto.
HIPÓTESIS (2)	los adjetivos de los cuales se necesita un equivalente se emplean para caracterizar a los caminos utilizados en el esquí de fondo, <i>cross-country ski trails</i> . El equivalente deberá aparecer en un contexto próximo a <i>trail</i> .

ACCIÓN/ACCIONES (2):	LM.b.4
CONSULTA (2)	<i>trail</i>
RECURSO (2)	corpus <i>ad hoc</i> monolingüe en inglés de textos turísticos, consultado con <i>WorSmith Tools</i>
RESULTADO (2a)	más de 1000 concordancias de la palabra <i>trail</i> en el corpus
INTERPRETACIÓN (2a)	resulta demasiado complicado analizar 1000 concordancias.
RESULTADO (2b)	lista de colocaciones que aparecen con mayor frecuencia 5 posiciones a la izquierda de <i>trail</i>
INTERPRETACIÓN (2b)	los adjetivos que sugería el diccionario no aparecen como colocaciones de <i>trail</i> . En 40 ocasiones el adjetivo <i>groomed</i> aparece como colocación en una posición inmediatamente a la izquierda de <i>trail</i> . Puede que <i>groomed</i> sea el adjetivo adecuado.
RESULTADO (2c)	agrupaciones más frecuentes de 3 palabras que contienen el término <i>trail</i>
INTERPRETACIÓN (2c)	hay agrupaciones de tres palabras que combinan <i>groomed</i> y <i>trail</i> . Puede que <i>groomed</i> sea el adjetivo adecuado.
HIPÓTESIS (3)	es necesario verificar que <i>groomed</i> es el equivalente adecuado
ACCIÓN/ACCIONES (3):	LM.a.verificación
CONSULTA (3)	<i>groomed</i>
RECURSO (3)	corpus <i>ad hoc</i> monolingüe en inglés de textos turísticos, consultado con <i>WorSmith Tools</i>
RESULTADO (3)	128 concordancias: <i>groomed bicycle and walking trails, groomed classic and skating trails, groomed cross-country ski trails, groomed fairways, groomed off-road trails, groomed runs, groomed slopes</i> y <i>groomed wilderness trails</i>
INTERPRETACIÓN (3)	existe <i>groomed cross-country ski trails</i> , por tanto <i>groomed</i> es un equivalente adecuado.
SOLUCIÓN	<i>groomed</i>

**Tabla 33. Ejemplo de búsqueda a partir de una o más hipótesis basadas en información relacionada del contexto + verificación (aceptación o rechazo del equivalente en LM).**

En ocasiones el traductor puede combinar una estrategia de búsqueda basada en una intuición (LM.b.3) con una estrategia de búsqueda basada en información relacionada del contexto (LM.b.4). Por ejemplo, en la búsqueda de los equivalentes en español de *storm* y *shower* (en el contexto del fenómeno de las Leónidas), el traductor intuye que los

equivalentes serán *tormeta* y *lluvia*, por un lado y, por otro lado, que deberán aparecer en un contexto cercano a *meteor*\* (Sánchez-Gijón 2004a: 225-230).

También es posible que una estrategia de búsqueda comience con una hipótesis basada en una intuición, pero que se descarte y sea necesario recurrir a una hipótesis basada en información relacionada del contexto. Por ejemplo, Aston (1999) describe la búsqueda de un equivalente o una denominación en inglés para la denominación en francés *clochemerlesques* de la cual entiende el significado (LM.b). A partir de una hipótesis basada en la intuición, piensa que el equivalente en inglés tendrá la misma raíz que el término en francés, por lo que consulta *clochemerle*\* en un corpus y en Internet (LM.b.3). Sin embargo, esta hipótesis se descarta. Entonces, se plantea una búsqueda a partir de adjetivos que aparecen en el contexto de *Clochemerle* (*polémique, querelle, discorde, guérilla, rivalité, conflit, rebelle*) y que describen su significado (LM.b.4). La consulta de equivalentes en inglés de estos adjetivos (*quarrel, squabble, row, bicker, dispute, rebel, battle, rivalry*) en las concordancias de un corpus y en Internet llevan al autor a deducir que *squabble* expresa el matiz que se está buscando, por lo que la solución que se adopta en el texto meta es *local squabbling*.

En la bibliografía se sugiere que este tipo de estrategias de búsqueda basadas en información relacionada para acceder a términos desconocidos tienen una naturaleza onomasiológica. Los autores se refieren más bien a búsquedas indirectas para acceder a términos desconocidos (Wilkinson 2007: 108), una búsqueda que también han comentado otras autoras (Bowker 2000: 37-45; Bowker y Pearson 2002: 200; Sánchez-Gijón 2004a: 212-215, 2009: 120; Wilkinson 2005a, 2007: 108).

Otra estrategia de búsqueda, si las demás fallan, sería tratar de acceder al equivalente mediante la lectura de textos paralelos<sup>115</sup> (LM.b.5). También es posible que el equivalente se encuentre por casualidad (LM.b.6). En este sentido, algunos autores indican que en ocasiones la localización del equivalente idóneo se produce de manera fortuita, lo que se ha denominado *serendipity* o *serendipitous finds* (Aston 2000; Wilkinson 2005b; Zanettin 2001b). Aunque, como también indican los autores, más que por el simple azar, los resultados de búsqueda son el fruto de la puesta en marcha de una serie de acciones donde unos términos llevan a otros términos, que a su vez dan nuevas pistas al traductor para efectuar nuevas búsquedas con las que finalmente se encuentra lo que se estaba buscando.

---

<sup>115</sup> Véase por ejemplo Schäffner (1998).

Estas son estrategias que, aunque a veces no se realizan de manera consciente, permiten al traductor llegar a los términos que precisa. También se ha mencionado la estrategia de búsqueda en foros<sup>116</sup> (LM.b.7).

Hemos visto que casi todas las búsquedas de equivalentes en LM suelen concluir con una estrategia de verificación, aceptación o rechazo del equivalente en LM. Por ejemplo, verificación de un posible equivalente, verificación de dos o más posibles equivalentes, verificación de dos o más variantes ortográficas.

Las estrategias de verificación son de recepción en LM (LM.a.verificación). Para verificar la validez de un equivalente en LM, el traductor a veces compara el significado de los términos supuestamente equivalentes en LO y en LM. Este tipo de estrategias pueden ayudar al traductor a darse cuenta de que existe un equivalente más adecuado, por ejemplo si un texto va dirigido a un público no experto el término *mucose* es demasiado especializado y sería recomendable emplear términos más generales, como por ejemplo *lining* o *tissue* (Varantola 2002a).

La verificación puede realizarse en un corpus monolingüe o en Internet, de términos encontrados normalmente en un recurso bilingüe (diccionario o corpus paralelo), o en el mismo corpus monolingüe o Internet. Asimismo la verificación puede obtenerse con la obtención de información visual o imágenes (por ejemplo en Internet).

En este sentido, se ve claramente la utilidad de consultar Internet como un corpus. Como ya se ha mencionado en algunas obras, el corpus es un recurso muy preciado desde el punto de vista lingüístico, pero adolece de la ausencia de contenido multimedia (imágenes, archivos de sonido, videos), elementos fácilmente accesibles en la red y de enorme valor para el traductor (Bowker 2002: 31).

Además de la información sobre los equivalentes en LM, el traductor tiene una necesidad de producción en LM (LM.c), principalmente de información de uso y colocaciones. Hemos visto estrategias de búsqueda para acceder a las colocaciones en un corpus (LM.c.1.1), en Internet (LM.c.1.2) o en un diccionario monolingüe (LM.c.1.3). Los estudios sobre el uso de diccionarios de colocaciones no nos han proporcionado ejemplos sobre cómo los traductores realizan búsquedas en estos diccionarios.

En un corpus, si se accede a las concordancias de una palabra, ordenando las concordancias por las palabras a izquierda y derecha puede agilizarse la localización de

---

<sup>116</sup> Véase Enríquez (2011).

colocaciones y términos compuestos, por ejemplo los adjetivos en inglés suelen situarse en una primera posición a la derecha de los sustantivos. La opción de búsqueda de colocaciones y agrupaciones de palabras en un corpus también agiliza esta búsqueda. Asimismo es interesante la consulta en corpus etiquetados con la que pueden especificarse las categorías gramaticales de las colocaciones que se desean obtener. Se consulta una palabra (de la cual se necesita una colocación) en combinación discontinua con una categoría gramatical. En las concordancias y listas de colocaciones resultantes, las colocaciones recuperadas se corresponderán con la categoría gramatical introducida. Por ejemplo, pueden buscarse todos los adjetivos que aparecen en una posición a la izquierda del sustantivo *printer* para acceder a tipos de impresoras (Bowker y Pearson 2002: 188-189).

Para una mejor comprensión de la clasificación preliminar de estrategias de búsqueda que acabamos de explicar, hemos esquematizado las estrategias de búsqueda en una tabla que presentamos en el anexo 3. La tabla consta de dos columnas. En la columna de la izquierda identificamos la necesidad o el problema y en la columna de la derecha las acciones de búsqueda llevadas a cabo por los traductores para resolver cada problema.

En la Tabla 34 incluimos una muestra del anexo 3, en la que vemos que ante una necesidad o un problema en LO de recepción (LO.a) en el que el traductor conoce una denominación en LO pero desconoce su significado, las acciones de búsqueda pueden ser la consulta de la denominación en LO conocida en un recurso monolingüe para obtener el significado desconocido (LO.a.1). Dentro de esta modalidad de acciones, el traductor puede consultar la denominación en LO conocida en un diccionario monolingüe y obtener una definición (LO.a.1.1). Para acceder a la tabla completa con el resumen de la clasificación preliminar de estrategias de búsqueda, véase el anexo 3.

<p><b>NECESIDAD o PROBLEMA</b></p> <ul style="list-style-type: none"> <li>- LO o LM</li> <li>a. Recepción</li> <li>b. Transferencia</li> <li>c. Producción</li> <li>* verificación</li> </ul>	<p><b>HIPÓTESIS – ACCIONES DE BÚSQUEDA</b></p> <ul style="list-style-type: none"> <li>- consulta</li> <li>- instrumento</li> <li>- resultado</li> </ul>
<b>LO</b>	
<p><b>LO.a Recepción</b></p> <p>CONOCE: denominación en LO o sigla/abreviatura/acrónimo</p> <p>&gt;</p> <p>DESCONOCE: significado</p>	<p><b>LO.a.1</b></p> <ul style="list-style-type: none"> <li>- Consulta de la denominación en LO conocida</li> <li>- en un recurso monolingüe</li> <li>- para obtener el significado desconocido</li> </ul>
	<p><b>LO.a.1.1</b></p> <ul style="list-style-type: none"> <li>- Denominación en LO / o sigla/abreviatura/acrónimo</li> <li>- diccionario monolingüe en LO (disponible o no en línea) / o diccionario de acrónimos</li> <li>- definición / o también forma desarrollada de un acrónimo)</li> </ul>

**Tabla 34. Muestra del resumen de la clasificación preliminar de estrategias de búsqueda del anexo 3.**

#### 4.5.2.3 Valoración sobre las estrategias de búsqueda identificadas en los estudios

De todo lo que hemos visto hasta ahora, el análisis de estrategias de búsqueda identificadas en los estudios nos ha permitido identificar determinados patrones de búsqueda de los traductores. No obstante, hemos comprobado que los patrones no siempre siguen un modelo fijo. La utilización y combinación de unas estrategias u otras depende de los recursos de los que disponga el traductor, de cada situación de búsqueda y de la propia experiencia del traductor tanto en el tema tratado como en su habilidad para efectuar búsquedas. Algunos estudios sugieren que los traductores con menor experiencia efectúan búsquedas más simples, por ejemplo las investigaciones con estudiantes de traducción muestran que, en muchos casos, su recurso principal de búsqueda es el diccionario bilingüe y que tienen dificultades a la hora de realizar búsquedas eficaces en contexto, en corpus o Internet. Por tanto, sería arriesgado tratar de extrapolar directamente los resultados sobre estrategias de búsqueda de estudiantes de traducción a la traducción en general y, especialmente, a la traducción profesional.

Por otro lado, en cuanto a las búsquedas que más nos interesan, las búsquedas onomasiológicas, del significado a la denominación, deducimos que las necesidades y las



búsquedas onomasiológicas en traducción se producen en la fase de expresión, sobre todo en la fase de transferencia en la búsqueda de equivalentes, cuando el traductor comprende un concepto o idea en la LO, pero desconoce una denominación o equivalente en LM para expresar dicha idea; por ejemplo, los autores se han referido a búsquedas indirectas para acceder a términos desconocidos en LM cuando la búsqueda en un diccionario bilingüe a partir del término original no funciona (véanse a modo de ejemplo Bowker 2000: 37-45; Bowker y Pearson 2002: 200; Rogers y Ahmad 1998: 199; Sánchez-Gijón 2004a: 212-215, 2009: 120; Wilkinson 2005a, 2007: 108). También dentro de la fase de producción, cuando el traductor busca colocaciones que expresan un significado determinado (Jousse et al. 2011; L'Homme y Leroyer 2009: 278).

En este sentido, hemos visto que una estrategia para acceder a la denominación desconocida en LM consiste en consultar a partir del término original en un recurso bilingüe, por ejemplo un diccionario bilingüe. Así, la búsqueda en recursos bilingües, desde nuestro punto de vista, podría considerarse una forma de evitar la búsqueda onomasiológica en traducción, ya que en lugar de partir del significado original para llegar a la denominación equivalente en LM, lo que se busca es una denominación en LO para llegar a otra denominación en LM, que supuestamente es equivalente. El mayor problema de este tipo de búsqueda bilingüe, como alegan los autores, es que los recursos bilingües no proporcionan siempre soluciones a los problemas de expresión.

Por tanto, la búsqueda onomasiológica será necesaria cuando el traductor no pueda utilizar el término original como punto de acceso al término en LM, y cuando no tenga una intuición previa a partir de la cual el traductor podría buscar el equivalente para comprobarlo; a veces el término original no aparece en el diccionario bilingüe y el equivalente intuido no siempre se confirma.

En este caso, el traductor deberá desarrollar estrategias, que podrían considerarse de naturaleza onomasiológica, y que en el análisis hemos visto que pueden basarse en información relacionada con la denominación desconocida. La mayor parte de estas estrategias tienen lugar en recursos monolingües, principalmente corpus e Internet, donde el traductor busca a partir de información que obtiene del contexto: una relación de tipo sintagmático (una colocación) y también de tipo semántico (por ejemplo un sinónimo o un hiperónimo). La búsqueda de denominaciones que están relacionadas con el equivalente que el traductor desconoce puede ayudar al traductor a localizar el equivalente en un contexto próximo a dichas denominaciones. Por ejemplo, hemos visto que los traductores pueden acceder a equivalentes desconocidos buscando a partir de colocaciones típicas

conocidas, o a partir de siglas, nombres propios y números invariables. Asimismo, las categorías gramaticales en un corpus etiquetado permiten acceder a las formas que suelen aparecer en una determinada estructura sintáctica, por ejemplo los verbos que se combinan con determinado sustantivo. La consulta de imágenes es útil en la comprobación de conceptos.

Las búsquedas en recursos bilingües del término original para acceder al equivalente y las búsquedas basadas en una intuición de equivalente que se comprueba pueden considerarse de naturaleza más bien semasiológica. Cabe señalar que estas estrategias semasiológicas son más frecuentes que las estrategias de naturaleza onomasiológica, seguramente porque hasta ahora en los recursos disponibles estas estrategias resultan más sencillas, rápidas y directas.

Las estrategias de búsqueda de naturaleza onomasiológica, de lo que hemos podido observar hasta el momento, son de las más complejas que puede efectuar un traductor, requieren un esfuerzo de reflexión y análisis y la puesta en marcha de diferentes acciones de búsqueda.

Por tanto, desde nuestro punto de vista, las búsquedas onomasiológicas deberán ir por el camino de agilizar la tarea al traductor en la realización de búsquedas a partir de denominaciones relacionadas, para identificar el equivalente deseado con el menor tiempo y esfuerzo posible.

Teniendo en cuenta lo que hemos visto hasta ahora, en el siguiente capítulo aportaremos nuestro propio estudio empírico sobre las estrategias de búsqueda que utilizan los traductores, centrándonos en aquellas estrategias que podríamos considerar de tipo onomasiológico. Nos fijaremos en los recursos más útiles para realizar búsquedas onomasiológicas, las combinaciones de búsquedas más utilizadas para detectar patrones o modelos de búsqueda onomasiológica. Con ello, tratamos de ordenar y ofrecer un panorama lo más amplio posible de las estrategias de búsquedas onomasiológicas en traducción. Las estrategias de búsqueda identificadas en este capítulo nos servirán como punto de partida en el diseño de nuestra investigación y en el posterior análisis de los resultados.

## 5 Estudio empírico sobre estrategias de búsqueda onomasiológica en traducción

[...] in order to look up a term in a TL dictionary, the translator has to know what term to look for, and if the translator already knows the term, chances are there is no need to look it up. The main challenge then is to devise a method for looking up “unknown” terms in resources where the SL term cannot be used as an access point.

(Bowker 2000: 22)

A lo largo de este trabajo hemos visto que la evolución de las tecnologías de la información ha propiciado una transformación de los recursos, de diccionarios en papel a diccionarios electrónicos y diccionarios en línea. Esta transición de formatos aparentemente debería implicar una revolución de los modos de búsqueda tradicionales, aprovechando las ventajas que aporta el formato electrónico.

Sin embargo, se ha observado que todavía en la actualidad los diccionarios no siempre explotan todas las ventajas que ofrece el formato electrónico. En este sentido, tampoco parece que se esté evolucionando tan rápido como se desearía en el desarrollo de diccionarios que incorporen nuevas técnicas de consulta, como por ejemplo las consultas onomasiológicas, del concepto a la denominación (Alcina 2009a: 37).

En los capítulos anteriores, por un lado, hemos efectuado un análisis de los recursos (diccionarios, corpus e Internet) y de las posibilidades de búsqueda que pueden ofrecer, *técnicas de búsqueda*, que hemos sintetizado en el apartado 3.5. Por otro lado, hemos observado que actualmente existe una fuerte corriente investigadora que trata de averiguar qué necesidades o qué formas de búsqueda pueden interesar a un usuario, es decir, un acercamiento basado en los usuarios más que en los recursos (*user-oriented approach*). Uno de los usuarios potenciales de un diccionario es el traductor. Por tanto, el análisis de otros estudios sobre las *estrategias de búsqueda* que utilizan los traductores en diccionarios, corpus e Internet nos ha permitido recopilar y ordenar una serie de estrategias de búsqueda que se utilizan en traducción, y las hemos presentado en nuestro modelo preliminar de estrategias de búsqueda en traducción propuesto en el apartado 4.5.

No obstante, en estos análisis previos hemos detectado que las estrategias de búsqueda de tipo onomasiológico apenas han recibido atención, salvo menciones en algunos trabajos. Por ejemplo, se sugiere que los traductores tienen necesidades de naturaleza onomasiológica, de acceso a las denominaciones que desconocen (Bowker 2000: 22; Rogers y Ahmad 1998: 199; Sánchez-Gijón 2009: 120; Verlinde et al. 2010: 3;

Wilkinson 2005a, 2007: 108). En este sentido, hemos encontrado estudios que tratan de recopilar estrategias de búsqueda de estudiantes de traducción para el acceso a denominaciones desconocidas en un corpus (Bowker 2000), que intentan comprobar la utilidad para los usuarios de las funciones de búsqueda onomasiológica en un diccionario de colocaciones (Jousse et al. 2011), o también algunos proponen la elaboración de diccionarios adaptados a los usuarios que permitan el acceso onomasiológico a las colocaciones (L'Homme y Leroyer 2009; Leroyer 2006).

Asimismo, la mayoría de los estudios analizados utilizan como sujetos a estudiantes de traducción, no a traductores profesionales, lo que no nos permite extrapolar directamente los resultados reunidos en estas investigaciones a la traducción profesional. De hecho, algunos de los autores que sí han investigado las formas de búsqueda de traductores profesionales han constatado diferencias en las estrategias de búsqueda de estos y de los estudiantes de traducción (Fraser 1999: 25), como por ejemplo que los traductores profesionales utilizan mayor variedad de recursos y de estrategias de búsqueda (Künzli 2001: 520; Massey y Ehrensberger-Dow 2011: 197-199). De hecho, algunos autores sugieren la necesidad de más estudios empíricos de las estrategias de búsqueda de los traductores profesionales (Massey y Ehrensberger-Dow 2011: 194-195; Ronowicz et al. 2005: 592).

Todas estas premisas justifican que nos hayamos planteado un estudio empírico con el que pretendemos realizar una aportación a la investigación sobre estrategias de búsqueda de los traductores profesionales, centrándonos además en un tipo concreto de estrategia, las *estrategias de búsqueda onomasiológica*.

En dicho estudio partimos de la base de que los traductores a veces se encuentran en la situación de que comprenden una idea o concepto de la LO, pero desconocen su denominación o equivalente en la LM. En otras ocasiones, conocen una denominación o un equivalente pero no lo recuerdan en ese momento, fenómeno que se conoce como «tener algo en la punta de la lengua». En este caso, no pueden buscar alfabéticamente la denominación en un diccionario y habrán tenido que recurrir a una serie de estrategias de búsqueda para encontrar la denominación. En nuestro estudio queremos contar con la experiencia de los traductores, que posiblemente han necesitado realizar este tipo de búsquedas, que hemos llamado *estrategias de búsqueda onomasiológica* (porque van del significado a la denominación).

Nuestra hipótesis es que la dificultad en el acceso a las denominaciones a partir de los conceptos es un problema de traducción real y que no se han diseñado todavía herramientas que permitan solucionar este problema de forma directa o que, aunque en algunos recursos es posible hacer este tipo de consultas, los creadores no lo han previsto y no lo explicitan a los usuarios. Creemos que los traductores han tenido que desarrollar una serie de estrategias de búsqueda (onomasiológicas) para acceder a las denominaciones con las fuentes terminológicas de que disponen. La ordenación alfabética de los diccionarios tradicionales no permite realizar este tipo de búsquedas y, por eso, los traductores han tenido que recurrir a otro tipo de herramientas como Internet y los buscadores o los corpus electrónicos para satisfacer sus necesidades parcialmente. Por tanto, consideramos que sería útil que los recursos implementaran técnicas de búsqueda o funciones que facilitaran la búsqueda onomasiológica a los traductores.

Para ello, entendemos que primero conviene conocer y tener en cuenta las estrategias de búsqueda onomasiológica que emplean los traductores, para después adaptar las funciones de los recursos con el fin de facilitar a los traductores el uso de aquellas estrategias de búsqueda en las que el traductor accede a una denominación que desconoce a partir de un concepto o de otra información que sí que conoce.

Los objetivos generales de este estudio son responder a las siguientes preguntas:

- primero, confirmar que los traductores tienen necesidades de tipo onomasiológico;
- segundo, conocer qué recursos pueden resultar útiles a los traductores en la búsqueda onomasiológica (diccionarios, corpus, Internet, etc.); y
- tercero, detectar qué estrategias les resultan más útiles con cada recurso para resolver este tipo de situaciones, centrándonos en indagar a partir de qué información buscan los traductores (una definición, el área temática, un sinónimo, etc.) para encontrar una denominación que desconocen o no recuerdan.

Con ello, pretendemos detectar aquellas estrategias de búsqueda onomasiológica que utilicen los traductores actualmente o las que crean que podrían desarrollarse en un diccionario o en un recurso onomasiológico y que facilitarían el trabajo del traductor en la búsqueda de palabras desconocidas u olvidadas.

Para la consecución de estos objetivos, nuestra investigación consta de dos fases. Primero, efectuamos un estudio piloto consistente en una encuesta (apartado 5.1). Con ella, evaluamos las posibles preguntas que nos permitirán obtener información por parte de los traductores respecto de sus búsquedas, y además realizamos una primera aproximación a las búsquedas onomasiológicas en traducción para comprobar la pertinencia de nuestro

estudio. Segundo, planteamos nuestro estudio principal en el que efectuamos una entrevista y profundizamos en las estrategias de búsqueda onomasiológica en traducción (apartado 5.2). A continuación, definiremos objetivos específicos dentro de cada uno de estos dos estudios.

### **5.1 Estudio piloto: encuesta a traductores sobre técnicas y estrategias de búsqueda onomasiológica**

El objetivo de nuestro estudio piloto es evaluar la metodología y las posibles preguntas para obtener información sobre las búsquedas onomasiológicas de los traductores y obtener unos primeros resultados sobre este tipo de búsquedas. Para ello, elaboramos un cuestionario en el que intentamos recoger aquellas preguntas que pensamos que nos ayudarían a recabar la información necesaria. Finalmente, el cuestionario lo enviamos a un grupo de traductores profesionales. Los resultados obtenidos nos permitieron extraer unas primeras conclusiones sobre el uso de los recursos por parte de los traductores y las estrategias de búsqueda onomasiológica empleadas<sup>117</sup>.

A continuación presentamos este estudio con más detalle. En el apartado 5.1.1 resumimos los objetivos, en el 5.1.2 hacemos referencia a la metodología empleada, a continuación en el 5.1.3 incluimos los resultados y, finalmente, en el 5.1.4 realizamos una valoración de los mismos.

#### **5.1.1 Objetivos del estudio piloto**

Los objetivos del estudio piloto son los siguientes:

- Realizar una primera aproximación a los recursos que el traductor considera más útiles para la resolución de problemas que requieren una búsqueda onomasiológica.
- Realizar una primera aproximación a las diferentes estrategias de búsqueda onomasiológica que se emplean para acceder a una denominación desde un concepto.
- Evaluar posibles preguntas y el instrumento para la recopilación de información sobre estrategias de búsqueda en traducción.

---

<sup>117</sup> El estudio piloto completo se presentó en el trabajo de fin de máster (Pastor 2008b), en el Máster en Tecnologías de la Traducción y Localización, de la Universitat Jaume I. También se puede acceder a un resumen sobre los resultados de esta investigación previa (Pastor 2008a).

## 5.1.2 Metodología

Elaboramos una encuesta que debía responderse por escrito y la enviamos a cuatro agencias de traducción y a la lista de distribución *Infotrad*. Recibimos un total de 15 respuestas de traductores profesionales que analizamos para obtener los primeros resultados sobre las búsquedas onomasiológicas en traducción.

A continuación exponemos el método de investigación utilizado. En primer lugar, en el apartado 5.1.2.1, explicamos el diseño y la estructura del cuestionario que utilizamos como instrumento de recogida de datos en este estudio piloto. Después, en el apartado 5.1.2.2, describimos el modo de acceso a los sujetos y el perfil de los traductores que participaron en nuestro estudio preliminar. Finalmente, indicamos la manera de administrar el cuestionario y analizar la información (apartado 5.1.2.3).

### 5.1.2.1 Diseño del cuestionario

Para la realización del cuestionario, planteamos a los encuestados una situación de traducción en la que necesitaban encontrar, a partir de un concepto conocido, una denominación que desconocían en la LM. A partir de este supuesto de partida, orientamos las preguntas en dos sentidos. La utilización, por un lado, de distintos recursos para solventar este tipo de dificultades terminológicas y, por otro, el uso de las estrategias de búsqueda que se emplean con cada recurso. Para la elaboración de las preguntas tuvimos en cuenta un análisis previo de recursos (Pastor 2008b: 39-110), que hemos completado en el capítulo 3 de este trabajo.

El cuestionario constaba de 10 preguntas con las que pretendíamos obtener la siguiente información.

- Conocer el perfil de los encuestados.
- Qué recursos utilizan preferentemente los traductores para realizar una búsqueda onomasiológica. Con cuáles se obtienen mejores resultados o cuáles prefieren los traductores. (Preguntas 1-3).
- Qué estrategias de búsqueda onomasiológica emplean los traductores con los diccionarios disponibles y qué carencias detectan o qué modos de consulta les resultarían útiles con esta herramienta. (Preguntas 4 y 5).
- Qué estrategias se emplean en los corpus para realizar búsquedas onomasiológicas. (Pregunta 6).
- De qué manera se utiliza el foro para realizar una búsqueda onomasiológica. (Pregunta 7).

- Qué estrategias se emplean en Internet, y especialmente en los buscadores, para realizar búsquedas onomasiológicas. (Preguntas 8 y 9).
- Qué otros recursos o estrategias se emplean para realizar búsquedas onomasiológicas. (Pregunta 10).
- Evaluación de la encuesta por parte de los encuestados. Posibles mejoras o sugerencias.

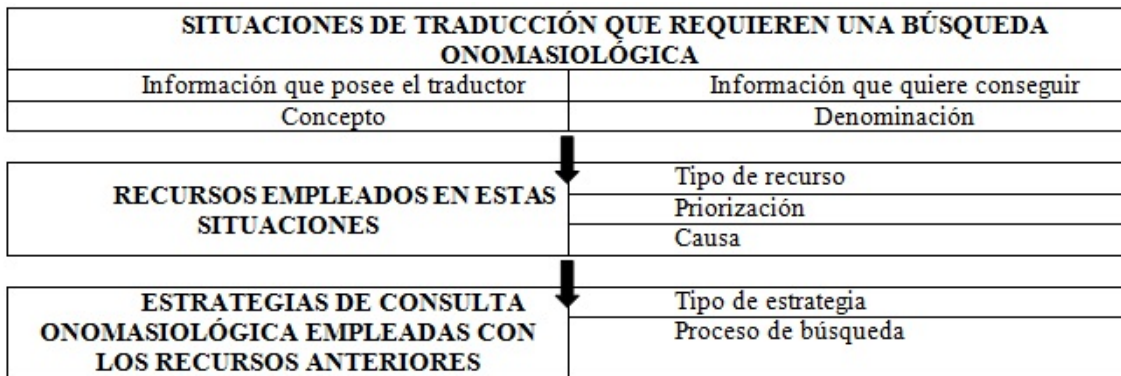


Tabla 35. Esquema del cuestionario.

### 5.1.2.2 Sujetos

Para el acceso a los sujetos contactamos con traductores profesionales de las empresas colaboradoras del proyecto ONTODIC y del Máster en Tecnologías de la Traducción y Localización, así como suscriptores de la lista *Infotrad*.

Recurrimos a traductores profesionales ya que estos tienen mayor experiencia en traducción y, por tanto, también habrán desarrollado un mayor número de estrategias para la búsqueda de terminología, entre las que pueden incluirse estrategias de búsqueda onomasiológica.

En el estudio piloto participaron 15 traductores profesionales: el 14% con más de 10 años de experiencia, el 53% con 2 a 10 años de experiencia y el 33% con menos de 2 años de experiencia.

Partimos de una muestra de 15 sujetos para extraer unas primeras conclusiones y comprobar la utilidad de nuestro modelo de encuesta y de las preguntas. En investigaciones posteriores, se modificaría la encuesta en función de las necesidades observadas y se ampliaría el grupo de sujetos.



### 5.1.2.3 Recogida y análisis de la información

Las encuestas se enviaron por correo electrónico a cuatro agencias de traducción españolas que colaboran en el proyecto ONTODIC y en el Máster en Tecnologías de la Traducción y Localización, y también se mandó un mensaje a la lista *Infotrad*. En dichos envíos presentábamos al grupo TecnoLeTTra, esta investigación y la propia encuesta e incluimos un archivo de Word adjunto con la encuesta.

Las respuestas de las encuestas se recibieron por correo electrónico en el plazo de tres semanas desde que se enviaron. El total de encuestas recibidas y analizadas fueron 15.

Para el análisis de las respuestas, clasificamos las contestaciones del cuestionario en función de las preguntas, de manera que para cada pregunta teníamos recogidas una detrás de la otra todas las respuestas a dicha pregunta, lo que nos facilitó el análisis de los resultados.

### 5.1.3 Resultados

Los resultados se obtuvieron del análisis de las 15 encuestas que recibimos y analizamos.

En la primera pregunta, se pedía a los encuestados que ordenaran cinco recursos: el diccionario bilingüe o multilingüe, el diccionario monolingüe, el corpus, el foro o las listas de distribución e Internet, en función de su utilidad para realizar una búsqueda onomasiológica (véase la Tabla 36).

En la primera posición (recurso de mayor utilidad), el 60% situó a Internet, el 20% al diccionario bilingüe o multilingüe, el 14% al corpus, el 6% al foro y nadie mencionó al diccionario monolingüe.

En la segunda posición, el 34% eligió el foro, el 20% Internet, el 20% el corpus, el 13% el diccionario monolingüe, y el 13% el diccionario bilingüe o multilingüe.

En la tercera posición, el 28% escogió al corpus, el 26% al diccionario monolingüe, el 20% al diccionario bilingüe o multilingüe, el 13% al foro y el 13% a Internet.

En la cuarta posición, el 27% seleccionó el foro, el 27% el diccionario monolingüe, el 20% el diccionario bilingüe o multilingüe, el 20% el corpus y el 6% Internet.

En la última posición (recurso de menor utilidad), el 34% mencionó el diccionario monolingüe, el 26% el diccionario bilingüe o multilingüe, el 20% el foro, el 20% el corpus y nadie Internet.

1ª posición	Internet (60%)	Diccionario bilingüe o multilingüe (20%)	Corpus (14%)	Foro (6%)	Diccionario monolingüe (0%)
2ª posición	Foro (34%)	Internet (20%)	Corpus (20%)	Diccionario monolingüe (13%)	Diccionario bilingüe o multilingüe (13%)
3ª posición	Corpus (28%)	Diccionario monolingüe (26%)	Diccionario bilingüe o multilingüe (20%)	Foro (13%)	Internet (13%)
4ª posición	Foro (27%)	Diccionario monolingüe (27%)	Diccionario bilingüe o multilingüe (20%)	Corpus (20%)	Internet (6%)
5ª posición	Diccionario monolingüe (34%)	Diccionario bilingüe o multilingüe (26%)	Foro (20%)	Corpus (20%)	Internet (0%)

**Tabla 36. Prioridad dada por los traductores a los recursos según su utilidad para realizar una búsqueda onomasiológica**

El recurso más valorado por casi todos los encuestados fue Internet, que se situó sobre todo en primera posición y nunca en última posición. Los siguientes recursos más útiles fueron el corpus y el foro. Salvo excepciones, los diccionarios, bilingüe y monolingüe, aparecieron en las últimas posiciones.

En la segunda pregunta, sobre otros recursos útiles para la búsqueda onomasiológica, dos encuestados se refirieron a las memorias de traducción (en este caso para la gestión de textos paralelos), otros siete participantes coincidieron en la consulta directa a especialistas en la materia a partir de la información del concepto, principalmente el contexto y el área temática. Seis traductores sugirieron los diccionarios de sinónimos y antónimos, los tesauros, los diccionarios ideológicos o de ideas afines, diccionarios visuales, buscadores temáticos y enciclopedias digitales. Un encuestado mencionó el *Diccionario de ideas afines* de Corripio, el diccionario visual *Pictorial* y el motor de búsqueda restringido *Google Scholar*.

En la tercera pregunta, se indagaron las razones del orden que se le había dado a los recursos de la pregunta 1. Para determinar la utilidad de los recursos los encuestados mencionaron razones de accesibilidad, rapidez y autoridad, principalmente. Los que preferían Internet indicaron que es el recurso más accesible y el más rápido en la obtención de resultados. Se destacó la posibilidad de acceder por palabras clave a textos en los que puede aparecer la palabra que se busca.

Los que situaron al corpus en las primeras posiciones lo hicieron por la fiabilidad de los contenidos, la contextualización de los conceptos y la búsqueda de colocaciones. Los

que lo descartaron indicaron que se debía a razones de tiempo y accesibilidad. Comentaban que es difícil acceder a corpus existentes y compilar uno requiere una inversión de tiempo considerable. Tres participantes indicaron que nunca habían consultado un corpus.

El foro se valoró por ser útil para preguntar a traductores que pueden haberse encontrado con un problema similar, o consultar a especialistas en la materia. Algunos encuestados indicaron que la ventaja de los foros se encuentra en que la consulta a personas no está tan limitada como con las máquinas. Los que rechazaron el uso de foros lo hicieron por razones de tiempo y fiabilidad. Señalaban que no se sabe cuándo se obtendrá una respuesta a la pregunta que se ha formulado ni si será correcta. Muchos encuestados afirmaron que utilizan el foro pero no para realizar consultas sino para buscar intervenciones anteriores en las que pueda haberse resuelto un problema similar.

En las valoraciones de los diccionarios monolingües comentaron que los utilizan sobre todo para acceder a definiciones, aunque sea en otra lengua, y facilitar la búsqueda posterior. Sobre los diccionarios bilingües los participantes explicaron que los utilizan para buscar equivalentes, pero nadie mencionó la búsqueda onomasiológica. La mayoría los criticó por estar obsoletos o incompletos. Sin embargo, dos participantes mencionaron la utilidad de los recursos multilingües, no precisamente diccionarios, porque cuando no se recuerda o se desconoce una denominación en una lengua, puede que la recordemos o conozcamos en otra lengua y podamos buscar a partir de esta. Ningún encuestado utilizaría un diccionario ordenado alfabéticamente para realizar una búsqueda onomasiológica. La razón principal que indicaron es que el usuario desconoce el término que está buscando y estos diccionarios solo permiten buscar a partir de una palabra conocida.

En la cuarta pregunta, sobre estrategias de búsqueda onomasiológica en diccionarios en papel, se mencionó la búsqueda a partir de sinónimos y palabras relacionadas. Una encuestada comentó que no entiende cómo puede realizarse una consulta onomasiológica en un diccionario impreso ordenado alfabéticamente. Los participantes mencionaron el uso de diccionarios ideológicos, de sinónimos y de imágenes y la búsqueda por áreas temáticas. Sin embargo, la mayoría coincidió en que es difícil realizar una consulta onomasiológica en este tipo de diccionarios impresos. Explicaron que los diccionarios de ideas o de imágenes en papel son demasiado complejos, que la búsqueda es lenta y complicada, y que en muchos casos no se encuentra una respuesta satisfactoria.

En la quinta pregunta, sobre estrategias de búsqueda onomasiológica en diccionarios electrónicos se observó una preferencia de los traductores por los diccionarios electrónicos, frente a los diccionarios en papel, debido a la rapidez y accesibilidad de los diccionarios

electrónicos. Para realizar una búsqueda onomasiológica, se nombraron por igual la búsqueda dentro de definiciones y ejemplos y la búsqueda por área temática. Destacó el uso de hipervínculos para acceder a términos relacionados, pertenecientes a una misma categoría, sinónimos, etc. Otras estrategias que aportaron los encuestados fueron la búsqueda de palabras por similitudes de escritura o de fonética, el uso de imágenes y mapas conceptuales y la búsqueda en línea de diccionarios con la introducción de expresiones entrecomilladas en un buscador. Otra función que se mencionó de los diccionarios electrónicos fue la inclusión de foros que complementan al diccionario.

Respecto a la sexta pregunta, sobre estrategias de consulta onomasiológica en corpus electrónicos, casi todos utilizaban todas las estrategias propuestas: la búsqueda por patrones lingüísticos, la búsqueda de concordancias de un término y la búsqueda de palabras cercanas por contexto. La estrategia de consulta más valorada fue la de palabras cercanas utilizando sinónimos e hiperónimos. También se mencionó la búsqueda por intuición y la búsqueda con comodines. Un encuestado indicó que otra estrategia que resultaría útil a los traductores con los corpus electrónicos sería incluir en los programas de análisis textual un mecanismo que permitiera la búsqueda en paralelo de un término en el corpus y en otras fuentes de datos terminológicas, para comparar posibilidades de polisemia. Otro encuestado mencionó la organización de los corpus en árboles conceptuales o temáticos. Una encuestada respondió que nunca había utilizado una herramienta que permitiera realizar las búsquedas que proponíamos en un corpus y otra que no estaba familiarizada con programas para gestionar los corpus. Otros señalaron el difícil acceso a corpus representativos y de calidad.

En la séptima pregunta, sobre estrategias de consulta onomasiológica en foros, la información más utilizada para realizar una consulta de tipo onomasiológico fue el área temática a la que pertenece el término y el contexto en el que aparece. Se mencionaron otros modos de consulta, como el uso de significados (incluso en otros idiomas o explicado con las propias palabras del traductor), la indicación del registro de la palabra y ejemplos de uso.

En la octava pregunta, sobre estrategias de consulta onomasiológica en buscadores de Internet, se empleaban todas las propuestas: la búsqueda de patrones lingüísticos entre comillas, la combinación de palabras relacionadas con el término (sinónimos, hiperónimos, palabras del mismo contexto, etc.) y operadores (AND, OR, NOT), búsqueda de palabras pertenecientes al contexto o al significado en páginas sobre la temática del texto o en los idiomas de la traducción y búsqueda de imágenes.

También se mencionaron las siguientes estrategias: la introducción por intuición de una traducción entre comillas para comprobar si es la correcta, la búsqueda de textos sobre la temática de la traducción para tratar de buscar la palabra en contexto y el uso de herramientas de búsqueda avanzada de *Google*, como *define:*, *info:*, *link:*, *related:*, *site:*, *cache:*, etc.

En la última pregunta, podían añadirse otros recursos y estrategias que resultan útiles para la búsqueda onomasiológica. La mayoría de traductores volvieron a hacer referencia a la consulta directa a especialistas en la materia. Un encuestado de una agencia de traducción mencionó la búsqueda en las bases de datos y en los corpus propiedad de la empresa.

Por último, pedimos a los encuestados que evaluaran la encuesta. A la mayoría le pareció interesante y fácil de comprender. Sin embargo, señalaron que requería demasiado tiempo para pensar y contestar. Algunos encuestados indicaron que era una encuesta demasiado larga y de respuestas demasiado abiertas para ser contestada por escrito. También nos pidieron que incluyéramos más ejemplos de una situación real de traducción en la que se busque una denominación desde un concepto para ayudarles a ponerse en situación.

#### 5.1.4 Discusión

Los objetivos de este estudio piloto eran principalmente tres. En el primero, queríamos conocer qué recursos resultan más útiles a los traductores para la búsqueda onomasiológica y, en el segundo, qué estrategias de búsqueda onomasiológica utilizan. Con ello, nos planteábamos averiguar si los traductores necesitan utilizar estrategias de consulta onomasiológica para resolver sus necesidades terminológicas, es decir, para buscar una denominación desde un concepto, y realizar una pequeña aproximación a las búsquedas onomasiológicas de los traductores. En el tercer objetivo, pretendíamos evaluar si la encuesta que habíamos diseñado para detectar las estrategias de búsqueda onomasiológica utilizadas por los traductores era válida para recabar esa información, con el propósito de mejorar las deficiencias y aplicarla a un número mayor de sujetos.

Respecto a los dos primeros objetivos, recopilamos respuestas que nos resultaron interesantes y nos orientaron para el desarrollo de las siguientes fases de la investigación. De dichas ideas, concluimos que los traductores sí que realizan búsquedas onomasiológicas en los recursos para solucionar determinadas dificultades de terminología. Para ello, se sirven de la información que poseen del concepto (un contexto, una definición o una

palabra relacionada, entre otros). La mayoría de recursos son semasiológicos, lo que significa que no facilitan la implementación de búsquedas de tipo onomasiológico o las posibilidades que ofrecen son limitadas. Sin embargo, de las respuestas de los traductores vislumbramos una serie de estrategias de búsqueda que pueden utilizarse para realizar consultas de tipo onomasiológico, por ejemplo aplicadas a los sistemas de búsqueda de recursos electrónicos, corpus gestionados con programas de análisis textual, foros y buscadores de Internet.

Los recursos más valorados para realizar consultas onomasiológicas son los buscadores de Internet. Todos los encuestados coinciden en que son herramientas imprescindibles que utilizan en todas las traducciones. Las técnicas de búsqueda disponibles en los buscadores permiten recuperar la terminología de los textos de una manera rápida.

Respecto al uso de los corpus para realizar búsquedas onomasiológicas encontramos dos tipos de valoraciones muy diferentes. Por un lado, la mayoría los sitúa por detrás de los buscadores de Internet, e incluso algunos por detrás de los diccionarios electrónicos por motivos de accesibilidad. Los traductores con menos años de experiencia muestran poca familiarización con el uso de programas de análisis textual para la gestión de corpus. Sin embargo, los que dominan este tipo de herramientas consideran a los corpus recursos muy útiles para realizar consultas onomasiológicas. Aún así, todos coinciden en que la creación de corpus *ad hoc* es una labor compleja y, por tanto, se descarta cuando los plazos de entrega de una traducción son reducidos. Los traductores que trabajan para una empresa rechazan la creación de corpus *ad hoc* y mencionan el uso de memorias de traducción y corpus paralelos proporcionados por la empresa.

De todo esto, deducimos que el corpus es un recurso bastante útil para la búsqueda onomasiológica de terminología y que la razón de que no se utilice se debe a su acceso restringido, precisamente por la dificultad de recopilar un corpus representativo de calidad. El escaso uso de los corpus para resolver cuestiones terminológicas vemos que también se debe al desconocimiento de algunos traductores sobre el manejo de corpus.

En cuanto a la valoración de los diccionarios, los monolingües tienen prioridad en comparación con los bilingües porque la información que se obtiene en los diccionarios monolingües es de tipo conceptual (una definición, un ejemplo de uso, un sinónimo o un área temática), mientras que los diccionarios bilingües únicamente aportan equivalentes de palabra a palabra, en la mayoría de ocasiones sin información contextual ni temática.

A partir de la encuesta también deducimos que los traductores consideran que los diccionarios en papel no son útiles para realizar búsquedas onomasiológicas, a no ser que la denominación que se desconoce sea un sinónimo. En este caso, podría buscarse en un diccionario de sinónimos. Asimismo, podría investigarse sobre los diccionarios visuales, ideológicos o temáticos para realizar búsquedas onomasiológicas, ya que se mencionan en las respuestas de la encuesta.

En las encuestas observamos distintas valoraciones de los foros por parte de los traductores. Cuando se valoran en las últimas posiciones es por la necesidad de rapidez en la recuperación de la terminología y por cuestiones de fiabilidad. Sin embargo, esto contrasta con la utilidad que le reconocen otros traductores por el tipo de consulta que permiten, es decir, la consulta a partir de cualquier información relevante del concepto: una definición, un área temática, un contexto, una palabra relacionada, etc., para obtener la denominación adecuada.

En cuanto a las estrategias de búsqueda onomasiológica empleadas en diccionarios electrónicos, corpus, foros e Internet, se confirman las que hemos propuesto en el cuestionario y se incluyen otras. Según esto, la información de la que disponen los traductores para realizar consultas onomasiológicas es muy variada. Se puede partir de una definición, un contexto, un sinónimo, un hiperónimo, un antónimo, un área temática, un registro de uso, una característica, un sonido, una imagen, una palabra que se escribe de forma similar, etc. En definitiva, los modos de acceso a las denominaciones a partir de los conceptos pueden ser muy variados dependiendo del recurso y del propio traductor.

Respecto al tercer objetivo, evaluamos el instrumento que habíamos creado, la encuesta. Comprobamos que las preguntas planteadas son válidas para conseguir la información que precisamos. Las preguntas de la encuesta resultaron ser un buen punto de partida para nuestra investigación. Sin embargo, en próximas investigaciones deberían revisarse y orientarse para obtener aún más datos sobre las estrategias de búsqueda que utilizan los traductores.

Asimismo, nos pareció muy interesante la sugerencia de plantear situaciones concretas de traducción para estimular la imaginación de los traductores y obtener mejores resultados. De hecho, en el estudio principal incluimos un ejemplo real de traducción que requiere aplicar estrategias de consulta onomasiológica.

Finalmente, respecto al instrumento empleado para recopilar la información, la encuesta escrita, detectamos los siguientes inconvenientes. Las preguntas de la encuesta eran de respuesta abierta, lo que resultó en una encuesta demasiado larga o que requería

demasiado tiempo para contestar por escrito, según los propios comentarios de los sujetos, en comparación con otras encuestas de respuesta múltiple que se pueden cumplimentar en menos tiempo (utilizan métodos cuantitativos). Esto incidió en el número de informantes que completaron la encuesta; la encuesta no nos permitió obtener muchos datos en poco tiempo. A su vez, no nos beneficiamos del análisis sencillo de los datos por medios estadísticos, ya que nuestra encuesta era cualitativa.

Por todo lo anterior, nos planteamos que quizá la encuesta escrita no fuera el instrumento de medida idóneo para los objetivos de nuestra investigación. Pensamos que este tipo de preguntas abiertas de tipo cualitativo es mejor plantearlas a modo de entrevista oral semiestructurada, como veremos en los siguientes apartados.

## ***5.2 Estudio principal: entrevista sobre estrategias de búsqueda onomasiológica en traducción***

Nuestro objetivo en el estudio principal es recopilar y estructurar las necesidades y estrategias de búsqueda onomasiológica que utilizan los traductores y observar cómo pueden adaptarse los recursos con el fin de satisfacer estas necesidades. Para ello, realizamos una entrevista oral a un grupo de 48 traductores profesionales. Los resultados obtenidos nos permiten extraer conclusiones sobre el uso de los recursos por parte de los traductores y las estrategias de búsqueda onomasiológica. Con ello, proponemos una serie de ideas para adaptar los recursos a las necesidades (onomasiológicas) de los traductores.

A continuación presentamos el estudio de manera detallada. En el apartado 5.2.1 especificamos los objetivos. Posteriormente, en el apartado 5.2.2 describimos la metodología utilizada: resumimos la teoría en la que nos hemos basado para la realización de nuestra entrevista cualitativa; asimismo, explicamos el diseño de nuestra entrevista, los criterios de selección de los sujetos y los detalles sobre el acceso y contacto, la realización de las entrevistas o recogida de los datos y el análisis de los mismos. Por último, mostramos los resultados (apartado 5.2.3) y exponemos nuestra valoración (5.2.4).

### **5.2.1 Objetivos del estudio empírico**

Los objetivos que nos planteamos al diseñar el estudio empírico son los siguientes:

- Verificar que los traductores tienen **necesidades de tipo onomasiológico**, de acceso a una denominación desde el concepto o significado, y conocer en qué situaciones aparecen dichas necesidades.



- Recopilar y clasificar las diferentes **estrategias de búsqueda** que emplean los traductores profesionales y, dentro de ellas, situar a las **estrategias de búsqueda onomasiológica**, para acceder a una denominación desde el significado.

- Identificar qué tipos de **recursos** consideran más útiles para la resolución de problemas que requieren una búsqueda onomasiológica.

La idea es indagar a partir de qué información o informaciones buscan los traductores para encontrar una denominación que desconocen o no recuerdan. Con ello, pretendemos detectar patrones de estrategias de búsqueda onomasiológica que utilicen los traductores actualmente en las diferentes situaciones que puedan plantearse.

También queremos averiguar qué ventajas y desventajas observan los traductores en cada uno de los recursos que utilizan, así como las demandas de estos usuarios en lo que respecta a la adquisición de recursos que permitan efectuar búsquedas onomasiológicas satisfactorias.

Así, el objetivo último es averiguar cuáles de estas estrategias deberían tenerse en cuenta y potenciarse a la hora de diseñar un diccionario o un recurso onomasiológico, ya que facilitarían el trabajo del traductor en la búsqueda de palabras desconocidas u olvidadas. Con todo ello, pretendemos proponer una serie de requisitos o criterios que deberían tenerse en cuenta en el diseño de recursos útiles para los traductores en una búsqueda onomasiológica.

### 5.2.2 Metodología

La finalidad de nuestra investigación es obtener un panorama general lo más completo posible sobre las estrategias de búsqueda onomasiológica en traducción. Para ello, tuvimos que determinar, en primer lugar, cuál era el instrumento más adecuado para recabar esta información, de entre los que se habían utilizado en otros estudios sobre estrategias de búsqueda de los traductores.

Decidimos que la metodología idónea para satisfacer nuestros objetivos era de tipo cualitativo y que el instrumento que mejor se adaptaba a nuestras necesidades era la entrevista<sup>118</sup>. Como apenas contábamos con estudios previos en traducción que utilizaran y describieran suficientemente la metodología de la entrevista, recurrimos a manuales específicos sobre entrevistas cualitativas que nos resultaron de utilidad a la hora de

---

<sup>118</sup> Véase apartado 5.1.4.

desarrollar nuestra entrevista. A partir de estas sugerencias, diseñamos nuestro instrumento de medida (la entrevista), seleccionamos y accedimos a los sujetos que después entrevistamos y, finalmente, transcribimos y analizamos las entrevistas.

A lo largo de este apartado describimos de forma detallada la metodología empleada en nuestro estudio. En el apartado 5.2.2.1 resumimos la teoría que hemos tenido en cuenta para el diseño de la entrevista, la selección de los sujetos, la realización de las entrevistas y el análisis de la información.

A continuación, en el apartado 5.2.2.2 pasamos a explicar cómo hemos diseñado nuestra entrevista. En el apartado 5.2.2.3 y subapartados, vemos los criterios de selección de los sujetos y el modo de contacto con los mismos. En el apartado 5.2.2.4 nos adentramos en la fase de realización de las entrevistas y, finalmente, en el apartado 5.2.2.5 la fase de análisis de la información.

#### 5.2.2.1 Teoría consultada sobre cómo realizar entrevistas

Para el diseño de la entrevista, la selección de los entrevistados, la recogida de datos y el análisis de los resultados hemos tomado como referencia una serie de manuales sobre *entrevistas cualitativas de investigación o entrevistas en profundidad* (Kvale y Brinkmann 2009; Olaz 2008; Valles 2002; Wengraf 2001). Estos manuales se centran en la entrevista sobre todo aplicada al ámbito de la sociología. No hemos encontrado manuales ni estudios que profundicen en la metodología de la entrevista aplicada específicamente a la investigación en traducción.

Existen distintos tipos de entrevistas cualitativas en función de una serie de criterios. Según el grado de estandarización, si es mayor o menor, puede ser una entrevista *estandarizada estructurada* cuando las preguntas y el orden son igual para todos los entrevistados, *estandarizada no estructurada* cuando a todos los entrevistados se les pregunta por la misma información pero el guion de la entrevista depende del desarrollo de la misma y *no estandarizada* cuando no hay un conjunto fijo de preguntas para todos los entrevistados. Según el modo de realización puede ser una entrevista *oral y por escrito o autoadministrada*. Según el número de participantes, puede tratarse de una *entrevista individual* y *entrevista o discusión en grupo* (Olaz 2008: 27-28; Valles 2002: 24-25; Wengraf 2001: 61).

En la entrevista cualitativa suelen emplearse preguntas de respuesta abierta y suele distinguirse entre preguntas de entrevista y preguntas de investigación, ya que no todas las preguntas se formulan para recabar información, sino que también hay preguntas

*introduccionarias, de seguimiento, especificación, estructuración, interpretación, etc.* (Valles 2002: 59-61). La entrevista cualitativa consta de distintas etapas en las que puede emplearse un guion más o menos estandarizado. La entrevista suele comenzar con una introducción del investigador al tema de la entrevista, donde además se informa sobre la confidencialidad y se pide permiso para grabar, tomar notas, etc. Para ganarse la confianza del entrevistado, el entrevistador puede referirse a la experiencia y la autoridad del entrevistado como experto en la materia. Se suele comenzar con preguntas sencillas con las que el sujeto se sienta familiarizado y, de forma paulatina, se introducen las distintas preguntas de investigación, y después preguntas de seguimiento, especificación, interpretación, etc. Se concluye la entrevista reiterando los agradecimientos por el tiempo y la dedicación del entrevistado y se le pide permiso para mantener el contacto por si hubiera que hacer alguna aclaración o continuación de la entrevista (Olaz 2008: 70; Valles 2002: 101).

Para la selección de los entrevistados, la representatividad de la muestra debe ser de naturaleza más cualitativa que cuantitativa, ya que se prima que los sujetos entrevistados cumplan un determinado perfil y que la calidad de la información que puedan aportar sea relevante, frente a que el número de participantes y de respuestas repetidas sea elevado (Olaz 2008: 25).

El número de entrevistados debe delimitarlo el investigador y este límite debe atender a los objetivos del estudio; autores como Valles (2002: 68) se refieren al *punto de saturación teórica*, cuando hacer más entrevistas no aporta más información relevante. Esta es una característica significativa de las entrevistas cualitativas, ya que el mayor número de participantes no implica necesariamente mayor representatividad de los resultados, en palabras de Kvale y Brinkmann (2009: 109): «Perhaps as a defensive overreaction, some qualitative interview studies appear to be designed on a misunderstood quantitative presupposition –the more interviews, the more scientific».

En la administración de la entrevista, se recomienda efectuar una serie de entrevistas de prueba para familiarizarse con los instrumentos utilizados para recabar los datos (entrevista, toma de notas, grabadora), con el ambiente de la entrevista y las técnicas entrevistadoras. Además, se evalúa la aceptación de las preguntas para observar si se precisan mejoras que favorezcan la obtención de mejores datos y resultados (Valles 2002: 90).

Por regla general, se graban todas las conversaciones, previo permiso de los entrevistados, para su posterior análisis (Olaz 2008: 65). Asimismo, durante la entrevista se

practica una escucha activa (Valles 2002: 110) y se toman notas, *probe notes* (Gorden 1975: 439), es decir, notas breves con palabras clave o expresiones del entrevistado, que sirven para ahondar a lo largo de la entrevista en alguna idea propuesta por el entrevistado.

Para el análisis de las entrevistas se recomienda transcribir las grabaciones (Olaz 2008: 75; Valles 2002: 114 y 135). Las primeras transcripciones suelen ser completas, incluyen toda la conversación grabada. No obstante, a medida que se avanza en las transcripciones, pueden efectuarse transcripciones parciales en función de la teoría que se haya generado con las transcripciones previas (Glaser y Strauss 1967), con la idea de simplificar y economizar los esfuerzos (Kvale 1996).

El investigador deberá decidir hasta qué punto incluye en las transcripciones detalles no verbales como silencios, repeticiones, tonos de voz, propios del lenguaje hablado (Valles 2002: 136). Se advierte de las posibles pérdidas de información que supone realizar la transcripción por escrito de un archivo de sonido (Wengraf 2001: 222), ya que el lenguaje escrito no refleja el tono de voz, la entonación o la respiración del entrevistado, que puede resultar importante en la interpretación de los datos (Kvale y Brinkmann 2009: 178). Asimismo, deberán eliminarse de la transcripción aquellos datos necesarios para garantizar el anonimato de los entrevistados (Valles 2002: 141).

Se recomienda que el investigador transcriba personalmente las entrevistas, ya que de la primera audición de la conversación grabada que se produce en el momento de la transcripción pueden surgir ideas muy valiosas, tanto de método, de análisis e interpretación (Kvale y Brinkmann 2009: 180; Valles 2002: 142; Wengraf 2001: 209-210) que se desperdician si se encarga la transcripción a una tercera persona. Asimismo, independientemente de que se encargue la transcripción a otra persona, se requiere una revisión por parte del entrevistador (Kvale 1996: 168-172).

Los manuales consultados coinciden en que los resultados del análisis cualitativo son igualmente válidos y fiables si este se realiza de forma manual, sin ayuda de programas informáticos. De hecho, incluso con ayuda de la máquina, para ciertos aspectos del análisis, la labor manual del investigador de leer e interpretar los datos es ineludible (Valles 2005: 108-110).

No obstante, tal y como también recomiendan los autores, en la investigación cualitativa, donde se produce una ingente cantidad de información para analizar, los

programas de análisis cualitativo<sup>119</sup> facilitan considerablemente la labor del investigador a la hora de procesar y organizar la información y, aunque no son obligatorios, sí que resultan altamente recomendables (Miles y Huberman 1994: 43 y 56). Por ejemplo, Miles y Huberman (1994: 43) afirman que una semana de trabajo de campo realizando entrevistas puede generar entre 200 y 300 páginas para el análisis.

La mayoría de programas de análisis cualitativo existentes presentan funciones muy similares<sup>120</sup>, aunque los expertos parecen inclinarse por *Atlas.ti* como el más completo e intuitivo de todos (Lewis 1997; Valles 2002: 147; Weitzman 1999: 1261, 2000: 819). Este programa lo desarrolló Muhr (1991) y, como el propio autor indica, fue concebido para la investigación cualitativa en ámbitos como la sociología y la lingüística, pero también pretendía expandirse a otras áreas. Actualmente, se emplea con frecuencia para el análisis de entrevistas cualitativas sobre todo en sociología<sup>121</sup>.

*Atlas.ti* ha resultado útil, además de en estudios sociológicos, en disciplinas tan diversas como las ciencias políticas, la antropología, las ciencias de la educación, la criminología, la economía, la historia, el derecho, los estudios de literatura, entre otros. En cambio, esta metodología cuenta todavía con escaso conocimiento y difusión en lingüística (comparada) y especialmente en el ámbito de los estudios de traducción (Ivanova y Vilar Sánchez 2008: 90). Aún así, se ha demostrado la utilidad del programa *Atlas.ti* en lingüística<sup>122</sup> (véanse, a modo de ejemplo, los estudios presentados en Tabares Plasencia et al. 2008), por lo que también podría resultar de interés en traductología.

*Atlas.ti* se basa en la Grounded Theory (Glaser y Strauss 1967), que puede definirse como sigue: «Stated simply, grounded theory methods consist of systematic, yet flexible guidelines for collecting and analyzing qualitative data to construct theories ‘grounded’ in the data themselves» (Charmaz 2006: 2). Según esta metodología, las teorías resultantes de una investigación se generan de forma inductiva, lo que significa que se basan en el análisis de los datos, por oposición al método deductivo, en el que el investigador cuenta con unas categorías analíticas previas que trata de confirmar con el estudio de los datos.

---

<sup>119</sup> Conocidos generalmente como CAQDAS (Computer Assisted Qualitative Data Analysis Software).

<sup>120</sup> Véanse por ejemplo *Nvivo*, *NUD-IST*, *Aquad*, *Ethno*, *HyperQual*, *HyperResearch*, *Qualog*, *Qualpro*, etc.

<sup>121</sup> Véanse a modo de ejemplo los estudios sociológicos reseñados en Valles (2001) que hacen uso del programa *Atlas.ti* en sus análisis.

<sup>122</sup> Este programa ha demostrado ser útil por ejemplo en la exploración y etiquetado de contextos de un corpus terminológico de la cerámica (Valero y Alcina 2010).

Según afirma Glaser (2001: 127), «GT requires staying open to the emergence of substantive and theoretical codes as opposed to preconceiving them before the research».

Aunque, como también reconoce Charmaz (2006: 47), se puede partir de unos códigos iniciales y después aumentar la lista de códigos o mejorar los existentes a medida que se avanza en el análisis.

Según lo visto en la teoría sobre entrevistas, concluimos que nuestra entrevista sería cualitativa, oral, estandarizada, semiestructurada e individual. Elaboraríamos un guion con preguntas de respuesta abierta que se formularían al entrevistado en función de la información que aportara durante la conversación.

La representatividad de la muestra de sujetos sería cualitativa, en el sentido de que daríamos prioridad a que los sujetos se ajustaran a un perfil, traductores profesionales. Asimismo, decidimos grabar y tomar notas durante las entrevistas para después transcribirlas y analizarlas de una forma más cómoda y ordenada con el programa de análisis cualitativo *Atlas.ti*. Todo ello lo especificamos a continuación.

#### 5.2.2.2 Diseño de la entrevista

Diseñamos una entrevista cualitativa, con preguntas abiertas, oral, individual, estandarizada y semidirigida o semiestructurada, ya que utilizábamos una guía general pero flexible, variábamos o añadíamos preguntas para profundizar o aclarar las respuestas del entrevistado.

En cuanto a la estructura de la entrevista, se tuvieron en cuenta las directrices de los manuales sobre realización de entrevistas consultados; también las sugerencias aportadas en el estudio piloto en el que distribuimos preguntas similares a las de la entrevista, pero a modo de una encuesta escrita. Finalmente, los primeros entrevistados, a los que se les pidió que evaluaran la entrevista, sugirieron otras propuestas que nos permitieron mejorar las entrevistas siguientes.

##### 5.2.2.2.1 El guion y las preguntas de la entrevista

La entrevista comenzaba con una **introducción** en la que se explicaba al entrevistado la investigación y se le informaba sobre el tipo de información que se le iba a requerir: las estrategias de búsqueda onomasiológica en traducción. Además se le recalca la relevancia de la información que podía proporcionar como experto en la materia. Al final de esta introducción, se le garantizaba que la información proporcionada sería anónima, pero se le pedía permiso para citar su nombre en los agradecimientos del estudio; se le

preguntaba si aceptaba participar en la entrevista y se solicitaba el consentimiento para grabar la conversación y tomar notas.

Después de la introducción, se mostraban al traductor una serie de **ejemplos de situaciones de traducción** que pueden motivar una búsqueda onomasiológica, con el fin de que el informante empezara a pensar en ejemplos reales que le hubieran sucedido en el ejercicio de su profesión.

La entrevista que diseñamos constaba de distintas preguntas divididas en bloques. En un **primer bloque** pretendíamos conocer el **perfil del entrevistado**. Se les preguntaba por:

- Años de experiencia en traducción.
- Lenguas de trabajo.
- Especialidades de traducción.
- Modo de ejercer la actividad de traducción (de manera autónoma, en plantilla).
- Los recursos de búsqueda utilizados habitualmente.

Dependiendo del perfil, por ejemplo de la experiencia que tuvieran los traductores, podríamos ver si utilizan estrategias de búsqueda más complejas; también dependiendo de cuáles fueran las lenguas de trabajo o las especialidades que tradujeran, si hacen uso de determinados recursos o búsquedas, o si algunas especialidades son más proclives a generar necesidades de tipo onomasiológico.

En el **segundo bloque** de preguntas se interrogaba sobre las **búsquedas onomasiológicas en traducción**. Para la formulación de las preguntas nos basamos en las estrategias detectadas en los análisis efectuados previamente<sup>123</sup>. Las preguntas de este bloque estaban orientadas a conocer la siguiente información:

- Si el traductor precisa realizar búsquedas onomasiológicas: si tiene esta necesidad en su trabajo diario, con qué frecuencia, en qué casos o con qué tipo de traducciones.
- A partir de qué información buscan los traductores cuando quieren acceder a una palabra que desconocen.
- Qué recursos les resultan útiles para realizar este tipo de búsquedas, qué funciones de los recursos les resultan más provechosas y qué estrategias de búsqueda utilizan.
- Qué reclamarían en los recursos que podría facilitarles la realización de búsquedas onomasiológicas.

---

<sup>123</sup> Véase la propuesta de modelo de clasificación de estrategias de búsqueda en el apartado 4.5 y los resultados del estudio piloto en el apartado 5.1.3.

Incluimos un **tercer bloque** de preguntas para la **evaluación de la entrevista**, que solo utilizamos en las primeras entrevistas para conocer si la entrevista era suficientemente comprensible y clara o si podían sugerirse mejoras.

Se concluía la entrevista reiterando los **agradecimientos** por el tiempo y la dedicación del entrevistado y se le pedía permiso para mantener el contacto por si hubiera que hacer alguna aclaración sobre las respuestas.

El guion de la entrevista se redactó en español. Asimismo, se realizó una traducción al inglés para utilizarla con aquellos entrevistados que prefirieran realizar la entrevista en este idioma. Dos de las entrevistas<sup>124</sup> se efectuaron en inglés, a petición de los entrevistados. La entrevista completa puede consultarse en el anexo 4.

### 5.2.2.3 Sujetos

La representatividad de la muestra de sujetos pretendimos que fuera de naturaleza cualitativa, ya que priorizamos que los sujetos entrevistados tuvieran un perfil concreto (traductores profesionales y expertos) y que pudieran aportar información de calidad; no procuramos que el número de participantes y de respuestas repetidas fuera muy elevado.

A continuación detallaremos qué clase de traductores seleccionamos para participar en la entrevista y cómo accedimos y contactamos con ellos.

#### 5.2.2.3.1 Selección de informantes

El perfil de los sujetos seleccionados para participar en nuestro estudio debía cumplir un requisito esencial: ser traductores profesionales y, además, que su única o principal actividad profesional fuera la traducción.

Este criterio atendió a que pretendíamos conocer las estrategias de búsqueda onomasiológica en el proceso de traducción «real», o en el mundo profesional. Pensamos que con vistas a desarrollar recursos que faciliten las búsquedas onomasiológicas a los traductores interesaban las opiniones de los usuarios potenciales de estos recursos, que serían previsiblemente los traductores profesionales.

De los estudios examinados en el capítulo 4, habíamos observado que las estrategias de búsqueda de los estudiantes o traductores ocasionales son más pobres, o escasas. Los traductores profesionales utilizan más herramientas y estrategias de búsqueda avanzadas, debido a la complejidad de su trabajo. Según Jensen and Jakobsen (2000: 114) «the

---

<sup>124</sup> Entrevistas 044 y 046.



translational behaviour of young professionals may be closer to that of semi-professionals (graduate students) than to that of professionals with eight or more years of experience». Por tanto, los traductores profesionales expertos podrían aportarnos más ejemplos de búsquedas avanzadas para acceder a las denominaciones de forma onomasiológica.

En definitiva, a diferencia de la mayoría de estudios sobre estrategias de búsqueda en traducción que utilizan como sujetos a estudiantes de traducción, principalmente por su mayor accesibilidad, nosotros decidimos centrarnos únicamente en traductores profesionales. Al mismo tiempo, dimos prioridad a los traductores profesionales con bastantes años de experiencia y cuya dedicación durante dichos años hubiera sido exclusivamente el sector de la traducción. Además, recurrimos a traductores con variedad de lenguas, especialidades y situaciones; con ello, pretendíamos vislumbrar posibles tendencias generales en las búsquedas onomasiológicas de los traductores.

#### 5.2.2.3.2 Acceso y contacto

Ideamos un modo de acceso y contacto con los traductores profesionales para preguntarles si estaban dispuestos a participar en la entrevista.

Para organizar el acceso a los sujetos de la entrevista, elaboramos un registro en Excel en el que incluimos a la persona de contacto, su dirección de correo electrónico, teléfono, dirección postal y comentarios que debiéramos tener en cuenta al realizar el contacto. Los primeros contactos se obtuvieron principalmente a través de los gerentes de agencias de traducción colaboradoras en el Máster en Tecnologías de la traducción y Localización de la Universitat Jaume I, el Servicio de Lenguas del Consejo de la Unión Europea, agencias de traducción o traductores autónomos que se anuncian en Internet o en las páginas amarillas y otros contactos.

Comenzamos con 105 contactos iniciales. En primer lugar, probamos a contactar con ellos por teléfono, pero ese método no funcionó, ya que fomentaba el rechazo de los traductores. Por este motivo, empleamos otro modo de contacto que fue el correo electrónico. Se les envió un mensaje inicial, en el que se les explicaba el tipo de investigación que estábamos realizando, la metodología empleada y se les sugería participar en la entrevista. El mensaje completo puede consultarse en el anexo 4.

A este correo electrónico algunos traductores no respondieron y otros no participaron en la entrevista por distintos motivos (sobre todo razones de tiempo y económicas, ya que se pedía una participación desinteresada; algún traductor tampoco participó a causa del

perfil, porque ya no se dedicaba a la traducción de forma profesional). Algunos de los traductores que no participaron nos proporcionaron nuevos contactos.

A aquellos traductores que respondían de manera positiva a nuestra petición, se les facilitaba información adicional, como el posible modo de realización de la entrevista (en persona, por teléfono o por Skype) y la duración aproximada de la entrevista (20 minutos). Además, se les enviaba una introducción a la entrevista con un ejemplo de situación que puede propiciar una búsqueda onomasiológica para que los traductores reflexionaran durante el ejercicio de su trabajo con anterioridad a la entrevista. Véase el anexo 4, donde incluimos una introducción a la entrevista y un ejemplo de situación.

Finalmente, en el caso de que la entrevista fuera a ser en persona, se acordaba con ellos un lugar de reunión, o si iba a efectuarse por teléfono o *Skype*, se concretaba el número de teléfono o el nombre de usuario de *Skype* con el que preferían que nos comunicáramos. También se concertaba una fecha y una hora para la realización de la entrevista.

#### 5.2.2.4 Realización de las entrevistas

Se realizaron **49 entrevistas**, de las cuales **17 fueron en persona** (en el despacho del investigador, en el del entrevistado o en una cafetería, si así lo pedía el entrevistado), **17 por teléfono** (con el uso de un manos libres para poder grabar la conversación) y **15 por Skype**, en todos los casos previo acuerdo con el entrevistado. Los contactos y las entrevistas con los traductores se efectuaron en un espacio de tiempo de 8 meses, desde diciembre de 2010 a julio de 2011.

La primera entrevista la efectuamos a modo de prueba (no la incluimos en el análisis de resultados). El objetivo fue familiarizarnos con los instrumentos de medida: el guion de la entrevista y las preguntas, el uso de la grabadora y la toma de notas, así como practicar las técnicas entrevistadoras aprendidas en la teoría. Además, pudimos evaluar las preguntas y realizar los últimos cambios para mejorar la comprensión de las mismas por parte de los entrevistados. En esta entrevista participó una docente universitaria de traducción.

Después, pasamos a la **fase de realización de entrevistas con traductores profesionales**. Entrevistamos utilizando un guion impreso de la entrevista como orientación para realizar las preguntas y que también nos sirvió para tomar notas. Sin embargo, no se trataba de un guion cerrado, en función de las respuestas del entrevistado, formulábamos unas preguntas u otras, o simplemente dejábamos hablar al entrevistado en lugar de hacerle más preguntas. Interveníamos solo cuando lo considerábamos oportuno

para asistir al informante en el caso de que se hubiera bloqueado en alguna respuesta, aunque la premisa que prevaleció fue la de intervenir lo mínimo posible.

Grabamos todas las conversaciones y además realizamos una escucha activa. Debajo de las preguntas del guion de la entrevista habíamos dejado espacio suficiente para anotar cuestiones que consideráramos relevantes. Tomábamos notas breves sobre aquellos comentarios del entrevistado en los que queríamos profundizar posteriormente.

Las entrevistas tuvieron una duración prolongada, de entre 20 y 90 minutos. Al tratarse de una entrevista cualitativa, abierta y semiestructurada pudimos ahondar en las búsquedas onomasiológicas de cada sujeto, que nos proporcionaba nuevos datos y perspectivas. No obstante, tras pasado el umbral de las 40 entrevistas, alcanzamos nuestro punto de saturación teórica, es decir, las respuestas empezaron a repetirse y no nos aportaban información nueva, por lo que con un total 48 entrevistas (sin contar la primera entrevista de prueba) decidimos concluir con la fase de realización de entrevistas.

Algunos traductores, días después de la entrevista, nos escribieron para explicarnos más búsquedas onomasiológicas que habían realizado durante su trabajo. Los correos electrónicos que recibimos nos ayudaron a completar los datos que habían aportado durante la conversación y nos proporcionaron información muy valiosa y real.

### 5.2.2.5 Análisis de la información

El análisis de la información recabada en las entrevistas constó principalmente de dos fases: transcripción y codificación. En este apartado describiremos, en primer lugar, cuáles han sido los programas empleados para el análisis. Después, explicaremos cómo fueron las fases de transcripción y posterior codificación de las entrevistas.

#### 5.2.2.5.1 Recursos utilizados para el análisis

Los programas que utilizamos para transcribir fueron *F4* (versión 3.1), un programa para transcripciones, y *Dragon Naturally Speaking* (versión 11), un programa de reconocimiento de voz. En la fase de codificación nos ayudamos del programa de análisis cualitativo *Atlas.ti* (versión 6.2.27). A continuación veremos con más detalle las funciones que nos han resultado útiles de estos programas.

#### 5.2.2.5.1.1 Programas utilizados para la transcripción: *F4* y *Dragon Naturally Speaking*

*F4* (versión 3.1) es un programa específico para la realización de transcripciones; lo elegimos porque es compatible con *Atlas.ti*, los propios creadores de *Atlas.ti* y los usuarios de los foros de su página web lo recomiendan. En el momento en el que utilizamos *F4* era software gratuito<sup>125</sup>, se descargaba fácilmente desde la página web del programa<sup>126</sup>, aunque previamente era necesario registrarse.

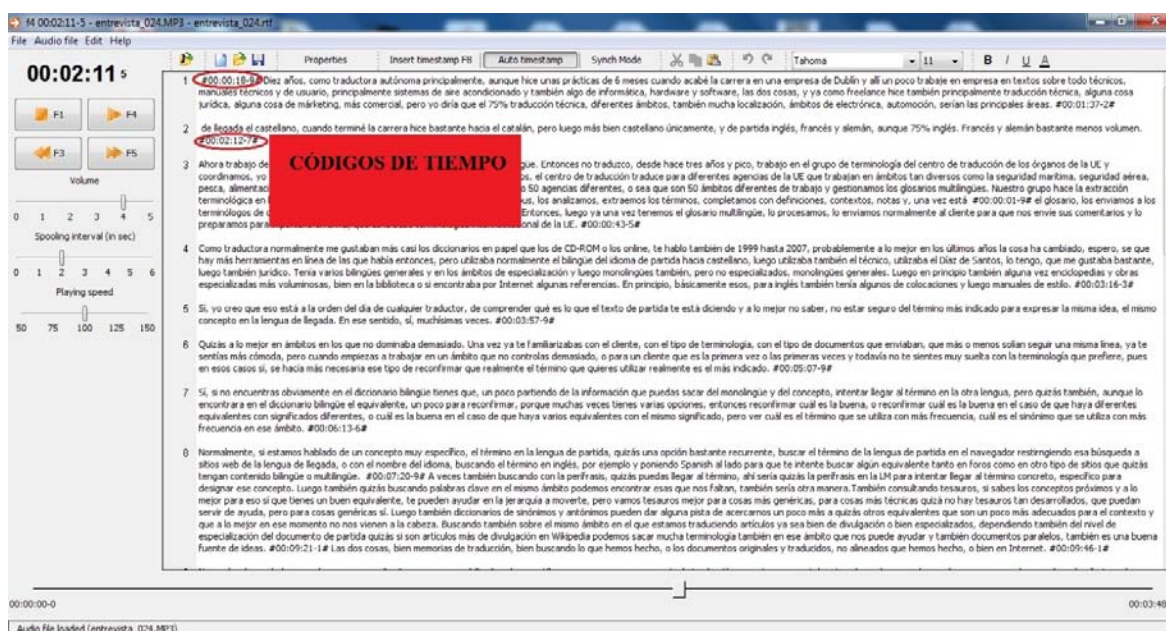
Con este programa es posible abrir un archivo de sonido, escucharlo y transcribirlo al mismo tiempo en el panel de edición. El uso de las teclas de función permite parar, avanzar o retroceder en la grabación sin necesidad de apartar las manos del teclado ni abandonar la pantalla de transcripción. También es posible aumentar o disminuir la velocidad de reproducción. Una vez terminada la transcripción, se puede guardar el archivo de texto resultante de la transcripción en formato rtf.

Además, la razón que nos llevó a utilizar el programa *F4* para la transcripción es que permite insertar códigos de tiempo a lo largo de la transcripción (como se muestra en la Ilustración 100) y, después, las transcripciones que se generan en formato rtf están vinculadas y sincronizadas con los archivos de sonido. Posteriormente, estos documentos vinculados pueden importarse en *Atlas.ti*.

---

<sup>125</sup> La fecha de descarga fue febrero de 2011. En dicha fecha el programa era gratuito y era posible realizar donaciones que facilitaban que sus creadores financiaran el proyecto. Actualmente, existen tres versiones de *F4*, una versión gratuita de prueba que solo permite reproducir hasta 10 minutos de grabación y una versión *F4 plus* y *F4 pro* de pago con funciones adicionales.

<sup>126</sup> <http://www.audiotranskription.de/english/f4.htm>



**Ilustración 100. Transcripción en el programa F4 con códigos de tiempo.**

*Dragon Naturally Speaking* (versión 11) es un programa privativo y comercial<sup>127</sup>. Presenta dos funciones que probamos en nuestro proceso de transcripción. La primera de ellas es la función de transcripción automática, con la cual el programa transcribe de forma automática el archivo de audio proporcionado. La segunda función es la de dictado, mediante la que el programa identifica el discurso dictado con un micrófono y lo convierte en texto.

La función de transcripción automática del programa de reconocimiento de voz *Dragon Naturally Speaking* no funcionó correctamente con nuestros archivos de audio. Probablemente, esto se deba a que contenían voces distintas (la del entrevistador y la del entrevistado) y además algunas entrevistas tenían sonido de fondo<sup>128</sup>. En cualquier caso, la transcripción automática proporcionó transcripciones inservibles incluso en aquellos archivos de audio más limpios, sin ningún tipo de ruido ambiental. El programa solo fue capaz de captar palabras sueltas de toda una grabación. El resultado fueron transcripciones sin sentido alguno.

<sup>127</sup> La adquisición de este programa se efectuó con fondos del proyecto *ONTODIC II: Metodología y técnicas para elaborar diccionarios de colocaciones basados en ontologías. Recursos terminológicos para la e-Traducción*, TIN2009-07690, subprograma TSI, financiado por el Gobierno de España.

<sup>128</sup> Algunas entrevistas se habían realizado por ejemplo en una cafetería, a propia petición de los entrevistados. En estos casos, dimos prioridad a que los entrevistados se sintieran cómodos, tal y como sugieren los expertos, aunque el ruido supusiera un esfuerzo añadido en la transcripción.

La **función de dictado** de este programa tampoco resultó cómoda ni agilizó el proceso de transcripción, incluso después de haber realizado los pasos previos de entrenamiento del programa y de haber transcrito varias entrevistas utilizando este método.

Para que el programa reconociera correctamente la voz, el ritmo de dictado debía ser lento, pronunciando todas las sílabas, por tanto, la velocidad de dictado siempre era más lenta que la del discurso del entrevistado en la grabación. Esto provocaba que fuera necesario parar la grabación constantemente.

Además, aunque el programa sí que reconocía la voz y transcribía correctamente, también cometía errores, tanto en la inserción de palabras equivocadas (en lugar de la palabra adecuada escribía otra) como en la ortografía incorrecta (por ejemplo muchas palabras las escribía sin acentos). El problema de los acentos pudo solucionarse rápidamente pasando posteriormente el corrector de textos, pero cuando se insertaban palabras erróneas era necesario volver a repasar la grabación. Así, el tiempo que se había adelantado al no tener que teclear manualmente la transcripción, se perdía por la necesidad de volver a escuchar y corregir los errores cometidos por el programa de reconocimiento de voz.

A continuación aportamos algunos cálculos que efectuamos para valorar la eficacia del programa de reconocimiento de voz *Dragon Naturally Speaking* en la transcripción de nuestras entrevistas.

Al principio del proceso de transcripción se efectuó una primera medición, una vez que se había completado el entrenamiento inicial del programa, pero sin haber realizado transcripciones con él.

- Medición del tiempo de transcripción con dictado, utilizando el programa *Dragon Naturally Speaking*:

- Código de la entrevista: Entrevista\_048
- Total de minutos originales transcritos: 6,85 minutos
- Tiempo dedicado a la transcripción: 32 minutos
- Por cada minuto de audio original, se tardó 4,67 minutos en transcribir<sup>129</sup>

---

<sup>129</sup> Este es el total de minutos dedicados a la transcripción y corrección de la transcripción. En esta primera medición, no calculamos por separado el tiempo de transcripción con dictado y el tiempo de corrección manual. En mediciones posteriores sí que tuvimos en cuenta estos parámetros.

- Medición del tiempo de transcripción manual<sup>130</sup>, sin el programa *Dragon Naturally Speaking*:

- Código de la entrevista: Entrevista\_045
- Total de minutos originales transcritos: 8,05 minutos
- Tiempo dedicado a la transcripción: 24 minutos
- Por cada minuto de audio original, se tardó 2,98 minutos en transcribir

En esta primera prueba la transcripción con la función de dictado del programa de reconocimiento de voz resultó más lenta que la transcripción manual. El principal inconveniente fue que el programa cometía muchos errores y se tardaba más tiempo en volver a escuchar la grabación para corregirlos que en hacer la transcripción de forma manual. No obstante, se pensó que este sería un problema inicial pero que el resultado del dictado mejoraría a medida que se entrenara al programa con las siguientes transcripciones.

Así, una vez que se había avanzado en el proceso de transcripción, se efectuó una nueva medición. Probamos a transcribir un mismo fragmento de entrevista<sup>131</sup> (del minuto 18:10 a 19:19 de la entrevista\_039), con la ayuda del programa *Dragon Naturally Speaking* y con transcripción manual.

El tiempo de transcripción manual sin el programa *Dragon Naturally Speaking* fue de 2,36 minutos. El tiempo de transcripción con dictado utilizando el programa *Dragon Naturally Speaking* fue de 2,11 minutos. Sin embargo, tuvimos que realizar una corrección de la transcripción, por lo que el tiempo final de transcripción con dictado fue de 3,46 minutos.

La conclusión que extraemos de las mediciones efectuadas es que el programa de reconocimiento de voz no resultó rentable para la tarea de transcripción de nuestras 48 entrevistas. Esto probablemente se deba a que el lenguaje oral tiene demasiadas marcas, muletillas, vacilaciones, cambios en medio de una frase, por lo que el reconocedor no es capaz de captar el sentido.

No obstante, el programa de reconocimiento de voz quizá podría ser útil o rentable a muy largo plazo, con mucho entrenamiento, si no fuera necesario realizar una corrección manual. De hecho, ha demostrado ser rentable en el dictado de otros textos, por ejemplo en el ámbito de la traducción (Serrahima 2009a, 2009b).

---

<sup>130</sup> Con una velocidad media de mecanografiado de 47 palabras por minuto y 233 pulsaciones.

<sup>131</sup> Con ello, tratamos de evitar que los resultados de las mediciones pudieran estar influenciados por utilizar dos fragmentos que presentaran distintas peculiaridades.

#### 5.2.2.5.1.2 Programa utilizado para la codificación: *Atlas.ti*

Elegimos *Atlas.ti* (versión 6.2.27), frente a otros programas similares, por motivos de disponibilidad<sup>132</sup> y también por las recomendaciones de los manuales sobre realización de entrevistas cualitativas que indicaban que *Atlas.ti* es el programa de análisis cualitativo más completo e intuitivo.

La herramienta de análisis cualitativo *Atlas.ti* es un programa privativo y comercial. En él, un proyecto se denomina *unidad hermenéutica o hermeneutic unit*. En la unidad hermenéutica pueden importarse los *documentos primarios o primary documents* (PD) para examinar (este programa permite el análisis de gran variedad de documentos primarios, desde archivos de sonido, de texto, imágenes, etc.).

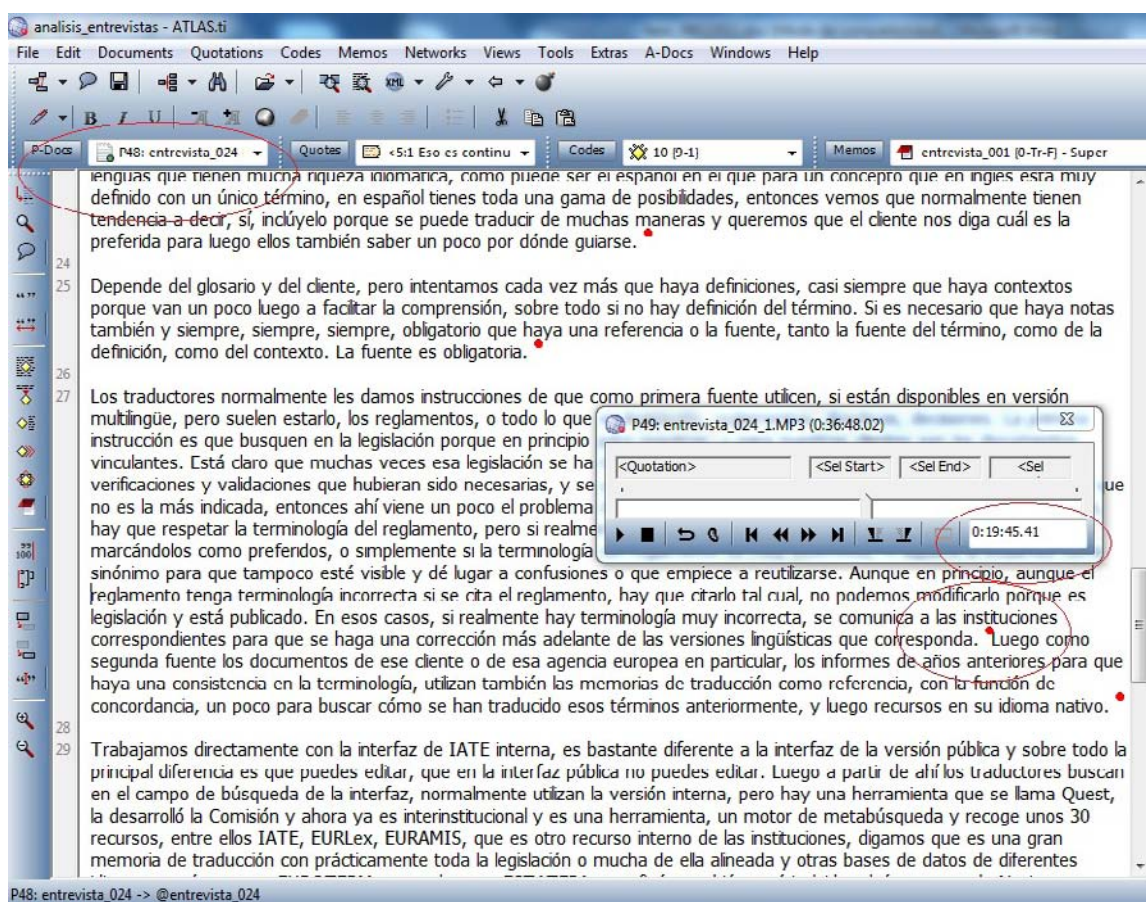
Por ejemplo, *Atlas.ti* permite importar los documentos de texto de las transcripciones y los archivos de audio originales, que se encuentran sincronizados previamente al haber insertado códigos de tiempo en las transcripciones con el programa *F4*. De esta manera, el texto transcrito es el que se etiqueta en el análisis, pero en todo momento se puede acceder rápidamente al fragmento de sonido original haciendo clic en el código de tiempo deseado.

Los códigos de tiempo se importan en el programa *Atlas.ti* y se muestran en forma de puntos rojos. Como podemos observar en la Ilustración 101, seleccionada una transcripción, si deseamos escuchar algún fragmento del archivo de sonido original, con el modo de sincronización activo, presionamos la tecla F4 y hacemos clic en el punto rojo que aparece delante del fragmento de texto «Luego como segunda fuente...». De esta manera, comienza directamente la reproducción de este fragmento, en el minuto 0:19:45.41.

---

<sup>132</sup> La adquisición de este programa se efectuó con fondos del proyecto *ONTODIC II: Metodología y técnicas para elaborar diccionarios de colocaciones basados en ontologías. Recursos terminológicos para la e-Traducción*, TIN2009-07690, subprograma TSI, financiado por el Gobierno de España.





**Ilustración 101. Sincronización de las transcripciones y los archivos de sonido originales en *Atlas.ti*.**

El análisis en *Atlas.ti* consta principalmente de dos fases, una textual y una conceptual (tal y como se describe en Muhr 1991: 349). Por un lado, se empieza con una *fase textual* en la que se exploran las entrevistas transcritas y se segmentan asignando *códigos (codes)*. Cada uno de estos fragmentos etiquetados se denomina *cita* o *quote*<sup>133</sup>.

Al lado de los códigos aparecen dos números entre llaves, por ejemplo {3-5}. El primer número hace referencia a la cantidad de citas que se han etiquetado con dicho código (*groundedness* o *fundamentación, validez, solidez*). El segundo número es la cantidad de otros códigos que están conectados de alguna manera a este (por ejemplo porque tienen palabras en común), determina la *densidad (density)* del código (véase Martínez Miguélez 2001). Por tanto, según este autor, la importancia de una categoría la determina la *validez*, que esté plasmada en muchas citas, y la *densidad teórica*, el conjunto de relaciones con otras categorías.

<sup>133</sup> Según Martínez Miguélez (2001), los *códigos* también se denominan *índices, palabras clave* o *categorías*, un término más apropiado, según este autor, porque los códigos «...son algo más, o mucho más, es decir, son “ideas”, “conceptos” o “interpretaciones” abreviadas de las citas, y son la “materia prima” para construir, después, las teorías o conclusiones teóricas, fin último de la investigación».

Asimismo, es posible insertar *comentarios (comments)* que expliquen el significado de documentos primarios o códigos, es útil para fijar reglas si el análisis se va a realizar en grupo. La presencia de comentarios en los códigos se marca con el símbolo ~.

Por otro lado, se procede a una *fase conceptual*, en la que se elaboran redes conceptuales que permiten diseñar estructuras teóricas que son la base en la presentación de los resultados del análisis. Las *familias o families* son agrupaciones de códigos, notas o documentos. Las familias pueden agruparse en *superfamilias o super families*. A su vez, pueden establecerse *redes conceptuales o networks* en las que se establecen las relaciones entre códigos, notas, documentos, familias, citas, etc.

A lo largo de todo el proceso, el investigador puede introducir también *notas o memos*, que contienen ideas que surgen a medida que se realiza el análisis.

Para la recuperación de los resultados del análisis, con la ayuda de la *herramienta de búsqueda o query tool* es posible localizar todas las citas etiquetadas con uno o más códigos o familias (es posible utilizar operadores de búsqueda avanzada).

También existen opciones de exportación asociadas con los documentos, códigos, citas y notas. Entre ellas, por ejemplo, es posible exportar a un archivo de texto el conjunto de códigos del análisis con sus citas correspondientes, las citas de los documentos seleccionados, una lista de notas con o sin el texto de las notas y las citas con las que están relacionadas, etc.; o exportar en Excel una tabla con todos los códigos del proyecto o los códigos seleccionados y el recuento de citas etiquetadas con cada código en los distintos documentos. También existen opciones de exportación en formato SPSS, XML y HTML.

Finalmente, como usuarios de *Atlas.ti*, nos gustaría indicar una serie de sugerencias que podrían incorporarse en futuras versiones del programa que podrían facilitar el análisis de los datos. En la vista *Network View Manager*, no se ofrece la función de búsqueda para localizar códigos o nodos de forma automática dentro de una red conceptual. Para encontrar un nodo, el investigador debe buscarlo manualmente en la red, con lo que puede perderse tiempo cuando se trabaja con redes complejas, compuestas de múltiples nodos y relaciones. Asimismo, el tamaño de la cuadrícula (*grid*) que dispone *Atlas.ti* para la creación de una red tiene un tamaño determinado y no puede ampliarse en el caso de que sea necesario crear una red más amplia.

#### 5.2.2.5.2 La fase de transcripción: preparación de las entrevistas para el análisis

Antes de comenzar a analizar las entrevistas, efectuamos una transcripción, es decir, pasamos la información proporcionada por los traductores que habíamos recopilado en formato de audio a formato texto para poder codificar posteriormente los datos. Aunque *Atlas.ti* permite analizar y codificar directamente los archivos de audio, sin necesidad de ser transcritos, decidimos trabajar con transcripciones de texto porque los fragmentos de texto de las transcripciones etiquetados durante el análisis nos servirían de apoyo para presentar los resultados del estudio. Además, actualmente este es el método que se emplea mayoritariamente en las investigaciones con entrevistas cualitativas.

Comenzamos el proceso realizando transcripciones completas de las entrevistas, que incluían toda la conversación grabada. No obstante, a medida que avanzábamos en las transcripciones, realizamos transcripciones parciales, solo de aquella información que posteriormente fuéramos a analizar, es decir, las respuestas de los entrevistados que consideramos relevantes para los objetivos de nuestro estudio.

Por ejemplo, transcribimos toda la información proporcionada por los traductores sobre el uso de recursos y estrategias de búsqueda. Sin embargo, descartamos los comentarios que no estaban directamente relacionados con la investigación. Por ejemplo, no transcribimos comentarios personales de los traductores externos a la investigación. Tampoco transcribimos las preguntas formuladas al entrevistado, a no ser que fueran necesarias para comprender las respuestas. Por ejemplo, no transcribimos las preguntas que formaban parte del guion de la entrevista.

Obviamos detalles no verbales como silencios, repeticiones, tonos de voz, propios del lenguaje hablado, salvo en el caso de que resultaran relevantes para el análisis posterior y eliminamos aquellos datos necesarios para garantizar el anonimato de los entrevistados.

Por último, nos encargamos personalmente de transcribir las entrevistas porque de cara al análisis de la información la primera audición de la conversación grabada, que se produjo en el momento de la transcripción, nos proporcionó ideas muy valiosas de análisis e interpretación.

### 5.2.2.5.3 La fase de codificación: análisis de las entrevistas

En la fase de análisis de las entrevistas, contábamos con un total de 418 páginas de información para analizar. Al tratarse de un **análisis de tipo cualitativo**, utilizamos un programa de análisis cualitativo (*Atlas.ti*), con el que armonizamos y agrupamos los datos obtenidos a partir de las respuestas. Esto nos facilitó la extracción de resultados susceptibles de valorar.

Los documentos primarios que analizamos fueron, por un lado, las transcripciones de las entrevistas que estaban sincronizadas con los archivos de sonido originales y, por otro lado, la copia de una serie de correos electrónicos enviados por algunos entrevistados con posterioridad a la entrevista, donde nos aportaban información adicional sobre estrategias de búsqueda.

Cada una de las transcripciones estaba identificada con el nombre *entrevista* seguido de un *número* (por ejemplo *entrevista\_001.rtf* era la transcripción de la primera entrevista realizada, tal y como figuraba en el registro de las entrevistas). A su vez, el archivo de sonido sincronizado con esta entrevista se denominaba *entrevista\_001.MP3*. Finalmente, en el caso de que el entrevistado hubiera enviado un correo electrónico con información adicional posterior a la entrevista, había un documento primario que se denominaba *entrevista\_001\_adicional.rtf*.

Nuestro análisis tuvo dos fases, textual y conceptual. En la **fase textual**, empezamos con la lectura de las transcripciones. Como en cada entrevista la transcripción y el archivo de sonido se encontraban sincronizados, durante el análisis escuchábamos el fragmento original de la entrevista en el caso de que tuviéramos dudas sobre la transcripción. Esto nos ayudaba a comprender mejor la idea que quería transmitir el entrevistado, que a veces no quedaba clara solo con la transcripción. También, en el caso de que hubiera un error de transcripción, lo detectábamos escuchando de nuevo la entrevista original y lo corregíamos. Con esta estrategia mitigamos la pérdida de información (tono de voz, entonación, pausas) que supone analizar la transcripción por escrito de un archivo de sonido.

Además, identificábamos aquellos fragmentos de texto que expresaban una idea relevante para nuestra investigación y los etiquetábamos con códigos. El mismo código se asignó a todos los fragmentos de entrevistas que reflejaban la misma idea. Asimismo, insertamos comentarios que describían algunos códigos para evitar que se etiquetara la misma idea con códigos distintos.

En este sentido, el método de análisis de la Grounded Theory (Glaser y Strauss 1967), en el que se basa *Atlas.ti*, nos pareció adecuado para nuestra investigación. Por un lado, partíamos solo de algunas ideas preconcebidas, a partir de las cuales habíamos generado unos códigos iniciales. Por otro lado, a medida que avanzaba el análisis descubrimos nuevas ideas o complementamos las existentes, por lo que aumentamos la lista de códigos o mejoramos los que ya teníamos.

Así, para el análisis de las primeras entrevistas tuvimos en cuenta los códigos generados del análisis de los estudios sobre estrategias de búsqueda en traducción (véase el apartado 4.5), a partir de los cuales habíamos creado una propuesta preliminar de clasificación de estrategias de búsqueda en traducción.

En esta clasificación utilizábamos códigos formados por una combinación de letras y números. Las letras señalaban los problemas o necesidades que motivaban una estrategia de búsqueda<sup>134</sup>: *LO* si es una necesidad en lengua origen y *LM* si es una necesidad en lengua meta. Además, la letra *a* si es una necesidad de recepción, la *b* si es una necesidad de transferencia y la *c* si es una necesidad de producción. Por otro lado, los números los añadíamos para catalogar las acciones de búsqueda en el orden habitual de utilización<sup>135</sup>. La lista inicial completa de los códigos y sus explicaciones sobre las estrategias de búsqueda que representan puede consultarse en el anexo 3. A continuación presentamos una síntesis:

- LO.a: estrategias de búsqueda para resolver problemas de recepción en lengua origen.
- LO.a.1: consulta de una denominación conocida en LO en un recurso monolingüe para obtener el significado desconocido.
- LO.a.2: consulta de una denominación conocida en LO a una persona (hablante nativo o especialista) para obtener el significado desconocido.
- LM.b: estrategias de búsqueda para resolver problemas de transferencia a la LM.
- LM.b.1: consulta de una denominación conocida en LO en un recurso bilingüe para obtener una denominación equivalente desconocida en LM.
- LM.b.2: consulta de una denominación conocida en una lengua intermedia entre la LO y la LM para utilizarla como puente en un recurso bilingüe para obtener una denominación equivalente desconocida en LM.
- LM.b.3: consulta de una hipótesis de equivalente en LM. A partir de la consulta de una intuición de equivalente, el traductor selecciona el equivalente desconocido.
- LM.b.4: consulta de una hipótesis de palabras relacionadas o palabras clave que pueden aparecer en el contexto del equivalente en LM desconocido; el traductor localiza el equivalente desconocido dentro de ese contexto.

---

<sup>134</sup> Véase el apartado 4.5.2.2.1.

<sup>135</sup> Véase el apartado 4.5.2.2.2.

- LM.b.5: lectura de textos paralelos para localizar un equivalente desconocido en LM.
- LM.b.6: localización por casualidad de un equivalente desconocido en LM.
- LM.b.7: consulta a un foro.
- LM.a.verificación: estrategias de recepción en LM (para verificar el significado de un equivalente en LM).
- LM.a.1: consulta de una denominación en LM en un recurso monolingüe para verificar el significado.
- LM.a.2: consulta de una denominación en LM a una persona (hablante nativo o especialista) para verificar el significado.
- LM.c: estrategias de producción en LM.
- LM.c.1: consulta de una denominación en LM para obtener información que oriente al traductor sobre el uso correcto en contexto (por ejemplo, colocaciones).

Con estas ideas de partida, pudimos comprobar cuáles de estas estrategias utilizaban los traductores profesionales entrevistados en nuestro estudio y cuáles se utilizaban como estrategias de búsqueda onomasiológica. Además, a medida que avanzaba el análisis pudimos detectar otras estrategias que habían reseñado los traductores en nuestras entrevistas sobre búsquedas onomasiológicas y, por tanto, generamos nuevos códigos, siguiendo el mismo modelo de combinación de letras y números. Por ejemplo, algunas de las nuevas estrategias detectadas en el análisis se etiquetaron con los siguientes códigos (todas las estrategias se explican detalladamente en el apartado de resultados).

- LO.a.3: consulta de una denominación conocida en LO al cliente para obtener el significado desconocido.
- Estrategias de tipo LM.b.1
  - oLM.b.1.5: consulta de una denominación conocida en LO en una enciclopedia bilingüe en línea o en un buscador de Internet de páginas bilingües para obtener una denominación equivalente desconocida en LM.
  - oLM.b.1.6: consulta de una denominación conocida en LO en una herramienta de metabúsqueda para obtener una denominación equivalente desconocida en LM.
- Estrategias de tipo LM.b.4
  - oLM.b.4.1: consulta a partir de palabras relacionadas sintagmáticamente con la denominación desconocida en LM.
  - oLM.b.4.2.1: consulta a partir de un sinónimo para acceder a una denominación desconocida en LM.
  - oLM.b.4.2.2: consulta a partir de un hiperónimo para acceder a una denominación desconocida en LM.
- LM.b.8: consulta directa a especialistas.

También creamos nuevos códigos para etiquetar ideas relacionadas con las ventajas, desventajas y demandas que observaban los traductores respecto de distintos recursos, por ejemplo:

- descontento\_diccionarios
  - o descontento\_diccionario\_bilingüe
    - pocos\_diccionarios\_bilingües\_especializados
    - terminología\_compartimentada<sup>136</sup>
    - terminología\_descontextualizada<sup>137</sup>
    - terminología\_desfasada<sup>138</sup>
  - o descontento\_diccionario\_monolingüe
- ventaja\_Internet
  - o ventaja\_recurso\_online
  - o ventaja\_rapidez
  - o terminología\_poco\_habitual
- demanda\_Internet
  - o demanda\_filtrar\_búsqueda
  - o demanda\_búsqueda\_semántica\_imágenes
  - o destacar\_palabras\_búsqueda\_en\_contexto
  - o búsqueda\_paralela\_libros\_traducidos

Los nombres de diccionarios o de otros recursos utilizados por los traductores también se recogieron en forma de códigos, por ejemplo:

- BBI Combinatory Dictionary
- Redes
- Diccionario médico de Masson
- etc.

Asimismo, generamos códigos para etiquetar la información sobre los perfiles de los entrevistados: años de experiencia, situación de empleo, lenguas de trabajo y especialidades.

---

<sup>136</sup> Código utilizado para etiquetar los comentarios de los traductores sobre algunas herramientas de búsqueda que muestran la terminología aislada, sin relacionarla con otros términos del campo.

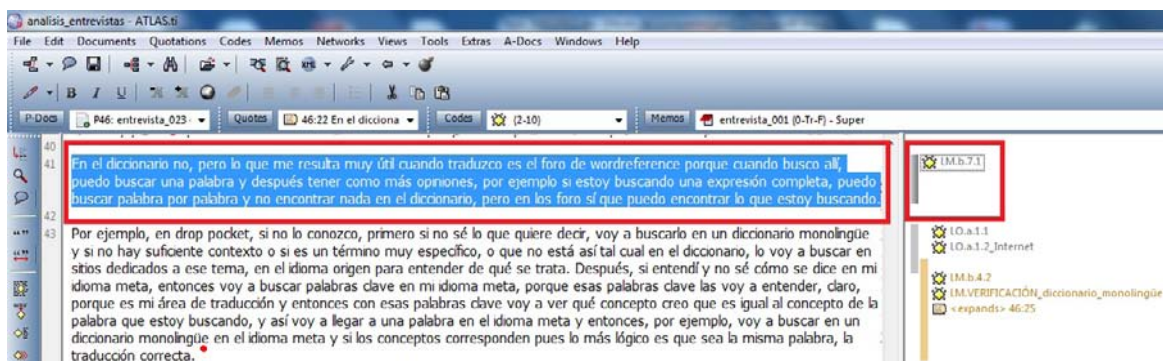
<sup>137</sup> Código utilizado para etiquetar los comentarios de los traductores sobre algunas herramientas de búsqueda que muestran la terminología sin proporcionar un contexto o ejemplo de uso.

<sup>138</sup> Código utilizado para etiquetar los comentarios de los traductores sobre algunas herramientas de búsqueda que no actualizan la terminología que contienen.

- 10 (años de experiencia)
- autónomo (situación de empleo)
- directa: inglés-español (lenguas de trabajo)
- traducción\_jurídica (especialidades)
- etc.

Por tanto, nuestro análisis adoptó un enfoque mixto, ya que partíamos de unos códigos previos, que habíamos obtenido de los análisis realizados anteriormente, pero estábamos abiertos a la aparición de nuevas ideas y categorías, ya que nuestro campo de estudio, las búsquedas onomasiológicas en traducción, se encontraban poco exploradas, por lo que no contábamos con demasiadas ideas preconcebidas.

En la Ilustración 102 mostramos un fragmento de nuestro análisis en el que asignamos códigos, que aparecen en la mitad derecha de la pantalla, a fragmentos o citas de una entrevista, en la mitad izquierda de la pantalla. Por ejemplo, siempre que los traductores mencionaban la idea de consulta en un foro, esta información la etiquetábamos con el código LM.b.7.1.



**Ilustración 102.** Asignación en el análisis con *Atlas.ti* del código LM.b.7.1 a la idea de consulta en un foro.

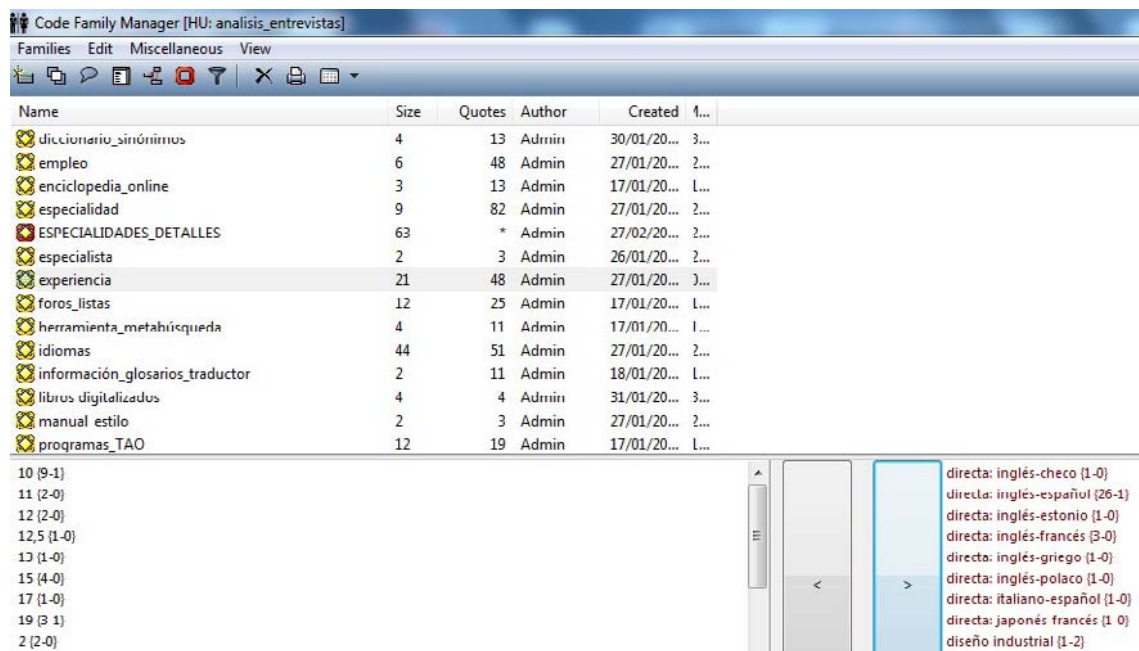
En la **fase conceptual**, agrupamos los códigos en *familias* o *families* y en *redes conceptuales* o *networks*. Estas redes nos ayudaron a estructurar las ideas del análisis. También utilizamos *notas o memos*, para registrar algunas conclusiones a las que llegábamos a medida que avanzaba el análisis.

Por un lado, las familias las utilizamos para dos tipos principales de agrupaciones. Primero, se organizaron en familias los nombres de recursos mencionados por los entrevistados. Por ejemplo todos los códigos referentes a los nombres de diccionarios de colocaciones (diccionario *BBI Combinatory Dictionary* y diccionario *Redes*) se incluyeron en la familia *diccionario\_colocaciones*.



Segundo, se utilizaron familias para agrupar los códigos de los perfiles de los entrevistados. Por ejemplo, todas las categorías referentes a las combinaciones de idiomas de los traductores se ubicaron en la familia *idiomas*, los códigos que determinaban el número de años de experiencia de los entrevistados se incluyeron en la familia *experiencia*, como observamos en la Ilustración 103. Aquí aparece seleccionada la familia *experiencia* con los códigos correspondientes a los años de experiencia de los traductores: *10*, *11*, *12*, etc.

A partir de estas familias, *Atlas.ti* nos permitió exportar datos cuantitativos sobre los perfiles de los entrevistados a tablas de Excel. Por ejemplo, en cuántas entrevistas se habían etiquetado los códigos *10*, *11*, *12*, etc., correspondientes a los años de experiencia de los entrevistados. Con estas tablas obtuvimos datos estadísticos sobre el perfil de los entrevistados (véase el apartado 5.2.3.1 y el anexo 5). Este es el único análisis cuantitativo que se hizo en el estudio.



**Ilustración 103. Agrupaciones en familias de algunos códigos de nuestro análisis en *Atlas.ti*.**

Por otro lado, en las redes conceptuales agrupamos y establecimos relaciones entre los códigos. Esta fue la parte más importante del análisis cualitativo, ya que las redes se convirtieron en mapas mentales en los que estructurábamos la información que obteníamos de las reflexiones de los traductores sobre las estrategias de búsqueda y los recursos.

Para indicar las relaciones en las redes entre los distintos códigos utilizamos las relaciones predeterminadas de *Atlas.ti* (*is associated with*, *is part of*, *is cause of*, *contradicts*, *is a*, *no name*, *is property of*). Solo creamos dos relaciones adicionales que

necesitábamos para reflejar las relaciones entre algunas estrategias de búsqueda. La relación *is alternative*, que establecimos entre dos o más estrategias de búsqueda para indicar que el traductor utiliza una u otra estrategia, pero no las dos. La relación *followed by* indicaba que una estrategia sucede a otra en el proceso de búsqueda de un traductor.

En total, creamos dos grandes redes conceptuales, resultado de todo el análisis. En una de las redes agrupamos y relacionamos las estrategias de búsqueda de los traductores. En la otra red organizamos la información aportada por los traductores respecto a la utilidad de los recursos existentes (ventajas, desventajas y demandas).

La primera red conceptual sobre estrategias de búsqueda la organizamos partiendo de la clasificación preliminar de estrategias de búsqueda a partir de la cual habíamos generado códigos iniciales (LO.a, LM.b, LM.a, LM.c, etc.). A partir de aquí, estructuramos las diferentes estrategias detectadas en el análisis y sus relaciones (entre otras, qué estrategias precedían o sucedían a otras). Por ejemplo, a una búsqueda en el diccionario bilingüe (LM.b.1.1) suele sucederle una verificación en Internet del equivalente encontrado en contexto (LM.a.VERIFICACIÓN\_Internet).

También relacionamos qué estrategias de las mencionadas por los traductores podían tener una naturaleza onomasiológica, como por ejemplo las estrategias de tipo LM.b.4, en las que el traductor realiza una búsqueda de palabras clave relacionadas con la denominación que desconoce y localiza en contexto la denominación desconocida porque aparece en una posición cercana a las palabras de búsqueda; las estrategias LM.b.7 de consulta en foros a partir de la información conocida sobre el significado para que los participantes del foro proporcionen la denominación adecuada; también las estrategias LM.b.8 de consulta directa a personas, por ejemplo a especialistas.

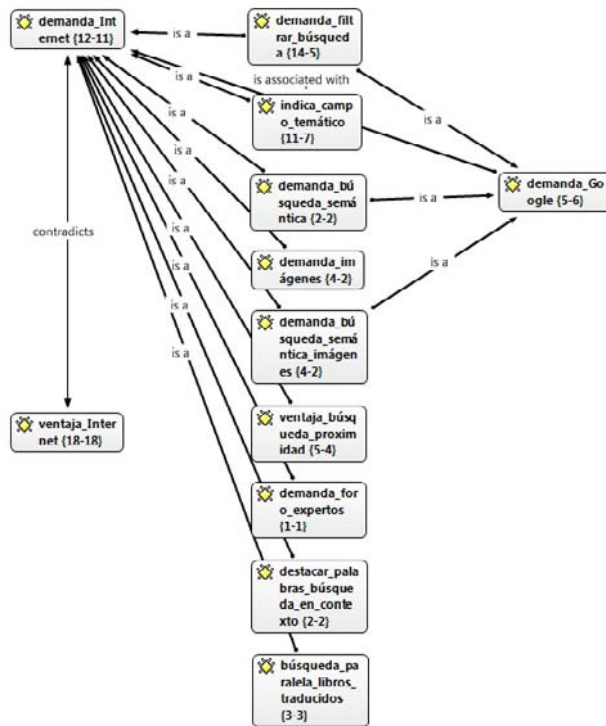
En la segunda red conceptual sobre los recursos indicamos, en primer lugar, los tipos de recursos mencionados por los traductores: diccionarios (diccionarios en papel y electrónicos, diccionarios generales y especializados, diccionarios monolingües, bilingües y multilingües, diccionarios de acrónimos, de sinónimos, de colocaciones, etc.), glosarios propios, herramientas de traducción asistida, corpus, recursos de Internet (buscadores como *Google*, *Google Libros* y *Google Imágenes*, la enciclopedia *Wikipedia*, foros, etc.), herramientas de metabúsqueda (que buscan simultáneamente en recursos de distinto tipo).

Además, en la red añadimos la información codificada sobre la frecuencia o preferencia de uso por parte de los traductores de cada recurso, por ejemplo vimos que el uso del diccionario bilingüe parece que disminuye con la adquisición de experiencia en traducción y que también ha disminuido desde la aparición de Internet. Después, añadimos

información a la red conceptual sobre el descontento o las desventajas que observan los traductores en cada uno de los recursos mencionados; también las ventajas que les aporta cada uno y finalmente las demandas.

La navegación dentro de los dos grandes grafos generados nos permitía hacer clic en cada uno de los nodos del grafo y observar las relaciones con el resto de elementos del grafo, por ejemplo, podíamos seleccionar el nodo sobre demandas de los traductores respecto de Internet y observar toda la información que habíamos etiquetado en las entrevistas sobre las demandas de los traductores en este recurso.

A continuación, en la Ilustración 104 mostramos una parte de la red sobre las ventajas, desventajas y demandas en los recursos. En la imagen se reflejan las **relaciones del nodo demanda\_Internet**. El grafo muestra que los traductores expresaron las siguientes demandas en Internet: la demanda de que puedan **filtrarse las búsquedas**, que a su vez está asociada con que se pueda filtrar por campo temático, y también se indica que esta es una demanda que realizan los traductores para la herramienta de búsqueda *Google*; la demanda de **búsquedas semánticas** en general y **búsquedas semánticas de imágenes**, que también los traductores piden que se incorpore en *Google*; la demanda de que puedan realizarse **búsquedas por proximidad**; que se creen **foros de expertos**; que en los resultados se **destaquen las palabras de búsqueda en contexto**; y que se facilite una **búsqueda paralela de libros traducidos** como en un gran corpus paralelo. Finalmente, el mapa conceptual indica que aunque los traductores tienen demandas en los recursos de Internet, también observan ventajas. En este caso, podríamos abrir la red conceptual de ventajas de Internet para observar las ventajas de este recurso mencionadas por los entrevistados.



**Ilustración 104.** Fragmento de la red conceptual sobre las ventajas, desventajas y demandas en los recursos (nodo sobre las demandas de los traductores en Internet).

Sirva de ejemplo esta pequeña muestra que forma parte de la red más amplia en la que relacionamos todas las ventajas, desventajas y demandas señaladas por los traductores respecto de los recursos. Debido a su amplitud, no mostraremos al completo los dos grandes grafos a modo de imagen, ya que la navegación a través de los nodos del grafo solo es posible dentro del programa *Atlas.ti*.

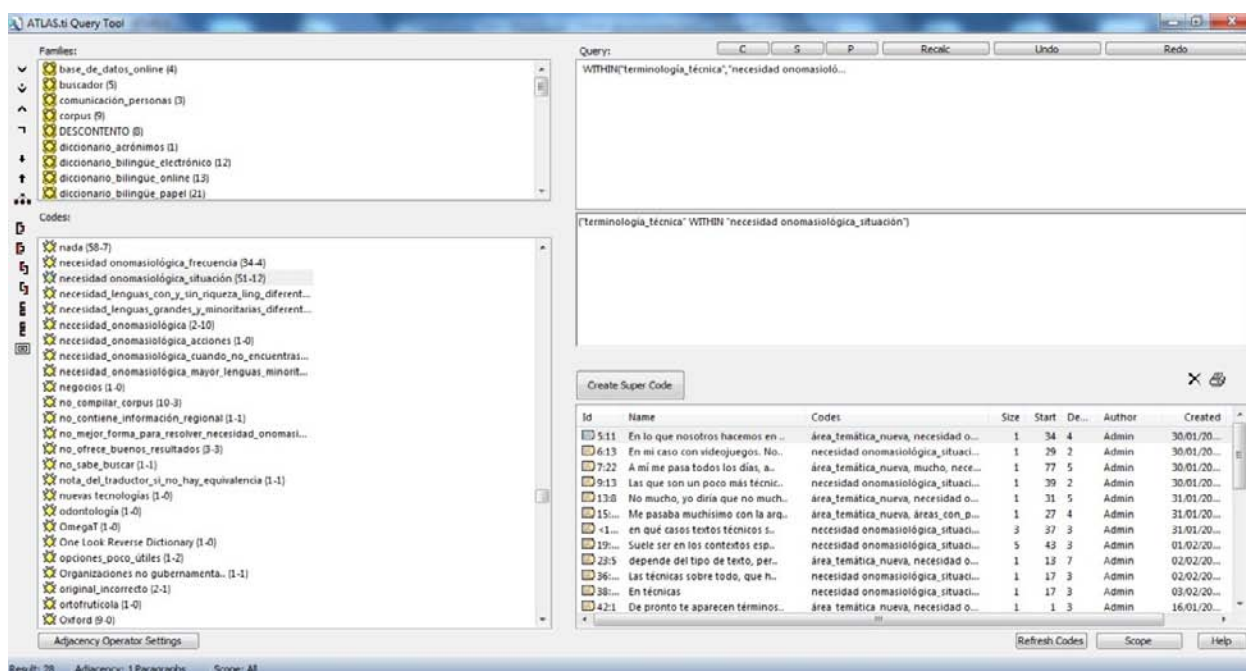
Llegados a este punto, cabe decir que, en nuestro análisis, las fases textual y conceptual fueron cogidas de la mano, ya que a medida que leíamos las entrevistas y asignábamos nuevos códigos, los agrupábamos en familias o los insertábamos y vinculábamos en las redes conceptuales. A su vez, al tiempo que desarrollábamos las redes conceptuales, era necesario modificar algunos códigos. Este proceso lo facilitó considerablemente el uso del programa *Atlas.ti*, ya que si eliminábamos o modificábamos el nombre de un código, este cambio se aplicaba a todos los códigos del análisis con dicho nombre.

Para la redacción de los resultados del análisis, tomamos como punto de partida las dos grandes redes conceptuales creadas, en las que habíamos estructurado las ideas obtenidas de las respuestas de los traductores sobre las estrategias de búsqueda y recursos utilizados. Así, navegábamos en las redes conceptuales, identificábamos y ubicábamos cada una de las ideas y las incluíamos en los resultados.

Por ejemplo, uno de los objetivos de nuestro estudio era conocer y ordenar las estrategias de búsqueda de los traductores y detectar aquellas estrategias que permiten resolver necesidades de tipo onomasiológico. Por tanto, para estructurar estos resultados partimos de la red conceptual sobre estrategias de búsqueda en la que habíamos organizado estas ideas en forma de códigos, y seleccionamos una a una las ideas y las ubicamos en los resultados siguiendo el orden y las relaciones entre las ideas que habíamos establecido en la red conceptual.

Después, para ilustrar cada una de las ideas de los resultados, necesitábamos acceder a aquellas citas que se encontraban etiquetadas con uno o varios códigos. En este sentido, al haber segmentado las entrevistas en citas y haber asignado códigos a las mismas pudimos encontrar con una sencilla búsqueda todas las citas que respaldaban cada una de las ideas. Realizábamos la búsqueda del código correspondiente a una idea a través de la herramienta de búsqueda de *Atlas.ti* y accedíamos a las citas que habíamos etiquetado en las entrevistas. A continuación, extraíamos aquellas citas que nos resultaban más representativas de cada una de las ideas expuestas y las incluíamos en los resultados.

Por ejemplo, si queríamos recuperar las citas en las que se había indicado que la necesidad de realizar búsquedas onomasiológicas puede motivarla la presencia de terminología técnica en las traducciones, realizábamos una búsqueda que recuperaba las citas etiquetadas con los códigos *necesidad\_onomasiológica\_situación* y *terminología\_técnica* simultáneamente (“terminología\_técnica” WITHIN “necesidad\_onomasiológica\_situación”). Véase la Ilustración 105.



**Ilustración 105.** Búsqueda de códigos y citas de nuestro análisis con la herramienta de búsqueda de *Atlas.ti*.

Finalmente, realizamos una exportación de todas las citas del análisis, agrupadas por códigos para incluirlas en el anexo 6. De esta forma, el lector podrá acceder, si así lo desea, a todas las respuestas de los traductores que respaldan cada una de las ideas que presentamos en los resultados, además de aquellas que nosotros hemos seleccionado como las más relevantes y que incluimos a continuación. También creamos el anexo 7, un archivo comprimido con las transcripciones completas de cada una de las entrevistas.

### 5.2.3 Resultados

Los **resultados** de nuestro estudio corresponden al **análisis de 48 entrevistas a traductores profesionales**.

Hemos dividido los resultados en cuatro partes. Primero, en el apartado 5.2.3.1 comenzamos con los resultados de un pequeño estudio estadístico en el que reflejamos las características de los traductores profesionales entrevistados: experiencia, situación laboral, lenguas de trabajo y especialidades.

Posteriormente, incluimos los resultados de nuestro análisis cualitativo sobre estrategias de búsqueda onomasiológica en traducción. En el apartado 5.2.3.2, recogemos resultados generales sobre las necesidades onomasiológicas en traducción, respondemos a la pregunta de si los traductores tienen necesidades onomasiológicas en su actividad de

traducción, con qué frecuencia, en qué situaciones y qué otros factores provocan la aparición de necesidades onomasiológicas.

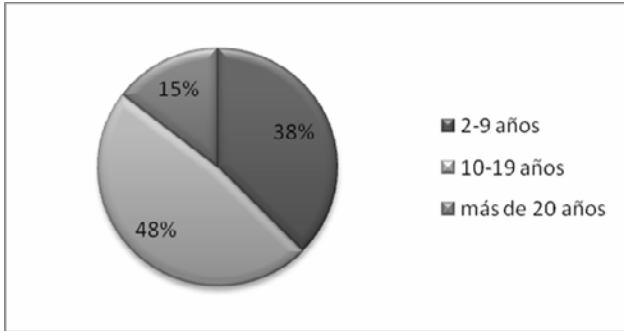
En el apartado 5.2.3.3, recopilamos y estructuramos todas las estrategias de búsqueda aportadas por los traductores y ubicamos, dentro de ellas, a las estrategias de búsqueda onomasiológica. Finalmente, en el apartado 5.2.3.4, explicamos las impresiones de los traductores como usuarios respecto de los recursos que utilizan, ventajas, desventajas y demandas.

Por cada resultado del análisis cualitativo, exponemos la idea que hemos obtenido del análisis de las entrevistas y la justificamos con las citas literales de los traductores extraídas de las transcripciones de las entrevistas. En cada idea solo incluimos las citas que consideramos más representativas o ilustrativas. Además, al lado de cada cita identificamos entre paréntesis el código de entrevistado, así como los códigos utilizados en el análisis para etiquetar las afirmaciones de los traductores. De este modo, el lector puede acceder, si desea ampliar información, a todas las citas asociadas a una determinada idea. En el anexo 6 se encuentran todas las citas del análisis agrupadas por códigos.

#### 5.2.3.1 Resultados sobre las características de los entrevistados

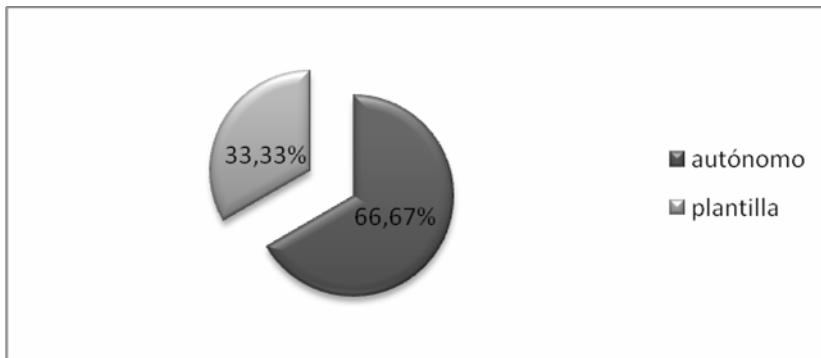
En el primer bloque de preguntas de la entrevista recabamos información sobre el **perfil de los entrevistados**. A continuación, presentamos para interés del lector información sobre los 48 traductores profesionales: años de experiencia, situaciones de empleo, combinaciones de lenguas y especialidades que traducen. Los resultados del perfil se obtuvieron de un pequeño estudio estadístico que realizamos sobre la muestra de los 48 entrevistados. Para más información, véase el anexo 5.

Todos los entrevistados contaban con más de **2 años de experiencia** dedicándose exclusivamente a la traducción profesional, 30 de los traductores (63%) tenían más de 10 años de experiencia, de los cuales 7 (15%) superaban los 20 años. La media de años de experiencia de los traductores entrevistados es 11,7 (número de años suficiente para considerar que un traductor es experto, según Jensen y Jakobsen 2000: 114). En la Ilustración 106 puede observarse la distribución de la muestra por cada uno de los años de experiencia que mencionaron los entrevistados.

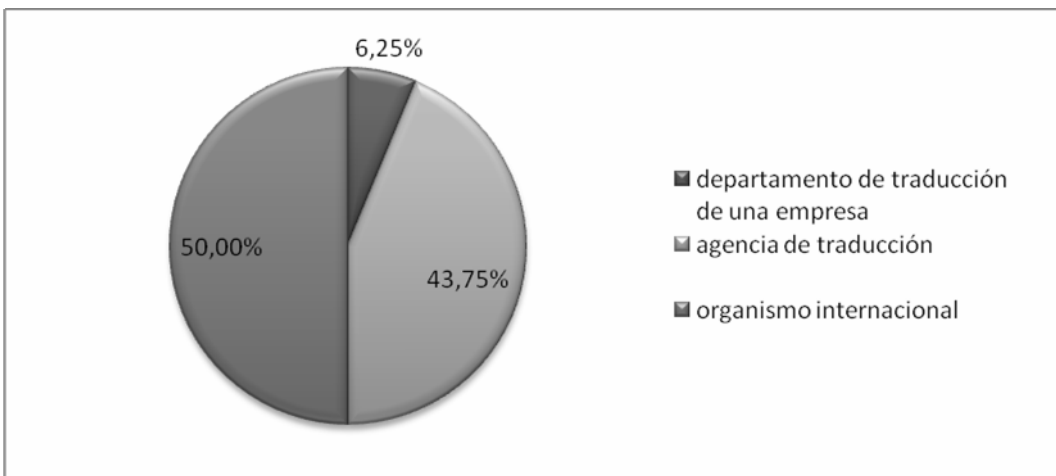


**Ilustración 106. Distribución de la muestra según los años de experiencia.**

En cuanto al modo en el que los entrevistados ejercen la actividad de traducción, su **situación de empleo**, 32 (66,67%) indicaron que trabajan por cuenta propia, como autónomos y 16 (33,33%) por cuenta ajena, en plantilla, véase la Ilustración 107. De los que trabajaban en plantilla, 1 (6,25%) de ellos lo hacía en el departamento de traducción de una empresa internacional, 7 (43,75%) en una agencia de traducción y 8 (50%) en un organismo internacional (véase la Ilustración 108).



**Ilustración 107. Distribución de la muestra según la situación de empleo.**



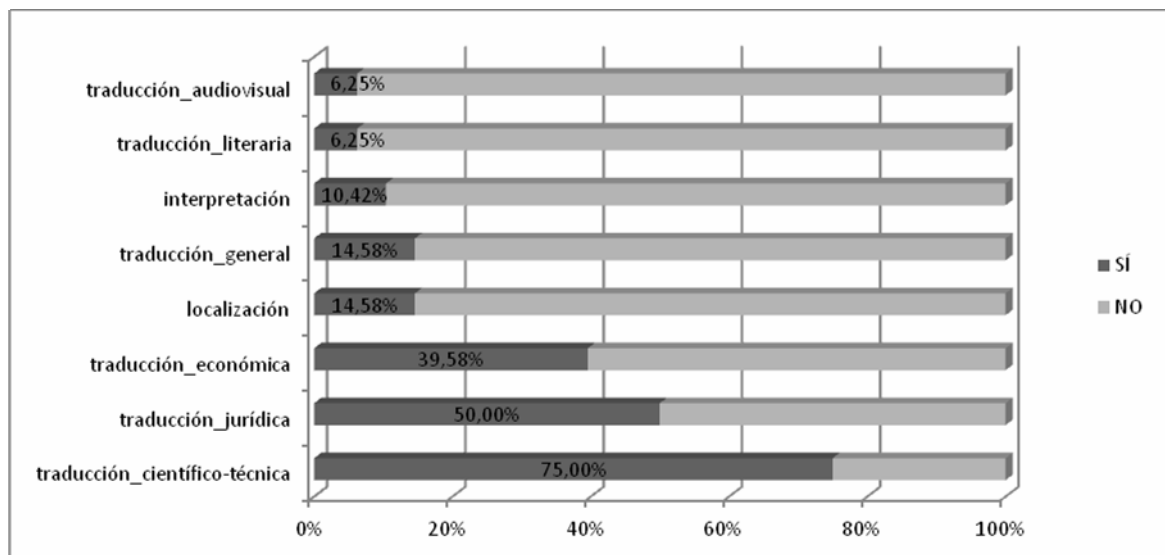
**Ilustración 108. Distribución de la muestra según la situación de los traductores que trabajan en plantilla.**



En cuanto a las **especialidades de traducción**, la mayoría de los entrevistados mencionó la necesidad de especializarse en más de un sector debido a las exigencias del mercado y al descenso del número de traducciones en los últimos años. En este sentido, casi todos los entrevistados señalaron más de una especialidad. Especialmente los autónomos afirmaron que traducen *de todo*, en los traductores en plantilla la especialización es ligeramente más elevada, aunque la mayoría también traduce varios campos de especialización.

Debido a que cada entrevistado indicó más de una especialidad, tratamos cada una de las áreas como variables independientes y representamos cada una en la gráfica sobre el 100% (véase la Ilustración 109) donde la parte de la barra en gris oscuro indica el porcentaje de nuestros 48 traductores que sí que traduce esta especialidad y la parte de la barra en gris claro el porcentaje que no la traduce.

La especialidad más recurrente es la traducción científico-técnica (que traducen 36 de los entrevistados, el 75%), le sigue la traducción jurídica (que realizan 24 de nuestros traductores, el 50%), después 19 traductores se dedican a la traducción económica (39,58%), 7 a la localización (14,58%), 7 a la traducción general (14,58%), 5 a la interpretación (10,42%), 3 a la traducción literaria (6,25%) y 3 a la traducción audiovisual (6,25%).



**Ilustración 109.** Distribución de la muestra según las especializaciones.

Finalmente, los traductores especificaron la **temática de los textos** que traducen. Como cada traductor mencionó la traducción de textos de muchas temáticas diferentes, también tratamos cada temática de forma independiente. A continuación incluimos los

datos más representativos sobre la temática de los textos traducidos. No obstante, todas las temáticas con sus porcentajes pueden consultarse en el anexo 5.

Las temáticas que más se mencionaron son informática (14 traductores, 29,17%), medicina (10 traductores, 20,83%), *marketing* (10 traductores, 20,83%), manuales de instrucciones (7 traductores, 14,58%), textos de la UE (7 traductores, 14,58%), automoción (6 traductores, 12,50%), ingeniería (5 traductores, 10,42%), entre otras.

En cuanto a las **lenguas de trabajo**, los traductores indicaron que trabajan más de una combinación de lenguas. Por tanto, cada una de las combinaciones de lenguas las tratamos también como variables independientes y las calculamos sobre nuestra muestra de 48 entrevistados.

De los resultados obtenidos, observamos una gran variedad de lenguas de trabajo y además los porcentajes de la mayoría de las lenguas no eran muy elevados, por eso decidimos que la representación gráfica de cada una de las combinaciones de lenguas no resultaría demasiado representativa. Por eso, realizamos agrupaciones de las lenguas para representarlas en un gráfico (véase Ilustración 110). Aquí vemos que para el 100% de los traductores su combinación de lenguas principal es de traducción directa; también mostramos que el 29% de los traductores mencionaron traducir una o más combinaciones de lenguas secundarias, finalmente observamos que solo el 6% indicó que se dedica a la traducción inversa.

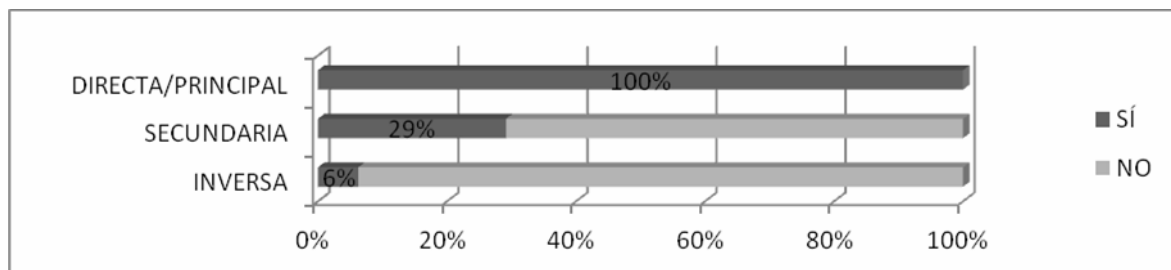
A continuación ofrecemos también los datos más representativos sobre las combinaciones específicas de lenguas mencionadas por los traductores. Los datos completos sobre todas las combinaciones de lenguas y sus porcentajes pueden consultarse en el anexo 5.

Respecto a las combinaciones de lenguas principales de los entrevistados (aquellas en las que realizan traducciones en mayor volumen o casi en exclusividad), la combinación mayoritaria es directa inglés-español, traducida por 26 (54%) de los entrevistados, 7 (15%) traducen directa español-inglés, 6 (13%) directa alemán-español, 4 (8%) directa francés-español, 4 (8%) directa español-francés y 3 (6%) directa español-alemán; también se mencionan otras combinaciones de estas lenguas con el catalán, chino, italiano, búlgaro, checo, estonio, griego, polaco, japonés, etc.

Estas combinaciones principales de lenguas son de traducción directa L2-L1, de una lengua extranjera a la lengua materna, solo 3 (6%) de los entrevistados indicaron que también realizan traducción inversa pero en servicios de interpretación y, de estos, 1 también en traducción aunque en muy poca cantidad. En cuanto a la combinación de

idiomas en traducción inversa, 1 entrevistado afirmó que traduce español-alemán y 2 español-inglés, pero en ninguno de estos casos como combinación principal.

Asimismo, algunos traductores señalaron que traducen a otras combinaciones de idiomas de forma secundaria, es decir, que tienen menos volumen de traducción de estas combinaciones. Estas combinaciones secundarias también son de traducción directa L2-L1, 5 (10%) alemán-español, 4 (8%) francés-español, 2 (4%) portugués-español y 2 (4%) italiano-español (estas combinaciones de idiomas secundarias las mencionaron traductores que realizan como combinación principal inglés-español). También se indicaron otras combinaciones de lenguas secundarias con sueco, noruego, rumano, etc.



**Ilustración 110.** Distribución de la muestra según agrupaciones de las lenguas de trabajo de los traductores.

En definitiva, nuestros entrevistados son traductores profesionales expertos (11,7 años de experiencia de media), que trabajan más de una combinación de lenguas (la mayoría traducción directa inglés-español, también directa español-inglés y alemán-español) y más de un área temática (la mayoría traducción científico-técnica, también jurídica y económica).

Nuestra muestra quizás tiene un sesgo hacia la traducción inglés-español y la especialidad científico-técnica. Sobre esto cabría explicar que sí que procuramos contar con traductores que se dedicaran a la traducción especializada porque en los capítulos anteriores vimos que en este tipo de traducción los traductores tienen una falta de terminología que desconocen y pensamos que esta podría generar necesidades de tipo onomasiológico. Sin embargo, incluimos otros perfiles de idiomas y especialidades, como por ejemplo jurídica, económica, literaria, etc. porque pretendíamos vislumbrar posibles tendencias generales en las búsquedas onomasiológicas de los traductores, un área de la que apenas contábamos con información previa.

### 5.2.3.2 Resultados sobre las necesidades en traducción. Situaciones y frecuencia de aparición de las necesidades onomasiológicas

Las primeras preguntas del segundo bloque de la entrevista estaban dirigidas a obtener unos datos iniciales sobre las necesidades del traductor y, concretamente, sobre necesidades de tipo onomasiológico, es decir, si el entrevistado había tenido necesidades de tipo onomasiológico, en qué situaciones o con qué tipo de traducciones y con qué frecuencia.

De las respuestas obtenidas, vemos que las **necesidades onomasiológicas** sí que están presentes en la actividad de traducción, por ejemplo en las siguientes **situaciones**.

Un problema que genera una necesidad de búsqueda onomasiológica a los traductores es la existencia de terminología poco habitual o desconocida en las traducciones, es decir, en ocasiones el traductor comprende el significado o concepto de la terminología utilizada en la LO, pero desconoce una denominación adecuada para expresar dicho significado en la LM.

Esto sucede sobre todo a traductores que trabajan con terminología técnica, jurídica y médica (con los nombres de piezas técnicas y los síntomas de enfermedades), también con los refranes. Por ejemplo, en el caso de piezas o máquinas, los traductores indicaron que en ocasiones saben para qué sirve una pieza o las funciones de una máquina, pero ignoran cuál es la denominación adecuada para designarla en el texto meta. Esta situación, señalan, suele aparecer con clientes nuevos o con áreas temáticas nuevas para el traductor.

Eso suele pasar con traducciones muy literarias, de muy alto nivel, o muy técnicas, generalmente, porque claro todo lo demás o lo conoces tú o sí que está en los diccionarios, pero puede haber cosas muy muy técnicas donde tengas, o muy literarias de una exposición, o que utilicen una terminología muy específica y que utilicen términos que... ten en cuenta que un traductor profesional que tenga un poco de experiencia, vocabulario tiene mucho, entonces para que sea una cosa de ese tipo, tiene que ser algo especializadísimo. En medicina, desde luego, en medicina muchas veces te puede ocurrir, y a veces en geotérmica, energía nuclear, nanotecnología, de cosas muy elevadas, cosas de ese tipo. (entrevista\_026: necesidad\_onomasiológica\_situación; terminología\_poco\_habitual; terminología\_técnica; terminología\_médica)

No, o sea pues si es un término muy específico: la camisa de no sé qué pieza de... que la encuentras dentro de la traducción en el contexto, te da la definición o para qué sirve, tú sabes exactamente qué es pero evidentemente es una terminología muy específica y no sabes cómo se dice en castellano. (entrevista\_038: necesidad\_onomasiológica\_situación; terminología\_técnica)

También parece que los traductores que residen durante un tiempo prolongado fuera de su país de origen tienen necesidades de tipo onomasiológico (de que entienden un concepto pero les falta la denominación). Esto se debe a que estos traductores tienen menos contacto con su lengua materna que es a la que traducen.

Me ha pasado algunas veces, es que ya llevo mucho tiempo aquí en España y hay conceptos, palabras o expresiones que yo he conocido aquí en España, entonces me pasa incluso que yo entiendo la idea que tengo que transmitir, pero me falta el equivalente en alemán porque, creo que me pasó con la gripe aviaria, que surgió aquí, yo en alemán no lo conocía, tenía que buscar el equivalente pero, igual no es un buen ejemplo porque ahí sí que es más fácil encontrar la traducción, pero sí que hay otros ejemplos de conceptos que he conocido aquí, o viceversa, en Alemania y aunque comprendas la idea, es muy difícil encontrar el equivalente. (entrevista\_032: necesidad\_onomasiológica\_situación; traductores\_fuera\_de\_su\_país)

Asimismo, esta necesidad la advierten traductores que trabajan con lenguas minoritarias, por la escasez de recursos de búsqueda en estas lenguas. En este sentido, acceder a la denominación adecuada que designa a un concepto resulta una tarea más compleja.

Desde luego y sobre todo creo que es el caso para los idiomas minoritarios, que quizás no hay tanto traducido, entonces es más difícil también encontrar en recursos bilingües y tienen que buscarse un poco la vida, digamos acabando por utilizar ya sea perfrasis, que realmente no encuentran el término, o incluso creando neologismos si es preciso y haciendo una labor de búsqueda muchísimo más avanzada de la que quizás hacen otros idiomas porque tienen muchísimos recursos bilingües disponibles en los que casi automáticamente vas a encontrar el equivalente adecuado sin tener que hacer búsquedas más allá. (entrevista\_024: necesidad\_onomasiológica\_mayor\_lenguas\_minoritarias)

En relación con la escasez de recursos, las necesidades onomasiológicas surgen en áreas en las que existen pocos diccionarios especializados.

Me pasaba muchísimo con la arquitectura, últimamente he tenido alguna cosita u otra pero no gran cosa y precisamente ayer estaba buscando un término que sin ser neologismo no viene en los diccionarios bilingües, puede que venga en algún glosario especializado pero lo que he hecho es buscar textos paralelos. Me ocurre también en el tema hortofrutícola porque están siempre inventando nuevos productos, te lo creas o no, nuevos métodos de cultivo, nuevos tipos de empaquetado, nueva maquinaria para realizar las tareas, todo ese tipo de cosas, me ocurre mucho. En los otros campos menos, aunque también quizás porque en económicas hay muchos más diccionarios y glosarios especializados. (entrevista\_008: necesidad\_onomasiológica\_situación; áreas\_con\_pocos\_diccionarios\_especializados)

Asimismo, la necesidad onomasiológica surge porque los recursos existentes no proporcionan soluciones satisfactorias, por ejemplo de terminología bilingüe. Los diccionarios bilingües no siempre resuelven el problema de obtener un término en la LM equivalente a un término en la LO. En este caso, el traductor parte del concepto o significado de la LO para tratar de encontrar una denominación adecuada en LM.

No obstante, aunque el diccionario bilingüe aporte posibles soluciones de denominación en la LM, los traductores indicaron que no consideran al diccionario bilingüe como la mejor forma de resolver una necesidad de tipo onomasiológico y, por tanto, realizan búsquedas de otro tipo para asegurarse de la equivalencia de significados.

Sí, si no encuentras obviamente en el diccionario bilingüe tienes que, un poco partiendo de la información que puedas sacar del monolingüe y del concepto, intentar llegar al término en la otra lengua. (entrevista\_024: necesidad\_onomasiológica\_cuando\_no\_encuentras\_terminología\_bilingüe; aunque\_se\_encuentre\_en\_bilingüe)

Normalmente el punto de partida es el bilingüe. Normalmente si en el bilingüe no está muy claro, o te viene una definición pero no es para el campo, imagínate estás haciendo una traducción de maquinaria y la acepción que viene es para centrales nucleares, buscas normalmente una segunda fuente. (entrevista\_034: necesidad\_onomasiológica\_cuando\_no\_encuentras\_terminología\_bilingüe; aunque\_se\_encuentre\_en\_bilingüe)

En cuanto a la **frecuencia** en la que los traductores se enfrentan a necesidades onomasiológicas, las respuestas fueron variadas, aunque todos coincidieron en que es un problema que cualquier traductor se encuentra en mayor o menor medida. Algunos entrevistados indicaron que les sucede con mucha o bastante frecuencia, sobre todo en las situaciones mencionadas arriba.

Sí, yo creo que eso está a la orden del día de cualquier traductor, de comprender qué es lo que el texto de partida te está diciendo y a lo mejor no saber, no estar seguro del término más indicado para expresar la misma idea, el mismo concepto en la lengua de llegada. En ese sentido, sí, muchísimas veces. (entrevista\_024: necesidad\_onomasiológica\_frecuencia)

Los traductores a los que les sucede con menos frecuencia suelen traducir textos más bien generales, o donde la terminología no es un problema; y también traductores muy especializados que siempre traducen textos de la misma temática.

Not very frequently because of the nature of the work we are doing here, it does not require you to know a lot about very different subjects, basically it is pretty much focused on what the Council has been doing in its proceedings, and as I said, most of the work has been done by the Commission. (entrevista\_046: necesidad\_onomasiológica\_frecuencia)

No mucho porque ya te digo que son temas en los que llevo años trabajando y está toda la terminología buscada, pero a ver a lo mejor un par de veces al mes tengo algún término rebelde que llamo yo. (entrevista\_042: necesidad\_onomasiológica\_frecuencia)

### 5.2.3.3 Resultados sobre las estrategias de búsqueda de los traductores y las estrategias de búsqueda onomasiológica

De las respuestas de los traductores a las preguntas sobre qué estrategias de búsqueda emplean en los recursos para resolver los problemas que les plantea una traducción, recopilamos y agrupamos una serie de **problemas y estrategias de búsqueda** reseñadas por los traductores. Además, concretamente, pedimos a los traductores que nos especificaran estrategias de búsqueda de tipo onomasiológico, por lo que pudimos ubicar estas estrategias dentro de la clasificación de estrategias de búsqueda.

Para la ordenación, utilizamos el mismo sistema de codificación (de letras y números) que en la clasificación preliminar de estrategias de búsqueda del capítulo 4<sup>139</sup>.

---

<sup>139</sup> Para más detalles sobre el modo de análisis y clasificación, véase el apartado 5.2.2.5.

Los **problemas** que indicaron los traductores solían darse por el siguiente orden cronológico: problemas de recepción en lengua origen (LO.a), es decir, en los que el traductor requiere comprender el significado de una denominación en la LO; después problemas de transferencia a la lengua meta (LM.b), en los que el traductor necesita una denominación en LM para un significado o concepto de la LO; a continuación problemas de recepción en LM (LM.a), donde el traductor precisa obtener el significado de un término de la LM sobre todo para realizar una verificación, comprobar que el significado en LO y en LM coinciden (LM.a.VERIFICACIÓN); finalmente problemas de producción en la LM (LM.c), de conocer la fraseología para expresarse en LM.

Para resolver cada uno de los problemas, los traductores utilizan una serie de **estrategias de búsqueda** que suelen darse en un orden determinado, o dando prioridad a unas estrategias antes que a otras.

Por ejemplo, ante problemas de recepción (LO.a) observamos que el traductor suele emplear primero recursos monolingües en LO (un diccionario monolingüe en LO, un corpus en LO, una búsqueda a través de un buscador de Internet o una enciclopedia en línea); si lo anterior no funciona accede a un buscador de imágenes, a algún diccionario específico, tesoro o diccionario de sinónimos; si no, pregunta a un nativo o experto en la materia y, finalmente, al cliente.

Para encontrar una denominación equivalente en LM a un concepto de la LO (problema LM.b), el traductor accede primero a recursos bilingües (un diccionario bilingüe, un corpus paralelo o una memoria de traducción, un buscador de Internet preferiblemente para acceder a páginas bilingües, una enciclopedia bilingüe, una herramienta de metabúsqueda).

Si la búsqueda bilingüe no da resultados, el traductor suele probar con una búsqueda en un recurso monolingüe en LM a partir de una intuición de equivalente en la LM.

Si la intuición falla, el traductor trata de localizar una denominación en LM a partir de relaciones que tiene dicha palabra con otras del contexto.

Si el traductor no es capaz de encontrar la denominación por sí mismo, la siguiente estrategia puede ser la consulta en un foro; después la consulta directa a otro traductor, también a un especialista, a terminólogos que trabajan en un proyecto, al gestor de proyectos o al cliente. Una vez el traductor accede a una denominación equivalente en LM, suele verificarla en un recurso monolingüe.

Además de denominaciones equivalentes para conceptos entre lenguas, el traductor deberá ser capaz de combinar de forma adecuada cada una de las unidades para producir

un discurso que suene natural en LM (problema LM.c). Para ello, al traductor puede resultarle útil acceder a colocaciones en LM a través de la consulta en un corpus o en Internet, también en el diccionario monolingüe y, últimamente, también en diccionarios específicos de colocaciones.

En los siguientes apartados ofrecemos detalles sobre cada una de estas estrategias de búsqueda utilizadas por los traductores. Cabe decir que la clasificación de estrategias de búsqueda que vamos a comentar corresponde a una serie de tendencias que hemos observado en la forma habitual de proceder de los traductores cuando realizan búsquedas y no necesariamente a patrones inalterables; dependiendo de las preferencias de los traductores y de la situación, estos pueden optar por unas estrategias u otras. No obstante, sí que observamos ciertas similitudes que se repetían en las estrategias de búsqueda que transmitían los traductores. Las estrategias las presentamos a continuación siguiendo el orden que determinamos por ser la forma más habitual de aparición en una secuencia de búsqueda, según vimos por los comentarios de los traductores. Para resolver cada uno de los problemas que se plantean, los traductores suelen poner en marcha primero unas estrategias y, si estas fallan, utilizan las siguientes.

#### 5.2.3.3.1 Estrategias de búsqueda para la resolución de problemas de recepción (LO.a)

Ante una necesidad de recepción en LO (LO.a), los traductores indicaron que adoptan diferentes estrategias de búsqueda como primera opción.

En el caso de que se trate de una duda del lenguaje general, suelen acudir a la consulta de una denominación en el diccionario monolingüe para obtener una definición (LO.a.1.1).

El diccionario monolingüe en lengua original lo utilizo para entender el concepto, efectivamente. (entrevista\_008: LO.a.1.1)

Además, los traductores señalaron que la estrategia de buscar el término original en el diccionario monolingüe para obtener una definición puede ser una primera aproximación para la puesta en marcha posteriormente de una estrategia de búsqueda de tipo onomasiológico con la que llegar a una denominación en LM que exprese dicho significado. En este caso se utiliza el diccionario monolingüe para obtener «pistas» sobre elementos del significado que ayuden al traductor a continuar la búsqueda.



Los monolingües yo eso sí que los uso mucho, buscando en la definición en el idioma origen de la palabra un indicio que me lleve a una búsqueda similar en mi idioma. Muchas veces al buscar la definición, ha sido pues eso, en monolingües, más que por buscar la definición por buscar indicios que me lleven a términos que me podrían llevar a definiciones parecidas. (entrevista\_002: LO.a.1.1; pistas)

Las necesidades de recepción en LO también se dan con acrónimos, siglas o abreviaturas. El significado del acrónimo se obtiene accediendo a la forma desarrollada de dicho acrónimo. Para ello, el traductor realiza la consulta del acrónimo en un diccionario de acrónimos, en el caso de que disponga de uno (LO.a.1.1\_diccionario\_acrónimos).

*Acronym Finder.* Para encontrar el significado de EER: Energy Efficiency Ratio. (entrevista\_021\_adicional: LO.a.1.1\_diccionario\_acrónimos)

Otra primera opción es la búsqueda del término original para acceder a un contexto (LO.a.1.2), en Internet (LO.a.1.2\_Internet), por ejemplo buscándolo directamente en *Google*, o utilizando el operador *define* de este buscador. En ocasiones la consulta en un buscador deriva a la entrada de una enciclopedia en línea, por ejemplo la *Wikipedia*, donde se explica el significado, o el traductor accede directamente a una enciclopedia en línea para buscar (LO.a.1.2\_enciclopedia\_online)<sup>140</sup>; los traductores también buscan el significado en la página web del cliente (LO.a.1.2\_sitio\_web\_cliente).

Lo que hago es buscarlo en *Google define*: lo que sea y leerme las definiciones en inglés, que al fin y al cabo es como un monolingüe. (entrevista\_011: LO.a.1.2\_Internet; filtro\_define)

Pues lo primero hay un gran invento que se llama *Wikipedia*, lo que yo hago normalmente es caso del castellano al alemán, el castellano no lo entiendo, bueno tengo una idea de lo que podría significar pero no lo encuentro en ningún diccionario, por ningún lado, yo lo primero que hago es teclearlo, con todo el morro que me caracteriza, en un buscador de Internet, y ahí ya veo un poquito el contexto, cómo se utiliza, veo si es España, si es Sudamérica, me miro a ver si las fuentes son más o menos fiables o algún chalado se ha puesto ahí a publicar algo y por ejemplo si luego te sale el término en *Wikipedia*, te lees la explicación, más o menos el contexto y cruzas los dedos a ver si el mismo artículo está también en la versión alemana, que está en la parte a la izquierda y entonces ya está el asunto bastante resuelto. Aún así ahí ponen, ojo aún falta revisar el texto, no sé qué no sé cuántos, pero para la primera orientación va muy bien, es una gran ayuda. (entrevista\_010: LO.a.1.2\_Internet; LO.a.1.2\_enciclopedia\_online; LM.b.1.5.1)

Normalmente intento buscar siempre la página de la autoridad, la página de la organización o del eso que se llama así. (entrevista\_013: LO.a.1.2\_sitio\_web\_cliente)

En el caso de los acrónimos, si el traductor no tiene un diccionario de acrónimos o no ha encontrado el acrónimo en el diccionario de acrónimos, realiza la búsqueda en Internet en contexto (LO.a.1.2\_Internet\_acrónimos).

---

<sup>140</sup> A esta última estrategia puede seguirle la estrategia LM.b.1.5.1, que trataremos más adelante, donde el traductor puede acceder en la enciclopedia al artículo equivalente en otra lengua.

Un ejemplo serían las siglas médicas que hay. Los médicos usan muchísimas abreviaturas y todo tipo de siglas y tal y, pues nada, hay diccionarios de siglas médicas donde puedes encontrar esto y así cada temática tiene sus listados especializados. Ahí la posibilidad de encontrar algo es más alta, pero en todo caso para solucionar casos extremos, lo que se tiene que hacer es, como siempre, adentrarse en el tema y buscar un contexto donde sale el acrónimo para luego poder ir más profundizando para a lo mejor averiguarlo. (entrevista\_009: LO.a.1.1\_diccionario\_acrónimos; LO.a.1.2\_Internet\_acrónimos)

Asimismo, en determinados casos, para conocer el significado de determinados nombres, por ejemplo de máquinas o de piezas de una máquina, más que una definición escrita, al traductor le resulta útil acceder a una imagen (LO.a.1.3), por ejemplo en Internet, sobre todo en *Google* imágenes (LO.a.1.3\_buscador\_imágenes). Asimismo, la visualización de videos puede ayudar al traductor a comprender determinados procesos. También, aunque en menor medida, algunos traductores cuentan con un diccionario visual (LO.a.1.3\_diccionario\_visual).

El acceso a imágenes a partir de la búsqueda del término original también es un primer paso de la búsqueda onomasiológica para los traductores, combinada o seguida de otras estrategias (véase más adelante LM.b.4.3 y LM.a.VERIFICACIÓN\_buscador\_imágenes).

En la moda se crean muchos términos nuevos todos los días. Entonces muchas veces son cosas que no se han llevado antes y que tienes que imaginar, que tienes que crear. Pero dentro de lo que son tejidos, estampados, telas, cremalleras, remates... todas esas cosas, una imagen te ayuda mucho porque igual no sabes exactamente qué es y cuando lo ves dices «ah, sí, esto es esto y se dice así». Sobre todo si tú no has vivido en ese ámbito en inglés. Tú vives en español y lo conoces en español. Una vez ves la imagen es mucho más clara que si en un texto te explican qué es. Y si te ponen «bordado que se cose con botones a la derecha y no sé qué» tú eso a lo mejor no te lo imaginas, pero si ves la foto dices «ah, sí, esto es esto» y muchas veces de ahí ya sacas como se dice. A mí las imágenes me ayudan bastante, en las piezas y en la moda mucho. (entrevista\_013: LO.a.1.3\_buscador\_imágenes; pistas)

Pues muchas veces con imágenes y en este caso recorro a Internet, a *Google images* por ejemplo, o también a diccionarios que funcionan con imágenes, diccionarios visuales para buscar lo que hay detrás de la idea. (entrevista\_032: LO.a.1.3\_diccionario\_visual)

Con menor frecuencia, y también con posterioridad a las estrategias de búsqueda anteriores, algunos traductores recurren a un diccionario ideológico, un tesoro, una red de ideas, u otros recursos de este tipo para conocer el significado de una palabra viendo cómo se estructura dentro de una familia de palabras o dentro de un campo (LO.a.1.4). Aunque, como veremos más adelante, esta es una estrategia de búsqueda poco común entre los traductores, ya que alegaron que la búsqueda en estos recursos es difícil y contienen terminología demasiado general.

Tesoro también. Pues para ayudar en la definición conceptual o en una familia de palabras relacionadas conceptualmente sobre todo. (entrevista\_019: LO.a.1.4)

Cuando las estrategias anteriores fallan, el traductor pregunta a especialistas nativos de la LO el significado de los términos (LO.a.2).

Aparte de eso, no sé, el boca a boca por ejemplo, cuando estoy harta de buscar y no encuentro para qué sirve una función de la máquina, pues le pregunto a mi jefe, que es inglés y que es ingeniero, entonces él generalmente me hace incluso un dibujo, o me lo muestra con un lápiz y una goma de borrar y me hace... suele funcionar bastante bien, lo que pasa que su explicación me ayuda a comprender el término, pero no me ayuda a encontrar un equivalente en español. (entrevista\_011: LO.a.2)

Como última estrategia de búsqueda de comprensión en LO, el traductor pregunta al cliente (LO.a.3). En ocasiones la consulta al cliente puede pasar de ser la última estrategia, a utilizarse de las primeras si el traductor sospecha que no encontrará el significado de un término en los recursos porque este término solo lo utiliza su cliente, o también si cree que puede ser un término erróneo o estar mal escrito en el texto original.

Otras veces nos encontramos también con problemas de... el otro día no me acuerdo qué término era, era de puertos, creo que era *dársena ciega*, es una dársena sin salida, pues eso lo buscábamos y no estaba en ninguna otra página más que en la de ese puerto. En *Google* lo metías y en ninguna otra parte, ahí estás condenado a llamar al cliente y decirle qué es una *dársena ciega*, porque claro no lo puedes buscar, no puedes buscarlo en ningún otro idioma porque ni siquiera sabes qué es en castellano, en este caso era una traducción castellano-inglés, no sabes ni siquiera lo que es en castellano porque es un término que prácticamente ha acuñado el propio cliente, si nadie más lo utiliza entonces es que es de tu cosecha, explícanos qué es una *dársena ciega*, vale una dársena sin salida, vale tiene sentido. (entrevista\_022: LO.a.3)

#### 5.2.3.3.2 Estrategias de búsqueda para la resolución de problemas de transferencia (LM.b)

A continuación nos referiremos a las estrategias de búsqueda para acceder a denominaciones en LM que designan un determinado significado de la LO (LM.b). En ocasiones, estas estrategias de búsqueda se ponen en marcha después de que el traductor haya utilizado las estrategias que hemos visto previamente para comprender el significado expresado en la LO. Otras veces, las estrategias de búsqueda LM.b son las primeras que utiliza el traductor, cuando este ya comprende el concepto al que se refiere la LO y lo que necesita es una denominación que exprese dicho significado en la LM.

La primera estrategia que solían mencionar los traductores para resolver una necesidad de transferencia en LM es consultar en el diccionario bilingüe a partir de la denominación en LO para encontrar una denominación equivalente en LM (LM.b.1.1). Los traductores emplean diccionarios y glosarios bilingües o multilingües, generales y especializados, en formato impreso, electrónico y en línea. Por regla general, hay preferencia por los formatos electrónicos o en línea, aunque no todos los traductores y no en todas las especialidades, como veremos más adelante.

Si bien es cierto que los traductores, sobre todo los más expertos, se quejaron de las deficiencias del diccionario bilingüe para acceder a las denominaciones en la LM (por problemas de fiabilidad, de terminología descontextualizada, etc.) y por eso indicaron que no utilizan con tanta frecuencia los diccionarios bilingües, sobre todo porque hay una falta de diccionarios bilingües especializados de calidad, muchos todavía son en papel, y los diccionarios y glosarios especializados disponibles en línea suelen ser demasiado escuetos o de escasa fiabilidad. En la mayoría de los casos, el traductor combina la estrategia de búsqueda en el diccionario bilingüe con otras estrategias de búsqueda.

Si es algo general, primero son los diccionarios bilingües generales. Si sale, muy bien, si no... Si es algo jurídico, yo tengo en inglés, yo tengo.... es casi la única cosa que tengo en papel en casa y que está muy bien hecho, entonces busco ahí. (entrevista\_014: LM.b.1.1)

El bilingüe, pues términos o palabras que tendrías que saber pero a lo mejor en ese momento no te acuerdas, cosas no muy especializadas, lógicamente, y si es especializado, porque también acabas con algunos diccionarios bilingües especializados de medicina sobre todo o cosas así, muchas veces no me fío mucho. (entrevista\_029: LM.b.1.1; descontento\_diccionario\_bilingüe)

La búsqueda de terminología bilingüe en el diccionario bilingüe no suele aportar una solución definitiva. A esta estrategia de búsqueda siempre le suceden estrategias de búsqueda de verificación (véanse las estrategias de tipo LM.a.VERIFICACIÓN).

Los traductores también mencionaron que utilizan el diccionario bilingüe, no como una búsqueda definitiva para encontrar una denominación en la LM, sino como una aproximación a una serie de términos que pueden ayudar a realizar otras búsquedas para acceder a la denominación adecuada, una vez más a modo de «pistas», o también para confirmar o buscar mejores sinónimos a un término en LM que el traductor tiene en la cabeza (LM.b.1.1\_sinónimos).

Para palabras generales, digamos, que no es que sean muy específicas, pero yo no las conocía previamente, un buen diccionario bilingüe siempre te ayuda. A lo mejor no te da exactamente la palabra que tú quieres pero sí te encamina hacia cómo realizar una búsqueda de una palabra. Igual te da otras palabras de la misma familia, que puedes utilizar para buscar la que tú quieres. (entrevista\_013: LM.b.1.1\_sinónimos)

Los traductores con mayor trayectoria en la traducción de textos de una determinada temática señalaron que consultan glosarios propios con terminología bilingüe (LM.b.1.1\_glosarios\_propios) o sus archivos propios (LM.b.1.1\_archivos\_propios). En las organizaciones internacionales crean sus propias bases de datos que los traductores están obligados a utilizar como primera opción de búsqueda. Las agencias de traducción también elaboran glosarios y bases de datos propios.

Mi principal recurso creo que es mi archivo. Mi archivo del ordenador... miro traducciones antiguas y de ahí, vamos, voy haciéndome glosarios a partir de las traducciones para no tener que ir a la traducción, pero vamos casi mi archivo me ayuda más que el diccionario. (entrevista\_017: LM.b.1.1\_glosarios\_propios; LM.b.1.1\_archivos\_propios)

We have been creating, ever since this unit was created, we've been developing our own electronic database, it's called *IATE*. So there have been ongoing contributions to the Polish part of *IATE* and we are obliged to use that as a primary reference. (entrevista\_046: LM.b.1.1\_glosarios\_propios)

La siguiente estrategia se emplea en algunos diccionarios bilingües para acceder a términos compuestos a partir de la consulta de uno de los componentes del término (LM.b.1.2). Algunos diccionarios electrónicos permiten consultar a partir de una parte de un término, no necesariamente el término completo. Los traductores mencionaron que es una estrategia útil, por ejemplo para los traductores de alemán, una lengua en la que los términos compuestos son muy habituales.

Una de las cosas que más útiles encuentro es que puedes buscar partes de palabras, que dentro del físico es imposible, en alemán yo creo que es muy buen ejemplo, las palabras compuestas en el electrónico, puedo hacer la búsqueda por la segunda mitad de la palabra y en el físico eso no puedo hacerlo. (entrevista\_038: LM.b.1.2)

Asimismo, los traductores aludieron a la búsqueda en corpus bilingües paralelos (LM.b.1.3) como los desarrollados por organizaciones internacionales (por ejemplo *EUR-Lex*). Los traductores que trabajan en organismos internacionales están obligados a consultar el corpus de la organización como primera opción. No obstante, los traductores señalaron que existen pocos corpus paralelos a los que puedan recurrir.

Entonces mucha terminología especializada de agricultura aparece en las bases de datos legales de la UE, como toda esa legislación está traducida a muchos idiomas, a todos los idiomas oficiales, lo que hacemos son búsquedas, búsquedas en la página web de *EUR-Lex*. Entonces buscamos un texto en castellano, que contenga un texto legal en castellano, porque no siempre está todo en *IATE*, entonces buscamos un texto legal en castellano que contenga el término que estamos buscando. Miramos que el contexto coincida con lo que nosotros realmente necesitamos. Luego buscamos ese mismo texto legal en el idioma meta y vemos si nos puede valer o no nos puede valer. (entrevista\_022: LM.b.1.3)

Otherwise, we have to search for existing legislation to find the equivalence of the terms that have been used in legislation that is replicating, modifying or making reference to the existing legislation, so even if you know that something, ok I know it 100% that is called like that, you look into the previous translation of that act and then you... it is, oh it is not really called like that, it is called something else, and very often you have to stick with that because otherwise you would destroy the structure, the cross-reference of the whole text. (entrevista\_046: LM.b.1.3)

El principal corpus paralelo señalado por los traductores fue la memoria de traducción (LM.b.1.3\_memoria de traducción), tanto del propio traductor, como las proporcionadas por el cliente. Las memorias de traducción, en el caso de que las haya, tienen preferencia frente a cualquier otro recurso bilingüe al que pueda acceder el traductor. Si la memoria no muestra coincidencias automáticamente, los traductores a

veces realizan búsquedas de concordancias dentro de la memoria de términos específicos cuya equivalencia desconocen.

A medida que voy haciendo la traducción sí que me van saliendo los segmentos que están repetidos, si no están pues busco la palabra en el menú de concordancia y, si no la encuentro en la memoria, busco en otro sitio. (entrevista\_040: LM.b.1.3\_memoria de traducción)

Los traductores también indican que utilizan la Web como corpus paralelo. En esta estrategia de búsqueda, el traductor introduce en un buscador de Internet el término original, combinado con filtros de restricción de idiomas (páginas web en la LM o páginas con la extensión del país de destino, por ejemplo España), también combinado con palabras clave como *diccionario* o *glosario*. El objetivo es obtener como resultado páginas web bilingües, diccionarios, glosarios o enciclopedias bilingües en los que acceder a equivalentes en LM (LM.b.1.4), o incluso entradas de foros en los que se ha discutido una posible traducción (relación con la estrategia LM.b.7.1). Esta estrategia de búsqueda viene sucedida por una estrategia de verificación de los equivalentes encontrados (véanse las estrategias de tipo LM.a.VERIFICACIÓN).

Por ejemplo, a veces cuando tengo una duda muy puntual pongo la palabra +glossary. Eso sería para ver si hay algún glosario en inglés que contenga esa palabra, y a ver qué dice. Y si no, pongo de nuevo la palabra en inglés +glosario. En ese caso ya tendría que ser un glosario bilingüe porque si yo estoy poniendo el término inglés y estoy luego diciendo glosario, el sistema de alguna forma me va a tener que hacer una relación entre los dos idiomas y, normalmente, el 90% de los casos sucede esto, o sea que me devuelve como resultado un glosario bilingüe. Eso en *Google*, o si no, en vez de glosario, diccionario. (entrevista\_021: LM.b.1.4\_filtro\_diccionario)

Si no, voy buscando por ejemplo en casos de traducciones, por ejemplo, del inglés al italiano, como el inglés es una lengua muy difundida, suelo a veces, o sea, poner en el *Google*, seleccionando las opciones avanzadas de la lengua, poniendo solo páginas en italiano, o sea, de lengua italiana en páginas italianas, suelo poner directamente el nombre en inglés. Porque a veces se pueden encontrar textos bilingües o a lo mejor a alguien que comenta y dice esa palabra se traduce así, o sea, y puede servir como ayuda para darte una ayuda, por empezar un poco a buscar, porque no es que si encuentro esa palabra una vez y después la doy como exacta, después la vuelvo a poner esa palabra que he encontrado en italiano, vuelvo a hacer otra búsqueda y veo si efectivamente, en ese contexto se encuentra esa palabra y efectivamente quiere decir esa cosa. (entrevista\_015: LM.b.1.4; filtro\_idioma; LM.b.7.1; LM.a.VERIFICACIÓN\_Internet; pistas).

Asimismo, los traductores indicaron otros filtros que utilizan en las consultas en Internet: filtro de fecha para restringir los resultados a páginas de un determinado período de tiempo, por ejemplo páginas actuales; filtro de formato para limitar por ejemplo a textos en PDF que en un principio podrían ser textos más fiables, o buscadores especializados como *Google Scholar* para limitar los resultados de la búsqueda a un determinado tipo de textos; también utilizan palabras clave para recuperar o excluir páginas de un determinado tema.

Sí, incluso por ejemplo, estos se pueden meter todos en el normal, pero también utilizo el filtro de fecha, solo en el último año, solo en el último mes, este tipo de filtros también. Como digo por países, o sea .es o .com, .ac.uk, .edu en el caso de los americanos, también el *Google Scholar*. (entrevista\_008: LM.b.1.4; filtro\_fecha; Google Scholar)

Si estoy buscando algún documento que lo que quiero es que sea un artículo científico, pues puedo ir acotando que me busque por ejemplo PDF, archivos en formato PDF y ahí acoto, probablemente lo que me suele aparecer son manuales o son artículos, o alguna imagen en algún formato. (entrevista\_037: LM.b.1.4; filtro\_formato)

Hombre, normalmente en *Google* si no metes filtros encuentras de todo de lo que no quieres, entonces sí que normalmente meto palabras clave que me actúen de filtro. (entrevista\_017: LM.b.1.4; filtro\_palabras\_clave)

Asimismo, manifestaron la utilidad de buscadores en línea que funcionan a modo de corpus bilingüe paralelo. Por ejemplo, en la herramienta *Linguee*, que ha cobrado éxito últimamente, los traductores acceden a páginas bilingües de Internet que se encuentran alineadas (LM.b.1.4\_buscador\_web\_bilingüe).

Últimamente uso un recurso que me gusta mucho, que se llama *Linguee*, creo que es un proyecto de unos alemanes, la verdad es que me gusta porque puedes hacer búsquedas por texto y te aparecen muchos documentos que son por ejemplo de *EUR-Lex* o del Parlamento Europeo a los que generalmente no había acceso y te aparece un poco como una búsqueda que puedas hacer de una concordancia de *Trados*. Entonces te dice cada uno de los segmentos de dónde procede y puedes comparar si viene de Naciones Unidas de *UNTERM*, este recurso lo estoy utilizando últimamente bastante. (entrevista\_028: LM.b.1.4\_buscador\_web\_bilingüe; Web as Corpus)

Otro ejemplo de estrategia de búsqueda de terminología bilingüe en Internet, es la búsqueda en enciclopedias bilingües en línea, sobre todo en *Wikipedia* (LM.b.1.5). En ese caso, el traductor busca el término original en la enciclopedia *Wikipedia* y accede al artículo equivalente vinculado en la LM (LM.b.1.5.1); si la entrada del artículo en LM no encaja en la traducción, el traductor lee el artículo completo en LM para tratar de localizar un equivalente adecuado (LM.b.1.5.2). Asimismo, la *Wikipedia* puede ofrecer «pistas» al traductor para realizar otras búsquedas. En cualquier caso, después de esta estrategia siempre se utilizan estrategias de verificación (véanse las estrategias de tipo LM.a.VERIFICACIÓN).

Además, en la *Wikipedia* se utilizan estrategias LM.b.2, que veremos a continuación, porque a veces no se encuentra el artículo correspondiente en la LM, pero puede accederse a artículos en otros idiomas que sirven como puente para localizar después el equivalente en LM.

Puedo incluso ir a *Wikipedia*, vamos a decir que estoy en castellano y quiero en inglés, voy a *Wikipedia*, lo pongo en inglés y luego comparo, pero son formas así un poco de hacer un tapiz, no es que en *Wikipedia* vas a encontrar la traducción literal o exacta, pero me va dando pistas y voy como reduciendo para encontrar lo que estoy buscando. (entrevista\_007: LM.b.1.5.1; pistas)

También. Por ejemplo, tragándome el artículo entero a ver si me tropiezo con lo que busco y, si no, una forma más rápida sería que el concepto que busco tiene su propio artículo y la página esté traducida al alemán, entonces le doy a Deutsch y me sale cómo tiene que ser. Me pasó el otro día con los *potenciales evocados*, que es una técnica médica y que en alemán, quién lo habría dicho, es *Evozierte Potentiale*, o sea, eso lo encontré rápidamente en la *Wikipedia*. (entrevista\_009: LM.b.1.5.2)

Los traductores también están utilizando últimamente estrategias de búsqueda en metabuscadores (LM.b.1.6), que permiten la búsqueda simultánea en distintos recursos. Sobre todo en los grandes organismos internacionales, como en la UE utilizan *Quest*, que realiza búsquedas en las bases de datos de la UE (sobre todo *IATE*) y también en las de otros organismos internacionales, en la legislación (*EUR-Lex*), en la memoria de traducción *EURAMIS* y en multitud de recursos más; además es personalizable, permite al traductor limitar los recursos en los que se realizará la búsqueda y puede utilizarse a través de una macro desde el mismo procesador de textos.

Otro ejemplo es *dtSearch*, una herramienta que se emplea en los servicios de traducción de la ONU y contiene también todos los textos traducidos por este organismo. Ambas herramientas se han convertido en los principales instrumentos de búsqueda de estas instituciones. Asimismo, hemos observado que algunos traductores autónomos también utilizan *dtSearch* e indexan y clasifican por áreas temáticas glosarios, documentación que han traducido anteriormente, textos descargados de la red, o incluso páginas en línea (dependiendo de la versión del programa).

Algunos traductores también utilizan *plugins* o complementos de búsqueda en línea desde el navegador, donde seleccionan una serie de recursos en los que se realizará la búsqueda.

We have this combined search tool called *Quest*, it's like a combination of all the databases, so you try a word in this box, and then it brings you all the results from *IATE*, from *EUR-Lex*, from our archives. This is the first thing I use, and then through there I can go deeper, like the next level, more detailed. (entrevista\_044: LM.b.1.6; Quest; ventaja\_herramienta\_metabúsqueda; consulta\_simultánea\_distintos\_recursos)

Hay un programa, el *dtSearch*, lo utilizamos continuamente donde se va introduciendo todo, todo, todo, y los bitextos y los tritextos y todas las resoluciones y en concreto los documentos de Ginebra prácticamente todos. Con lo cual, eso es, además es lo que da fe. (entrevista\_048: LM.b.1.6; dtSearch)

Yo lo que he hecho ha sido automatizarme un plugin de búsqueda directamente desde *Firefox*, me resultaba lo más cómodo, pero claro, eso no es lo más habitual, pero vamos sí tienen su sistema propio de campos de búsqueda etc., etc., lo que pasa es que a mí me resultaba más cómodo definirme una serie de parámetros que con un solo botón automatizar. (entrevista\_006: LM.b.1.6; consulta\_plugin\_navegador)

Otra estrategia de búsqueda de terminología bilingüe es la utilización de terceros idiomas como puente entre la LO y la LM (LM.b.2). Esta estrategia funciona bien por ejemplo con lenguas cercanas, por ejemplo entre las lenguas romances (para acceder a un



equivalente en español al traductor le puede resultar útil encontrar un equivalente en francés o italiano). Esta estrategia se utiliza a menudo cuando en la combinación de lenguas del traductor existen pocos recursos bilingües disponibles (por ejemplo, en las entrevistas analizadas, japonés-francés, español-chino, inglés-estonio, etc.).

Esta estrategia suele aparecer seguida de una estrategia de tipo LM.b.3.1, que veremos más adelante, ya que una vez el traductor accede a un equivalente en una lengua que se asemeja a la suya, puede intuir cuál será el equivalente en su idioma y comprobarlo por medio de la consulta del equivalente intuido.

Claro, eso sí que lo hago bastante y de hecho en *IATE* no hago búsquedas del inglés al español, hago búsquedas del inglés al español, al francés, al italiano, al portugués. Y pido que me dé resultados en todos esos idiomas porque a veces no aparece, no tienen la combinación para ese término en castellano, pero sí que está en francés y a partir de ahí puedo hacer búsquedas en castellano en *Google* que me confirmen que efectivamente lo que yo he intuido a partir de otro idioma es correcto. (entrevista\_022: LM.b.2; LM.b.3.1)

Por ejemplo, una de las lenguas que se utiliza como puente entre dos lenguas es el latín, esta estrategia es útil cuando se trabaja con términos que tienen un equivalente en latín (nombres de plantas, animales, enfermedades, etc.). Esta estrategia está relacionada con la estrategia LM.b.1.4 de acceso a páginas web bilingües con el uso de filtros de restricción de idioma.

Primero, el traductor consulta el término original con restricción de idioma latín (introduciendo en la consulta la palabra *latín*). Como resultado accede a páginas en las que el término original aparece junto con su equivalente latino. Una vez se ha accedido al término en latín, se introduce dicho término restringiendo la nueva consulta a páginas de la LM. Como resultado se accede a páginas en las que probablemente encontremos el término en LM adecuado, que aparecerá junto con la forma latina.

Un ejemplo sería una planta. Hago mucha cosmética, entonces salen muchas plantas también. Entonces si no encuentro cómo es la traducción de la planta en castellano al alemán, la pongo en el portal en castellano y casi siempre me sale el nombre botánico y a través del nombre botánico la puedo buscar en alemán diciendo solo páginas en alemán. Eso funciona con plantas o con enfermedades y cosas así que tienen una denominación común en inglés o en latín y así puedes encontrar cosas. (entrevista\_009: LM.b.2)

La siguiente estrategia de búsqueda que mencionaron los traductores es la consulta a partir de una intuición de un posible equivalente, en un diccionario monolingüe, en un corpus o en Internet (LM.b.3). Existen dos posibilidades. Una, que el traductor consulte una posible denominación completa en LM para un concepto en LO (LM.b.3.1). Otra, que el traductor desconozca la denominación completa pero intuya un componente de la posible denominación completa en LM (LM.b.3.2). Esta última estrategia es útil para acceder a términos compuestos. Por ejemplo, el traductor puede introducir un término

incompleto intuido entre comillas en *Google* y sustituir con un asterisco aquellas palabras que desconoce.

Lo que yo suelo hacer mucho es inventarme una palabra, teclearla en la lengua meta a ver si he acertado y más o menos decir podría llamarse así, y luego verificar si realmente se llama así, que es un poco costoso por tiempo pero por otro lado satisfactorio. (entrevista\_010: LM.b.3.1\_Internet)

Las palabras clave suele ser entendido el concepto en inglés, o bien, palabras que yo pienso que puede contener el término en español pero que no he encontrado en un diccionario o de las que no estoy segura, algo así como propuestas de traducción, o palabras que creo que deben formar parte del término. Si es un término compuesto, por ejemplo, igual puedo intuir o puedo hacer una traducción más o menos literal de los términos del inglés al español y buscar la preposición por ejemplo, si es de o para, que no estoy segura de que tipo de preposición se utiliza en esa palabra compuesta en español. (entrevista\_012: LM.b.3.2)

También se da el caso de que el traductor combine la consulta del término original con una propuesta de equivalente en la LM, sobre todo en Internet (LM.b.1&3). Esta estrategia está relacionada con la estrategia de búsqueda de terminología bilingüe en Internet (LM.b.1.4), ya que el objetivo es acceder a páginas bilingües donde seguramente aparezca el equivalente intuido.

Near-azeotropic HFC +azeotrópico. Para saber cómo se dice en español (azetrópico). (entrevista\_021\_adicional: LM.b.1&3)

Asimismo, los traductores mencionaron la búsqueda en contexto por palabras relacionadas (LM.b.4) cuando las estrategias anteriores fallan: cuando la búsqueda de terminología bilingüe (a partir de la consulta del término original) no ofrece resultados de equivalencias (estrategias de tipo LM.b.1 y LM.b.2) y si el traductor tampoco es capaz de intuir un posible equivalente que pueda comprobar (estrategias LM.b.3).

En tal caso, el traductor trata de acceder a contextos en los que es probable que aparezca la denominación que desconoce para localizar la denominación en estos contextos (LM.b.4). Hemos visto que el traductor encuentra estos contextos de referencia consultando a partir de términos que guardan una relación con la denominación que el traductor desconoce y, por tanto, es probable que en el contexto en el que aparezcan estos términos relacionados también se encuentre la denominación deseada.

Respecto al tipo de relaciones que más útiles resultan a los traductores para acceder a estos contextos, nuestros entrevistados indicaron dos tipos de relaciones principalmente, las relaciones sintagmáticas y las relaciones semánticas.

Por un lado, se refirieron a la consulta de relaciones de tipo sintagmático, colocaciones en la LM típicas de la denominación desconocida (LM.b.4.1). En este caso, si el traductor desconoce una denominación en LM, consulta en un recurso colocaciones que deberán aparecer en el mismo contexto de la palabra que desconoce y, una vez que accede

al contexto, trata de localizar la palabra desconocida en ese contexto. Hemos observado que los traductores realizan la búsqueda de colocaciones en buscadores de Internet (LM.b.4.1\_Internet); en un corpus o en una memoria de traducción (LM.b.4.1\_memoria\_traducción), ya que las memorias tienen en cuenta el contexto que rodea a los términos para la selección de los equivalentes que proponen; o en un diccionario de colocaciones (LM.b.4.1\_diccionario\_colocaciones).

Quizá dentro de los monolingües, por ejemplo, lo que te estaba comentando antes de los diccionarios de idioms en inglés, eso sí que son... buscas de manera alfabética, pero no buscas digamos por la primera palabra que tú quieres buscar. Si a lo mejor la colocación que tú estás buscando o el idiom que tú estás buscando lleva *bird*, tú buscas *bird* aunque no sea la primera palabra, para ver todos los idioms que te devuelve que utilizan esa palabra, entonces eso no es alfabético. (entrevista\_013: LM.b.4.1; LM.b.4.1\_diccionario\_colocaciones)

Sí recuerdo poner frases entre comillas en *Google* para que buscara solo esa frase sabiendo que el final de esa frase iba a tener el término que buscaba. (entrevista\_042: LM.b.4.1; LM.b.4.1\_Internet)

En un corpus, en una memoria de traducción. Quizá ahí ya metería búsquedas, porque bueno es tan específico y tan rápido que seguramente metiera búsquedas, en inglés o en español, de términos parecidos, bueno del término, y de quizá algún término relacionado, que aparezca en el mismo contexto, en la misma frase donde lo he encontrado, seguramente haría ya una búsqueda referencial. (entrevista\_002: LM.b.4.1; LM.b.4.1\_memoria\_traducción)

Por otro lado, los traductores también acceden a los contextos consultando a partir de relaciones semánticas de tipo conceptual (sinonimia, hiperonimia, metonimia) (LM.b.4.2). En este sentido, hemos detectado distintas posibilidades.

Una de ellas es acceder a una denominación que se desconoce a partir de la consulta de un sinónimo (LM.b.4.2.1).

Luego también diccionarios de sinónimos y antónimos pueden dar alguna pista de acercarnos un poco más a quizás otros equivalentes que son un poco más adecuados para el contexto y que a lo mejor en ese momento no nos vienen a la cabeza. (entrevista\_024: LM.b.4.2.1; LM.b.4.2\_diccionario\_sinónimos)

Otra posible relación que utiliza el traductor es la hiperonimia-hiponimia (LM.b.4.2.2), es decir, consulta un hiperónimo para acceder a un hipónimo que desconoce, por ejemplo consultando a partir de una especie genérica de plantas (hiperónimo) puede accederse a la denominación desconocida de un tipo específico de planta de esa especie (hipónimo).

La máquina esta que te digo que he traducido hoy puede utilizar diferentes materiales para crear dientes, puede utilizar por ejemplo titanio, circonio, y entonces había varios materiales que no sabía cómo se decían en español porque son específicos, por ejemplo tienen un número específico en la tabla periódica y yo no lo sé porque no soy experta. Entonces muchas veces sí que pongo sencillamente en *Google* tipos de materiales para dientes, de lo general, y me salen un montón de páginas y una vez que veo la lista, veo cómo se escriben en español y también la ortografía porque por ejemplo circonio, puede ser con z o con c. (entrevista\_011: LM.b.4.2.2; LM.b.4.2\_Internet)

También es posible hacer uso de la relación parte-todo (LM.b.4.2.3), por ejemplo a partir de la consulta del nombre de una máquina (todo) se puede acceder a la denominación desconocida de un componente de esa máquina (parte).

Por ejemplo, un embalaje, en los tubos tipo pastas de dientes, entre lo que es la cabeza del tubo con el cierre y el orificio por donde sale la sustancia embalada digamos o envasada hay una pequeña lámina que protege el contenido del tubo contra bacterias y otras cosas, o sea una medida higiénica y esto en castellano tiene un nombre, el *opérculo*. Entonces claro, ¿cómo se dirá *opérculo* en alemán? El problema, encontrar una palabra para decir *opérculo* en alemán porque no la encuentras en ningún libro. [...] Pues como sospechaba o como describían lo que era el *opérculo* yo sabía de qué se trataba, entonces se trataba de averiguar cómo era esto en alemán. Entonces me metí en Internet en páginas de fabricantes de tubos para investigar allí a ver si encontraba una descripción de los elementos del tubo y efectivamente lo encontré. (entrevista\_009: LM.b.4.2.3; LM.b.4.2\_sitios\_relacionados)

Asimismo, los traductores en ocasiones hacen una descripción del concepto (que incluye tanto relaciones como características de dicho concepto) para acceder a la denominación que designa dicho concepto (LM.b.4.2.4). Los mismos traductores indicaron que introducen todo aquello que consideren oportuno, de manera que describiendo un concepto (utilizando tanto características de un concepto como relaciones con otros conceptos) lleguen a la denominación que desconocen.

O pones *dársena sin salida* en inglés en *Google*, con comillas, sin comillas a ver qué te sale, porque a veces pones la explicación del término y te sale el término que tú estabas buscando, por ejemplo si tú pones *dársena sin salida* entre comillas a veces en páginas de otros puertos te puede salir, un *close stock* es una *dársena sin salida*, entonces tú ya sabes, tú ya tienes un término en inglés que has encontrado a raíz de la explicación del término, a veces tenemos que hacer esas búsquedas y entonces es hacer de Sherlock Holmes, de ir buscando, a veces encuentras y a veces no encuentras. (entrevista\_022: LM.b.4.2.4; LM.b.4.2\_Internet)

Por último, observamos que también realizan consultas a partir del área temática en la que se enmarca la denominación que desconocen (LM.b.4.2.5).

Si no, por ejemplo, esto ya para casos más concretos, lo que suelo hacer es buscar a lo mejor textos sobre esa situación o de ese ámbito, eso para terminología más especializada, si sé a qué sector, a qué campo pertenece, normalmente si es algo que no recuerdo, pero sé por dónde hay, puedo a lo mejor ir a buscar a una página de información de ese estilo. Por ejemplo, imagínate, si es de informática y es un término que sale en un ámbito de sistemas operativos pero en ese momento no sé cuál es, pues siempre intento buscar material de referencia sobre sistemas operativos e intentar buscar del texto que estoy traduciendo algo que trate de eso a ver si encuentro ahí algún término que me pueda llevar a eso. (entrevista\_003: LM.b.4.2.5; LM.b.4.2\_Internet)

Las búsquedas de relaciones semánticas de tipo conceptual se realizan en Internet (LM.b.4.2\_Internet), diccionarios de sinónimos (LM.b.4.2\_diccionario\_sinónimos), diccionarios o redes de ideas (LM.b.4.2\_diccionario\_ideas), sitios relacionados (LM.b.4.2\_sitios\_relacionados) o en la página web del cliente (LM.b.4.2\_sitio\_web\_cliente). En este punto cabe decir que los diccionarios de ideas, con terminología relacionada, parece que resultan interesantes también a los intérpretes.

Uso bastante y sobre todo ahora en interpretación, palabras afines, es un poco la idea que lleváis del diccionario, pones una idea y suelen aparecerte palabras relacionadas con esa idea, pones por ejemplo *alimentación* y te hace... ideasafines.com.ar, pones *comida* y te aparecen unos enlaces y te pone *nutrición*, puedes entrar dentro de *nutrición* y ahí dentro de *nutrición* hay una serie de palabras relacionadas, luego *paella*, *ingestión*, *desayuno*, cada uno de estos subtítulos contiene un número de palabras relacionados con el tema, entonces aquí tienes *banquete*, *comida*, *comilona*, *festejo*, *festín*, *fiesta*, *reunión*; *partes*, pues partes de un banquete; luego tienes *aperitivo*, *entrada*, *entremeses*, *principio*; *convidado*, pues *comensal*, *invitado*, *asistente*, *concurrente*; entonces la idea, esto lo hago a veces cuando tengo que hacer una especie de brainstorming para prepararme a lo mejor alguna interpretación y de alguna forma quiero generar vocabulario, entro aquí y hago una lectura en voz alta, o hago una extracción de términos que me puedan servir, o lo mismo como si utilizara un diccionario de sinónimos. (entrevista\_028: LM.b.4.2\_diccionario\_ideas)

Asimismo, en la búsqueda de palabras desconocidas por relaciones, el acceso a imágenes en las que se encuentran identificadas las palabras relacionadas puede resultar de utilidad (LM.b.4.3). Esto puede hacerse en buscadores de imágenes de Internet y también en diccionarios visuales (preferiblemente bilingües), aunque los traductores reclamaron que los diccionarios bilingües visuales escasean.

En alemán hay un diccionario que se llama *Duden*, que es de imágenes Tú por ejemplo tienes *carpintería*, y entonces tienes en esa sección un montón de... tienes el dibujito de la herramienta, la acción, de lo que sea, y entonces tienes la terminología en alemán, en inglés y en español, por ejemplo. Y entonces a veces sí que te sirve cuando vas a buscar... *pianos*, pues entonces tienes todas las partes de un piano. *Anatomía*, y entonces tienes todas las partes. Tú buscas por tema y dentro del tema está la imagen Lo tengo en alemán porque, que yo sepa, no existe en otros idiomas, es una cosa muy alemana, y ese diccionario alemán contiene inglés y español como mínimo, y puede contener otros idiomas, pero el *Duden* yo no sé que haya otras ediciones. (entrevista\_026: LM.b.4.3)

Me suele pasar sobre todo con tipos de tornillos o con tipos de tejados me pasó hace no mucho. Eran para placas fotovoltaicas para los tejados y claro, dependiendo del tipo de teja había que poner un soporte u otro y yo no sabía que existían tantos tipos de tejas, aparte que era un texto alemán y allí tienen tejas diferentes a las que se utilizan aquí pues por las condiciones meteorológicas, igual en otra parte de España también tienen tanta cantidad de tejas, pero yo aquí conocía 3 o 4, entonces claro, era bastante difícil saber cómo se llamaban ese tipo de tejas aquí si no las habías visto nunca. Busqué en empresas de tejas de España y en empresas de tejas en Alemania y me puse a buscar las formas de las tejas en cada una hasta que veía que coincidía el tipo de teja, entonces ya sabía cómo se llamaba en español. (entrevista\_031: LM.b.4.3)

Las estrategias de búsqueda LM.b.4, en las que se consulta a partir de palabras relacionadas, campo temático o una imagen, son ejemplos de búsquedas que los traductores identifican como de tipo onomasiológico, ya que en este caso tratan de acceder a una denominación desconocida a partir de una información de tipo conceptual (de las relaciones que guarda la denominación desconocida con otras denominaciones que el traductor sí que conoce).

Si todas las estrategias anteriores fallan, una posibilidad que le queda al traductor para acceder a la terminología que necesita es leer textos paralelos (LM.b.5) hasta que localiza la denominación que desconoce. Puede darse el caso de que el traductor conozca mucho un determinado tema y sepa que haciendo una lectura rápida en textos de ese tema,

sería capaz de identificar la palabra que no recuerda. Asimismo, la estrategia de lectura de textos paralelos se emplea especialmente cuando la necesidad del traductor no es de terminología puntual, sino que en general desconoce el tema que está traduciendo y precisa realizar una documentación previa, que le ayudará después a tomar decisiones más acertadas en cuanto a la terminología.

La lectura puede realizarla el traductor en páginas relacionadas con el tema en Internet (LM.b.5\_Internet), en los archivos propios del traductor (LM.b.5\_archivos propios), o incluso en documentación en papel si es necesario (LM.b.5\_documentación\_papel).

También es importante aclarar que las búsquedas anteriores a partir de terminología relacionada en contexto (LM.b.4 y sus variantes) requieren en la mayoría de los casos una mayor o menor fase de lectura en dichos contextos, ya que en ocasiones el equivalente deseado se localizará rápidamente, pero en otros casos es posible que el traductor deba leer más de lo esperado.

Sobre todo, si conozco muy bien el tema y sé que está centrado en algo y que buscando algo de ese tema puede que me salga la palabra y la sé reconocer sí, igualmente si no conozco para nada el tema, informarte antes acerca del tema, leer un poco en general de qué se trata para luego poder profundizar, si no sabes de qué va el tema es imposible que encuentres una palabra específica que no está en un diccionario, no la sabrías reconocer en el texto, por decirlo de alguna manera. (entrevista\_013: LM.b.5)

Me acuerdo en concreto de una pieza de un fusil neumático de pesca submarina que tenía toda la descripción en inglés de cómo era la forma y donde estaba el fusil y todo, pero no sabía, no encontraba por ninguna parte cómo se decía en español y al final buscando textos paralelos que hablasen de las partes de un fusil en español, al final lo encontré. ¿Pero tuviste que leer? Sí, mucho. (entrevista\_042: LM.b.5; LM.b.4.2.3)

Los traductores indicaron que, en ocasiones, el acceso a la terminología que desconocen se produce por casualidad (LM.b.6), aunque siempre como resultado de la puesta en marcha de varias de las estrategias de búsqueda anteriores.

Yes, and again Internet, this is mostly a hazard search, you try to see more or less, you know which area you are operating in, you see which other words might be used, you try to search your memory, what it tells, which area is that, what do I know of that area, sometimes from memory you do that, sometimes is helpful, you know, you can get images of the thing, you introduce the word, you get images, and you may want to try, to try, aha, maybe this, but mainly it is haphazard I might say. (entrevista\_046: LM.b.6)

Por último, los traductores mencionaron estrategias de búsqueda en las que consultan directamente a personas, lo que denominaríamos recursos humanos. La primera estrategia de búsqueda de este tipo está relacionada con la consulta en foros (LM.b.7). En la mayoría de ocasiones, los traductores antes de lanzar una pregunta a un foro, consultan la base de datos de preguntas y respuestas previas para comprobar si su duda ya se ha resuelto antes (LM.b.7.1), lo que suele dar resultados, ya que en foros de traducción, como por ejemplo

en *ProZ* (que es el foro que más traductores nombraron), las consultas que se han realizado anteriormente se clasifican por palabras clave, como en un diccionario, y los traductores reconocen que normalmente la duda que motiva la consulta en un foro ya la ha tenido otro compañero y está almacenada en la base de datos del foro.

Primero buscando en los archivos por si alguien lo ha preguntado antes, lógicamente, no vas a hacer perder el tiempo (entrevista\_008: LM.b.7.1)

En el caso de que la duda no se encuentre resuelta ya en el foro, el traductor en ocasiones lanza una pregunta al foro, que puede ser un foro de traductores (LM.b.7.2) o un foro de especialistas (LM.b.7.3). Hemos observado una mayor tendencia de los traductores a consultar en foros de traductores que en foros de especialistas, aunque algunos foros reúnen tanto a traductores como a especialistas en la materia. Por ejemplo, los foros *Tremédica* y *Medtrad* incluyen tanto traductores de medicina, como médicos. Más adelante, cuando nos refiramos a las ventajas y desventajas de los foros explicaremos por qué los traductores indicaron que realizan pocas preguntas en foros o que esta es una de las últimas estrategias de búsqueda que utilizan.

Sí, si no encuentro el término por otro método, pregunto en un foro. Normalmente son preguntas sobre un término que no conozco y no he podido encontrar en otro sitio. (entrevista\_030: LM.b.7.2)

Los foros también pueden ser muy interesantes porque en los foros habla gente del área en la que estás trabajando y del idioma al que quieres traducir. Entonces, ellos no se preocupan por cómo se decía en la LO, entonces eso puede ser muy interesante pero no tiene que ser la primera etapa. (entrevista\_023: LM.b.7.3)

Aunque los traductores no otorgan a los foros la máxima fiabilidad, reconocen que pueden ofrecer «pistas» útiles, que después pueden verificar en otros recursos (véanse las estrategias de tipo LM.a.VERIFICACIÓN).

Sí, la red de traductores *ProZ*, que uso bastante porque tiene muchos recursos, mucha gente que ha preguntado cosas, así que a veces puede no dar respuesta porque no te puedes fiar al 100%, pero sí puede dar pistas. (entrevista\_014: LM.b.7.1; pistas)

Los foros de traductores también sirven para forjar nueva terminología cuando esta no existe o no es adecuada.

En este caso particular sí, pero en otros casos en los que no estaríamos hablando, sobre todo en cuestiones donde en España... por ejemplo *paella*, esa traducción al chino no la encuentras, no existe, normalmente lo que he encontrado es que lo traducen como *arroz de mariscos*, no es correcto, paella valenciana no es arroz de mariscos, ni de lejos. Qué es lo que hago en estos casos, yo no puedo traducirlo como arroz con pollo y conejo, garrofón y tal, es ridículo, además nadie se lo leería. Yo lo que hago es una traducción fonética, yo traduciría pa-e-lla al chino, una traducción fonética, y a pie de página pondría arroz de no sé qué, que es la comida típica de la Comunidad Valenciana, de España, y tal. Estoy orgulloso porque ese término de *paella* lo puse yo y ahora me encuentro que todo el mundo lo llama así. Yo la verdad es que lo defendí bastante en los foros, con bastante pasión y yo creo que como Internet es una cosa tan extraña, que al final se ha difundido. Entonces, yo puedo decir que *paella* en chino lo he inventado yo. Ahora ya cada vez encuentro menos *arroz de marisco*. Este es un caso que te

comentaba que no existe, es algo que hay en España que no existe en China, porque de hecho en inglés se traduce tal cual, *paella*, o *Spanish rice*. También hay mucha confusión... *gazpacho*, lo que pasa es que con *gazpacho* no lo he conseguido, la traducción de *gazpacho*, la más extraordinaria es *sopa de color amarillo de verduras*, que de entrada no es amarillo, es de color naranja, pero bueno, yo intenté poner una especie de sopa fría y tal, y también intenté poner la traducción fonética de *gazpacho*, pero se ve que no le ha gustado a la gente y lo siguen llamando como *sopa amarilla*, pero como no soy andaluz me da igual. (entrevista\_027: LM.b.7.1)

Asimismo, como última o una de las últimas estrategias de búsqueda, los traductores se ponen en contacto con un especialista en la materia (LM.b.8.1). Los traductores suelen contactar con amigos o familiares que son especialistas de un determinado campo. A veces, también tratan de buscar a especialistas en la materia que no conocen personalmente. En cualquier caso, el especialista es el último recurso porque, según afirmaron, es un recurso muypreciado, al que no quieren molestar si no es absolutamente necesario. En las organizaciones internacionales, las preguntas a especialistas son más frecuentes ya que, como nos comentaron, los traductores pueden ponerse en contacto con especialistas de la división que necesiten y plantearles sus preguntas.

No, pero tengo una red de colaboradores, sobre todo en china, donde hay médicos, hay abogados, ingenieros, de manera que si yo tengo alguna duda en un campo en particular, cuando el texto lo requiere, cuando es preciso, en este caso sí que contacto con ellos, más que nada porque la terminología en china, por ejemplo en el campo de leyes, cambia completamente, no tiene nada que ver lo que hay en España con lo que hay allí, entonces buscar una correspondencia es difícil. En este caso tienes que ir al concepto. Yo intentaría explicar por teléfono o por e-mail el concepto que estás buscando al colega, de manera que él al final siempre te encuentra una terminología o una frase que exprese ese concepto. (entrevista\_027: LM.b.8.1; llamada\_teléfono; e-mail)

Alguna vez he hecho alguna pregunta en redes sociales como *Twitter*, no he preguntado en ningún foro porque digamos que es más difícil que te contesten en el momento, pero en redes sociales como *Twitter*, que es instantáneo, sí lanzo alguna pregunta específica y me han contestado en cuestión de segundos, pero salvo eso. Tanto especialistas como traductores, como el *Twitter* lo sigue por suerte bastante gente, tanto especialistas en mi campo, como traductores, pues es una gran ayuda. Yo lanzo la pregunta, ¿alguien me podría confirmar si aparece tal...? Eso es. Lo he usado poco, pero cuando lo he tenido que usar ha sido un resultado rápido y bastante fiable. (entrevista\_035: LM.b.8.1; LM.b.8.2; redes\_sociales)

En ocasiones, el traductor también se pone en contacto con colegas, otros traductores (LM.b.8.2).

Well, the collective expertise of our colleagues. I mean of our colleagues of our peers. Sometimes I think it might be a valuable shortcut if I know that at least four people will have been dealing with fishery in my department but I can't quite locate which for these were, and I happen to do a text on fishery, which is not my field of interest, there are some divisions within here so I am in the technical field, I'm coordinating the group of technical translation and there is fishery, and there is fiscal policy, and there is health, and there is education, groups of interest, groups of colleagues. Sometimes I just send out a circulate e-mail, please help me, have you ever come across this? This is a practice here. (entrevista\_046: LM.b.8.2; e-mail)

Finalmente, algunas grandes agencias de traducción y los organismos internacionales tienen grupos de expertos terminólogos a los que los traductores pueden remitir sus dudas terminológicas (LM.b.8.3).



Los foros muy pocas veces he tenido que recurrir a uno porque con las empresas con las que trabajo solemos tener a algún líder del grupo que es el que se encarga de la terminología y entonces solemos tener debates cerrados por *Skype*. (entrevista\_031: LM.b.8.3; skype)

También cada unidad tiene su grupo de terminólogos y si tienes algún término que no puedes encontrar, lo has buscado en diccionarios, has hablado con expertos, etc., tienes que referirlo a ellos. (entrevista\_043: LM.b.8.3)

El contacto con especialistas y compañeros suele realizarse por correo electrónico o por teléfono; los traductores también mencionaron que últimamente para este tipo de estrategia de búsqueda está funcionando bien la comunicación a través de las redes sociales.

Finalmente, si todos los recursos anteriores no han resultado útiles o el traductor sigue teniendo dudas, pregunta al cliente (LM.b.9). Normalmente, el cliente es el último recurso que utiliza el traductor, salvo en determinados casos en los que el traductor sabe que no va a encontrar la respuesta por otros medios, o cuando necesita una confirmación de la terminología que prefiere el cliente.

También puede ocurrir que te sale simplemente tradúceme este menú o esta lista de capítulos o de índice y no tienes ni idea, entonces si realmente existe la posibilidad de que pudiera ser distintas cosas, se pregunta al cliente porque no vas a poner el que no es, digo yo. (entrevista\_008: LM.b.9)

No, depende, hay búsquedas que sí, que a lo mejor les dedicas tres minutos, otras les dedicas diez y otras les dedicas una hora. Y cuando llevas una hora buscando, desesperada, entonces es cuando dices, duda al cliente, un correito, con una lista de términos y dices, hasta aquí, no le puedes dedicar más tiempo, por cuestiones de plazo y de rentabilidad. No puedes. (entrevista\_022: LM.b.9)

#### 5.2.3.3.3 Estrategias de búsqueda para la resolución de problemas de verificación

Como veníamos avanzando en la descripción de las estrategias de búsqueda anteriores, es una práctica habitual entre los traductores emplear estrategias de verificación de los términos en la LM que desconocían y que han encontrado por medio de la búsqueda en algún recurso. En este sentido, el traductor normalmente no se queda directamente con la primera opción que encuentra, sino que la comprueba en otro recurso. También es común que el traductor compare la validez de dos posibles equivalentes que ha encontrado para decantarse por uno. Estas estrategias de verificación son estrategias en LM de recepción (LM.a.VERIFICACIÓN), porque se parte de una denominación en LM para acceder a su definición, y comprobar que el significado en LO y LM coincide.

La verificación puede realizarse consultando la palabra en el diccionario monolingüe para comprobar si recoge el sentido que expresa el texto original, también para comprobar

la ortografía, variantes regionales, o para comprobar si es un término normativo, o si el término existe (LM.a.1.1; LM.a.VERIFICACIÓN\_diccionario\_monolingüe).

Y el diccionario monolingüe en LM sirve para estar seguro que realmente he encontrado el término correcto, para confirmar, o para buscar sinónimos. (entrevista\_030: LM.a.VERIFICACIÓN\_diccionario\_monolingüe)

Sí, lo utilizo muchísimo, utilizo el diccionario de la RAE, como digo intento ser siempre normativo, lo utilizo muchísimo y el *DPD*. Exacto, sobre todo como en informática hay mucho anglicismo pues para saber hasta qué punto está aceptado. (entrevista\_035: LM.a.VERIFICACIÓN\_diccionario\_monolingüe)

En pocas ocasiones, el traductor busca la verificación en un corpus (LM.a.1.2; LM.a.VERIFICACIÓN\_corpus), por ejemplo en el corpus *CREA*, de la Real Academia, por razones de fiabilidad, para contar con una base fiable que permita justificar la elección al cliente.

También consultamos mucho el *CREA* y el *CORDE* de la RAE. Pues muchas veces para aclarar dudas con los clientes. Hay muchos clientes que revisan textos muy alegremente y hacen cambios que no vienen a cuento, o incluso cambios incorrectos terminológicamente y los utilizamos como base para demostrarle que está equivocado en algunos casos. (entrevista\_019: LM.a.VERIFICACIÓN\_corpus)

Otras veces, la verificación se la da al traductor encontrar la terminología en la página web del cliente porque es al que va dirigida la traducción (LM.a.1.2; LM.a.VERIFICACIÓN\_página\_web\_cliente)

Los corpus los usé muy poco y eso justamente es para lo que tú dices, o sea para ver cómo funciona la palabra en contexto, pero lo que hago es, además de tener diccionarios y glosarios especializados, tanto en papel, como guardados en favoritos, también tengo sitios guardados en favoritos, por ejemplo tengo el sitio de la Asociación Estadounidense de Marketing, por ejemplo todas las ONG para las cuales hago trabajos también tengo los sitios guardados porque algunas ONG también tienen sus glosarios ya desarrollados. Entonces, en ese sentido, en ese caso hay que respetar la terminología que ellos usan, no puedes ir con una traducción y decir, bueno, en el diccionario me decía tal cosa. Ellos te van a decir, no, deberías haber verificado en nuestra página web oficial que nosotros utilizamos la palabra de tal forma, y también para comercio exterior, o sea para todas las especialidades que yo tengo, tengo enlaces guardados. (entrevista\_021: LM.a.VERIFICACIÓN\_página\_web\_cliente)

Otro medio de verificación que utilizan hoy día casi todos los traductores es Internet (LM.a.1.2; LM.a.VERIFICACIÓN\_Internet), donde los traductores encuentran más ejemplos de la terminología en contexto, siempre, como ellos indicaron, accediendo a sitios web fiables, que tienen la terminología actualizada.

Sí, normalmente... a ver, si no sale exactamente como yo lo quiero en el bilingüe, no lo saco solo del bilingüe, o sea siempre lo cotejo de otra manera, viendo que se utiliza de verdad, que la palabra que de verdad te devuelve el diccionario bilingüe se utiliza en ese ámbito, porque a veces el bilingüe lo denomina de una manera y tú luego lo buscas en el ámbito al que tú lo quieres aplicar y no se utiliza esa palabra. Entonces sí que suelo cotejarlo, sobre todo, poniéndolo por ejemplo en *Google* y viendo si salen por ejemplo páginas de autoridades que utilizan ese término en el ámbito en el que tú te estás refiriendo. (entrevista\_013: LM.a.VERIFICACIÓN\_Internet)

Otro medio de verificación que utiliza el traductor son las imágenes (LM.a.1.3; LM.a.VERIFICACIÓN\_buscador\_imágenes). Se trata de buscar una imagen a partir del equivalente que se ha encontrado y se compara si las imágenes que aparecen del término equivalente son las mismas que las del término original (en este sentido, esta estrategia de búsqueda está relacionada con la estrategia LO.a.1.3).

Por ejemplo hoy mismo me ha salido un término que era pues *Trapezblech* pues en principio la traducción que se te ocurre es *chapa trapezoidal*, es una traducción muy directa, pero al ser tan directa como que no me fiaba mucho, en *Google* escribo *chapa trapezoidal*, comparo el número de resultados con los de otras alternativas que podía haber pensado, pero sobre todo la confirmación me la da la búsqueda de imagen, *Trapezblech* y *chapa trapezoidal*, las imágenes eran iguales, era lo mismo. (entrevista\_038: LM.a.VERIFICACIÓN\_buscador\_imágenes)

#### 5.2.3.3.4 Estrategias de búsqueda para la resolución de problemas de producción (LM.c)

Finalmente, cuando los traductores han encontrado un equivalente satisfactorio en LM, también precisan conocer cómo se combina dicho equivalente en contexto, por lo que necesitan encontrar fraseología o colocaciones (LM.c). La necesidad de este tipo de información observamos que se da más en el caso de que los traductores realicen traducción inversa, ya que no dominan la fraseología como en su lengua materna; también les ocurre a los traductores que llevan tiempo viviendo fuera de su país de origen; y en la traducción técnica, ya que los traductores no son expertos y desconocen la fraseología del campo de especialidad. El traductor accede a las colocaciones en un corpus (LM.c.1.1), en Internet (LM.c.1.2), en un diccionario monolingüe (LM.c.1.3) o en un diccionario de colocaciones (LM.c.1.4).

Y en relación con esto, uso bastante también el *CREA*, el *CREA* sí que, lo mismo, el *CREA* lo que pasa es que como es electrónico, lo uso bastante porque este de Ignacio Bosque, tienes el tocho aquí que tienes que ir pasando, pero el *CREA* me viene muy bien, pones los términos y te aparece dónde ha aparecido, si es en la prensa, durante los últimos años. [...] A veces simplemente que notas que te falta algún adverbio, algún adjetivo, que no sabes muy bien, lo mismo estás viendo el que te aparece en inglés, te parece que lo usas demasiado y quieres cambiar un poquito, darle un poquito de color y lo pongo ahí para ver cómo lidian con esto otros escritores. Para colocaciones, sobre todo para colocaciones. Pongo la palabra, a lo mejor pongo dos o tres palabras que de alguna manera es muy obvio la palabra que va a seguir, si dijéramos que tengo una duda con una preposición y no tuviera muy claro que es *finalizó con*, pondría algo que finalizó, lo pondría en el *CREA* y trato de buscar si en alguno de los ejemplos viene esa preposición que no tengo clara. (entrevista\_028: LM.c.1.1)

Ahora puedes meter lo que se llama un wildcard, que es un asterisco y eso sustituye una cantidad indeterminada de texto; una palabra o varias. [...] Es útil, cuando dices no sé qué, no sé qué, no sé cuántos, tal, y entonces te falta un huequito en medio o tú has pensado que es un verbo pero realmente se utiliza con otro o has pensado que es con una preposición pero realmente es con otra. (entrevista\_016: LM.c.1.2)

No, lo que yo busco siempre es el verbo para el sustantivo y casi siempre en textos jurídicos, que es muy difícil en alemán esto, o bien porque después de tantos años sin estar en Alemania se te olvida o no estás seguro. Entonces eso se busca. Buscas el sustantivo e intentas ver con qué verbos se combina. Sí, o introduzco alguna sospecha que tengo en Internet a ver si se me confirma muchas veces, entonces tiene que ser verdad, o en el diccionario simplemente esperando que me ponga un ejemplo con este sustantivo, que me pongan el verbo. (entrevista\_009: LM.c.1.2; LM.c.1.3)

Si es un problema del español utilizo mucho el *DPD*. Sí lo tengo en la barra de *Mozilla Firefox* y el *DRAE* también. Esos dos los utilizo permanentemente. Más que nada el *DPD* porque el *DPD* trae por ejemplo en algunos casos de los verbos lo que sería el régimen preposicional del verbo. Entonces, a veces, cuando hay dudas en cuanto a las preposiciones, a pesar de que tengo otro diccionario de preposiciones que es el de Alicia Zorrilla, una doctora de argentina que es miembro de la Real Academia Española, pero un poco lo que te decía, tengo el diccionario aquí en la biblioteca pero la comodidad de ir al ordenador y poner la entrada y ya está ahí todo, a veces con las prisas de los proyectos eso es totalmente comprensible, pero igualmente en ambos casos son dos fuentes muy confiables. (entrevista\_021: LM.c.1.3; LM.c.1.4)

#### 5.2.3.4 Resultados sobre percepciones y expectativas de los usuarios respecto de los recursos

A continuación, incluiremos los resultados sobre la opinión de los traductores respecto de los recursos existentes, ventajas, desventajas y demandas.

El uso del **diccionario**, y en concreto del diccionario bilingüe, por parte de los traductores es menor con la adquisición de experiencia y también ha descendido su uso desde la aparición de Internet.

Al principio, cuando yo empecé a traducir debo admitir que no era así porque justamente por la falta de experiencia consultaba mucho o demasiado el diccionario, dependía demasiado. Después sucedió que, además de adquirir la experiencia en la traducción y las búsquedas, por supuesto que Internet vino después, porque yo me gradué en el 93 y empecé a tener Internet en el año 96. (entrevista\_021: uso\_diccionario\_bilingüe\_frecuencia; menos\_con\_adquisición\_experiencia; menos\_desde\_aparición\_Internet)

El formato en el que la mayoría de los traductores prefieren consultar el diccionario es el electrónico, la tendencia es que apenas se emplea el papel, salvo en algunas excepciones de diccionarios que no existen en formato electrónico, como veremos más adelante.

Yo todo en electrónico, no utilizo en papel. (entrevista\_006: uso\_diccionario\_papel\_frecuencia)

No, tenemos un diccionario específico técnico inglés-español y lo utilizo muy muy pocas veces, creo que este año, sin exagerarte, he buscado una palabra. (entrevista\_011: uso\_diccionario\_papel\_frecuencia)

Asimismo, hemos observado que los diccionarios actuales que podrían considerarse de naturaleza onomasiológica (tesauros, diccionarios ideológicos, etc.) no están resultando útiles a los traductores y, por tanto, apenas los utilizan.

No, y eso ahí sí que he intentado utilizar tesauros y nunca he entendido la lógica de los tesauros para buscar terminología, eso nunca me ha funcionado, alguna vez lo he utilizado pero no. A lo mejor alguna vez sí, pero me parece que hay que conocer la lógica de la materia para encontrar la solución a no ser que sea un tesoro informatizado y puedas hacer una búsqueda por secuencia, pero me parece que requiere ya un conocimiento de las relaciones entre los conceptos y si no la tienes pues yo no... la búsqueda es más incierta porque tienes que lanzarte a lo desconocido, tienes que adivinar prácticamente. (entrevista\_039: uso\_tesoro\_diccionario\_ideológico)

En cuanto al uso del **corpus**, los traductores utilizan más bien las memorias de traducción como corpus bilingües y también Internet como corpus (Web as Corpus) tanto monolingüe como bilingüe.

Corpus, buf, la verdad es que también poco. Alguna vez he hecho algún trabajo terminológico específico, pero en general no los suelo usar. A veces sí, cuando a lo mejor hablas de un producto nuevo, intento buscar un documento parecido en español para saber qué terminología y expresiones usan, pero un corpus grande no suelo usar. (entrevista\_035: uso\_corpus)

Los **foros** también son recursos menos utilizados, o al menos que no se emplean como primera opción de búsqueda.

También muy de vez en cuando se ha preguntado en algún foro. No descarto el buscar en los foros tampoco, pero esto no siempre sale y bajo mi punto de vista para no encontrar lo que busco y tener que ir a una fuente primaria de todas formas, pues voy directamente a las fuentes primarias. (entrevista\_008: uso\_foro\_frecuencia)

Finalmente, los traductores indicaron que **Internet** es el recurso más útil de todos los que utilizan, sobre todo el buscador *Google* lo han calificado como un recurso imprescindible.

Sin Internet no podría trabajar, obviamente. (entrevista\_003: imprescindible)

*Google* es Dios. Es el sitio que empiezas a buscar, a partir de ahí te lleva a los sitios que puedes encontrar una respuesta, y no existe en papel, San *Google*, *Google* es imprescindible, si hubiera alguna alternativa igual de bien o mejor, me daría igual, no tengo acciones en *Google*, pero... (entrevista\_029: imprescindible)

#### 5.2.3.4.1 Ventajas y desventajas observadas en los recursos

El uso mayoritario o minoritario de los recursos se debe a una serie de ventajas y desventajas que los traductores observan en ellos. A continuación indicaremos primero las desventajas y después las ventajas mencionadas por los traductores para cada uno de los recursos.

##### 5.2.3.4.1.1 Ventajas y desventajas observadas en los diccionarios

Respecto a las **desventajas del diccionario**, el motivo principal del descontento de los traductores, tanto de los diccionarios monolingües como bilingües, es que son demasiado generales y la información se desfasa rápidamente. Además, existen pocos diccionarios bilingües especializados.

Sí, la primera opción es el bilingüe, aunque ya si es algo que ya veo que es demasiado técnico a veces, ni siquiera busco en el bilingüe porque sé que ya no lo voy a encontrar. (entrevista\_015: descontento\_diccionario\_bilingüe; descontento\_demasiado\_generales)

Teniendo en cuenta que trabajo en el campo de las telecomunicaciones y las tecnologías de la información, la mayoría de los diccionarios me suelen resultar inútiles porque están desfasados. (entrevista\_006: descontento\_diccionarios; descontento\_terminología\_desfasada)

En el campo terminológico de las telecomunicaciones y la IT, la verdad es que diccionarios bilingües de calidad actualizados existen pocos, al final siempre acabas confiando en recursos abiertos de usuarios. (entrevista\_006: descontento\_diccionario\_bilingüe; pocos\_diccionarios\_bilingües\_especializados)

Asimismo, los traductores criticaron al **diccionario bilingüe** porque no añade información que ayude al traductor a seleccionar el equivalente más adecuado, solo contiene listas bilingües de denominaciones. Por ejemplo, no suele aportar información de tipo regional, además el vocabulario se encuentra compartimentado (los términos están aislados, sin indicar su relación con otros términos del campo) y descontextualizado (no incluye información sobre uso en contexto), y los equivalentes proporcionados en ocasiones son erróneos.

Por ejemplo los diccionarios bilingües donde solo tienes la equivalencia me parece que sirven de muy poco, para empezar porque no te permiten evaluar no solo la calidad de la equivalencia sino la correspondencia con lo que tú necesitas, puede que la correspondencia sea exacta, pero que el contexto, o el registro no sea el que tú estás buscando. Entonces, un diccionario solo con equivalencias inglés-español, por ejemplo, a mí me parece que no va muy lejos. (entrevista\_039: descontento\_diccionario\_bilingüe; solo\_terminología\_bilingüe)

Qué pasa cuando haces una búsqueda español-inglés o inglés-español, que te trae distintos resultados y de pronto tú lo aplicas pero no es la palabra que se usa en tu país o en el país al cual va a ir el texto, porque de pronto yo he traducido al español textos que iban a otros países de Latinoamérica, entonces tal vez no se entendían. Un defecto que presenta, por ejemplo de ese tipo, es el diccionario de Marina Orellana porque de pronto tiene una palabra como *copote*, que no es la que usamos nosotros aquí en la Argentina. Entonces, si tú entendiste de qué se trata, perfecto, lo que tienes que hacer luego es la adaptación, pero claro ya sería un diccionario que implica un trabajo de mayor envergadura, mucho más complejo, interdisciplinario, con gente de todas las academias, si se quiere, porque en nuestro caso existen muchas variedades regionales, entonces ya es más difícil. (entrevista\_021: descontento\_diccionario\_bilingüe; no\_contiene\_información\_regional)

No me gustan los diccionarios bilingües porque el diccionario bilingüe parte de, no sé, de premisas de que todo está muy estancado y muy compartimentado, que todo está metido en cajones y a veces el texto no es así, es decir, las personas que escriben en inglés no escriben teniendo en mente un diccionario bilingüe, tienen en mente lo que es su idioma. Entonces creo que hay que buscar en las fuentes de ese idioma. (entrevista\_004: descontento\_diccionario\_bilingüe; terminología\_compartimentada)

Es que no me fío del todo porque me da muchas posibilidades y entonces, claro, dependiendo del contexto, a lo mejor se puede dejar una palabra u otra, digamos que para mí es menos preciso. (entrevista\_018: descontento\_diccionario\_bilingüe; terminología\_descontextualizada)

Los diccionarios es que normalmente no siempre afinan, entonces... yo qué sé... la posibilidad de contextos es tan amplia, que el diccionario a veces te ayuda y a veces tienes que decir no, eso no, aunque tú lo digas, ¿no? Esa es básicamente mi experiencia con diccionarios. (entrevista\_017: descontento\_diccionarios; descontento\_terminología\_errónea)

Respecto de los diccionarios cuya naturaleza es más bien onomasiológica (**diccionarios de sinónimos, tesauros, diccionarios ideológicos, inversos, redes de ideas**) y también de los **diccionarios de colocaciones**, los traductores indicaron que son demasiado generales y que las relaciones que incluyen son demasiado obvias, que la mayoría están disponibles en papel y la búsqueda es lenta y, en muchas ocasiones, poco satisfactoria. Los traductores revelaron que la búsqueda onomasiológica en estos diccionarios les resulta difícil.

Tesauro no, en particular no. Tengo uno en inglés, pero es que al final no me sirve mucho porque las palabras son demasiado generales las que salen, pensaba que lo iba a usar pero al final muy poco. (entrevista\_014: descontento\_tesauro; descontento\_demasiado\_generales)

No, eso sí que no, eso en el monolingüe o vas a la entrada concreta o estás perdido, a no ser que sea un ideológico o algo, pero es que no sirven tampoco... No me han sido útiles. (entrevista\_012: descontento\_diccionario\_ideológico\_inverso; no\_ofrece\_buenos\_resultados)

Sí pero para la inmediatez que necesita el traductor no siempre lo puede utilizar todo lo que quisiera. El de Redes, por ejemplo, pues no es fácil de entrada utilizarlo, tienes que estar un poquito acostumbrado. (entrevista\_019: descontento\_diccionarios\_colocaciones; descontento\_redes\_ideas; descontento\_búsqueda\_difícil)

Respecto al formato de consulta de los diccionarios, los traductores mostraron su descontento en los **diccionarios electrónicos** por el límite de las licencias y porque todavía muchos ofrecen las mismas búsquedas que el formato papel. El inconveniente señalado de los **diccionarios en línea** es la escasa fiabilidad y que muchos diccionarios en línea están incompletos o tienen muy pocas entradas.

Lo que pasa es que las licencias de algunos diccionarios no te permiten tenerlos en todos los ordenadores que puedas llegar a utilizar. Tú puedes estar con el ordenador de otra persona y entonces no lo tienes. (entrevista\_016: descontento\_diccionario\_electrónico; descontento\_licencias\_limitadas)

En diccionarios bilingües tengo en inglés el Oxford electrónico, pero tengo en papel el Larousse, pero en esto (las búsquedas) prácticamente son igual los dos. No, con el tiempo supongo que sí porque se irá renovando el electrónico, con nuevas versiones y tal, pero en principio el que tengo no, hasta ahora... (entrevista\_020: descontento\_diccionario\_electrónico; descontento\_mismas\_búsquedas\_papel)

No, bueno mi jefe quiere desde hace tiempo que compre uno pero ahora por ejemplo que estamos con los nuevos productos dentales, no existe todavía un diccionario dental específico en español o inglés-español, hay muchos en Internet pero no son fiables. (entrevista\_011: descontento\_diccionario\_online; menos\_fiabilidad)

Lo que pasa es que a veces los diccionarios que aparecen en la red no son tales, en el sentido de que te dice diccionario de tal cosa y de pronto tiene 200 entradas. Eso no es un diccionario, es un glosario en todo caso, o a veces también sucede que hay algunos glosarios o diccionarios que lo único que tienes es una caja de búsqueda. Entonces tú introduces el término, pero en realidad no sabes qué cantidad de entradas contiene la base de datos. Entonces puedes pensar que es inmensa, como no se sabe, y la respuesta va a estar cuando el sistema te da el resultado, que te va a decir que no tiene nada, o de pronto justo aparece una que es la respuesta o te guía a un paso posterior para hacer otra búsqueda. (entrevista\_021: descontento\_diccionario\_online; incompleto)

En cuanto a las **ventajas del diccionario**, los traductores consideraron a los **diccionarios en papel** más fiables generalmente, más completos y más trabajados, además señalaron que existen más diccionarios especializados en papel (razón por la cual todavía existen bastantes diccionarios que los traductores consultan en papel). Además, el uso del diccionario en papel es una excusa para levantarse y apartar la vista de la pantalla del ordenador, es útil cuando se producen problemas de conexión y permite escribir anotaciones al margen.

Sobre todo la actitud, la filosofía, digamos, porque tú si hacías un diccionario de términos jurídicos, se suponía, porque siempre ha habido chapuzas, se suponía que tenías que tener un resultado trabajado, elaborado, documentado, revisado, corregido, etc., un resultado muy perfeccionado. Ahora como Internet te permite ir haciendo eso poco a poco, se tiende a hacer cosas muy de borrador y sin cuidar mucho la calidad, ni la documentación, ni nada, cada uno que va pasando, pues eso, la lógica wiki, va poniendo ahí cosas y hay cosas estupendas y hay cosas que no sirven para nada. Entonces no tienes una autoridad que respalde la obra, que digas, pues esto lo ha hecho la Real Academia de Ciencias, por lo tanto tiene que estar bien, normalmente, en principio estará bien. Para mí es un problema importante. (entrevista\_039: ventaja\_diccionario\_papel; muy\_trabajado; muy\_completo)

Y el papel muchas veces porque son algunos muy específicos de especialidades que no se encuentran electrónicamente. (entrevista\_019: ventaja\_diccionario\_papel; más\_diccionarios\_especializados)

Yo suelo empezar por los libros porque así me levanto y no estoy todo el tiempo sentado. (entrevista\_009: ventaja\_diccionario\_papel; ventaja\_obliga\_levantarse)

Sí, entre otras cosas porque puedes escribir en el margen, puedes consultarlo cuando no estás en línea, o si falla la línea. (entrevista\_008: ventaja\_diccionario\_papel; ventaja\_si\_falla\_red; ventaja\_escribir\_margen)

La principal ventaja del **diccionario electrónico** es la rapidez y la comodidad de uso. Por ejemplo, en los diccionarios multilingües, el cambio en la combinación de idiomas es muy rápido. Además, el diccionario electrónico permite el acceso desde el ordenador e incluso desde el propio procesador de textos, por lo que no se pierde tiempo en la búsqueda. Los traductores también señalaron la posibilidad de copiar y pegar la terminología, útil en palabras muy largas. Asimismo, la letra es más grande y el entorno es más visual. También ofrece la ventaja del transporte y de la cantidad de información que permite incluir, por lo que muchos añaden contextos y también información sobre pronunciación (útil para la traducción de algunas lenguas como el japonés). Además, el formato electrónico facilita que el usuario pueda editar los contenidos (por ejemplo, en la base de datos interna de la UE, *IATE*). También se actualizan más fácilmente.

95% electrónico. Muchísimas, simplemente con la postura ya del traductor que antes siempre andaba pues torcida, la cabeza en el diccionario y los brazos en el teclado y realmente es mucho más cómodo en este sentido, también es más rápido, siempre acabas pasando hojas en el de papel y, en este sentido, es mucho más rápido. (entrevista\_033: ventaja\_diccionario\_electrónico; ventaja\_rapidez; ventaja\_comodidad)



Es más rápido, es más fácil de transportar a cualquier sitio si lo necesitas. (entrevista\_029: ventaja\_diccionario\_electrónico; ventaja\_rapidez; ventaja\_transporte)

El año pasado incorporé a mi PC una herramienta que se llama *Dixio*, que te permite haciendo clic en el botón de Windows y en el mouse, se abre un cuadro de diálogo que te explica de qué se trata la palabra, cuáles son los sinónimos, antónimos y las posibles traducciones, es muy completo. (entrevista\_021: ventaja\_diccionario\_electrónico; acceso\_desde\_procesador)

O copiar y pegar, por ejemplo, a veces tienes un término que puede ser un nombre o un compuesto químico que es así de largo, pues es mucho más fácil resaltarlo, copiar y pegarlo que cómo se escribía eso... suponiendo que tuvieras un diccionario en papel de compuestos químicos, que no todo el mundo lo tiene. (entrevista\_029: ventaja\_diccionario\_electrónico; ventaja\_copiar&pegar)

Por un lado, es menos dañino para la vista porque en un diccionario la letra es chiquitina, estás ahí, la luz. (entrevista\_010: ventaja\_diccionario\_electrónico; ventaja\_letra\_grande)

También porque los diccionarios electrónicos permiten incorporar mucha más información. No están limitados por el espacio y entonces siempre encuentras más información en el diccionario electrónico, o podrías encontrar, que en el de papel. (entrevista\_012: ventaja\_diccionario\_electrónico; muy completo)

Electrónico, sí tengo, pero solamente el de japonés, que está más relacionado porque como son caracteres lo necesito para buscar los caracteres. Pero es más porque es un idioma específico, si no, no. [...] Sí, sí, sí, a nivel de estructura, o sea para los caracteres funcionan igual. Lo bueno de los diccionarios electrónicos, cuando tú ves un carácter, si no lo conoces ni lo sabes pronunciar. Así que primero tienes que buscar cómo se pronuncia para acceder a la palabra, por eso este paso adicional que lo permite hacer el diccionario electrónico, que ahora por Internet eso todavía no se puede hacer. (entrevista\_014: ventaja\_diccionario\_electrónico; ventaja\_información\_pronunciación)

Trabajamos directamente con la interfaz de *IATE* interna, es bastante diferente a la interfaz de la versión pública y sobre todo la principal diferencia es que puedes editar, que en la interfaz pública no puedes editar. (entrevista\_024: ventaja\_base\_de\_datos\_IATE\_interna; puede\_editarse)

Sí, sobre todo porque además permite actualizarlo mucho más, eso, que si todos los años tienes que ir y comprar un diccionario nuevo... (entrevista\_013: ventaja\_diccionario\_electrónico; ventaja\_actualización)

El medio electrónico ofrece, asimismo, más opciones de búsqueda, los traductores mencionaron la función de autocompletar, autosugerencias, búsqueda de palabras truncadas, combinación de palabras de búsqueda, búsqueda booleana, búsqueda dentro de las definiciones y enlaces a palabras relacionadas. Sin embargo, respecto a la amplitud de opciones de búsqueda que ofrece el diccionario, los traductores indicaron que existe la posibilidad de que se incluyan demasiadas opciones de búsqueda, que pueden resultar poco útiles, por ejemplo la opción de búsqueda de anagramas.

Sí, la función de completar palabras, si no estás seguro de cómo se escribe o ni siquiera sabes cómo es, es muy útil que los diccionarios tienen la función de completar palabras, si tú pones 4 o 5 letras, te ponen el resto y así ya tienes una pista más. [...] Sí, sobre todo en cosas médicas, a veces en textos económicos que no estás seguro y así enseguida tienes la palabra entera y con la definición. Eso está bien. (entrevista\_009: ventaja\_diccionario\_electrónico; ventaja\_autocompletar)

Ha mejorado porque ahora también salen términos parecidos a lo que buscas y a veces encuentras una cosa aquí si el término exacto no sale. (entrevista\_030: ventaja\_diccionario\_electrónico; ventaja\_autosugerencias)

El *Brandstetter* te permite por ejemplo en alemán, como hay tantas palabras compuestas, te permite utilizar el asterisco y buscar solo la segunda parte de la palabra, de manera que si no encuentras, o no está la que tú buscas sí puedes ver palabras similares con parte del término igual. (entrevista\_033: ventaja\_diccionario\_electrónico; ventaja\_palabra\_truncada)

Las búsquedas booleanas van muy bien, el poder buscar, por ejemplo en el *SuperLex* puedes buscar dos términos por ejemplo, puedes buscar en la definición del término. Si buscas términos en un contexto muy concreto puedes buscarlos directamente en el diccionario electrónico, entonces los resultados que te arroja son términos en los que tú ya sabes que muchos no te valen, pero tú vas a los dos o tres que sabes que sí que te van a valer y te resalta dentro de la propia entrada, te resalta las palabras que tú has buscado. (entrevista\_022: ventaja\_diccionario\_electrónico; ventaja\_búsqueda\_booleana; búsqueda\_a\_partir\_términos\_definición)

It's quicker, and sometimes you can do cross-references more easily than on paper, to jump from one word to the other. (entrevista\_047: ventaja\_diccionario\_electrónico; palabras\_relacionadas)

No es mi caso, pero el diccionario de la Real Academia te da la opción, creo, de buscar anagramas y ese tipo de cosas, concretamente te digo no me sirve de mucho a mí, son posibilidades que para mí no son útiles. (entrevistas\_029: opciones\_poco\_útiles; anagramas)

#### 5.2.3.4.1.2 Ventajas y desventajas observadas en los programas de traducción asistida

Los traductores también indicaron **desventajas de los programas de Traducción Asistida por Ordenador (TAO)**, una de ellas es la resistencia generacional de algunos traductores más mayores a las nuevas tecnologías, pero también el precio elevado de las licencias, el control del cliente sobre el traductor (sobre todo desde la aparición de programas TAO en línea, donde el cliente sabe cuándo el traductor se conecta y se desconecta), la perpetuidad de errores que existen en la memoria, las incoherencias terminológicas entre el glosario y la memoria que pueden producirse, la escasa fiabilidad y extensión de algunas memorias de traducción, la compartimentación de la traducción y el hecho de que no necesariamente agiliza el proceso de traducción, pero el cliente sí quiere pagar menos.

Yo, será porque soy vieja, pero soy bastante remisa a muchas de las herramientas de traducción asistida. (entrevista\_048: descontento\_TAO; resistencia\_generacional)

Lo que sí utilizo ahora, cada vez nos lo piden más, es *Across*, es un programa que permite al cliente gestionar mucho mejor las traducciones, trabajas online, es decir, el cliente ve cuándo has abierto el archivo, ve hasta qué punto tienes el archivo trabajado, ve cuándo lo has finalizado, lo envías directamente a un servidor y cada vez nos lo piden más, pero no me gusta nada. Te tienen demasiado controlada, luego el cliente en este caso es quien introduce la terminología en el diccionario. La manera como lo trabajamos nosotros es que no nos dejan hacer prácticamente nada más que la traducción y no estoy muy de acuerdo con este sistema. (entrevista\_033: descontento\_TAO; control\_cliente)

En ocasiones, en mi opinión, entorpece muchas veces, yo trabajo muchas veces como *proofreader* y el traductor suele fiarse de la terminología de la TM que a veces no es muy correcta y se le dan muchos errores, se mantiene allí un foco de errores heredados importante, en el aspecto terminológico. (entrevista\_038: descontento\_TAO; descontento\_persisten\_errores)

Si te digo la verdad, a veces son más un engorro que una ayuda porque si no concuerdan con la memoria de traducción ya no sabes cuál tiene prioridad, pero bueno básicamente es una regla más que hay que cumplir en el proyecto, más que una ayuda. (entrevista\_042: descontento\_TAO; descontento\_incoherencia\_glosario\_memoria)

Eso en cuanto a memorias de traducción que proceden de otros traductores porque, por un lado, si son de otros traductores puedo ahorrarme búsquedas pero también es verdad que no puedo pensar que son tan fiables como si lo hubiera hecho yo, así que a veces lo que va ahorrando por un lado, si es de otro traductor, hay veces que sí que hay que ir confirmando, si no te suena de nada el término. (entrevista\_037: descontento\_TAO; menos\_fiabilidad)

¿Sabes lo que pasa? Que normalmente nuestros clientes no tienen un corpus tan grande, tan grande, tan grande, como para que eso sea muy útil. Ahora sí que hemos cogido un cliente que nos ha hecho un contrato por 2 años y que sí que estamos creando un corpus más grande en el que sí estamos haciendo un glosario, guía de estilo y todo eso de este mismo cliente, pero normalmente nuestros clientes son PYMES, que generalmente se puede buscar simplemente en el encargo anterior o en dos o tres encargos anteriores, si ha salido alguna vez ya, pero generalmente no hay mucha terminología difícil que se repita... ¡por desgracia! porque podríamos ahí acumular valor, pero no la hay. (entrevista\_016: descontento\_TAO; descontento\_poco\_extensos)

Y además otra cosa muy personal, a mí personalmente me gusta traducir entonces a mí la memoria de traducción me parece que me coarta, porque me ofrece soluciones muy troceadas y tengo que traducir frase por frase, pierdo el contexto, es difícil ver la frase anterior y la siguiente, me parece que es una traducción muy encorsetada. (entrevista\_039: descontento\_TAO; traducción\_compartimentada)

Y también el programa que se ha hecho más popular no necesariamente es el mejor, pero el que se ha hecho más popular es el *Trados*, y el *Trados* es carísimo. Entonces, a mí no me sale rentable porque además trabajar con TAO no me reduce el tiempo de traducción que es lo que me aumentaría la productividad, tengo que revisar después todo porque los sistemas de traducción te ayudan, pero no te quitan, y en cambio el cliente sí que pretende que te cuesta menos tiempo y te paga la mitad, con lo cual no es en absoluto rentable. Entonces, bueno, cuando no hay más remedio, pero hasta ahora me he defendido... (entrevista\_026: descontento\_TAO; precio; descontento\_no\_agiliza\_traducción)

No obstante, los traductores también señalaron **ventajas de los programas TAO**, como que garantizan la coherencia terminológica y que agilizan las tareas repetitivas. Además, permiten clasificar los textos por áreas temáticas y la terminología se encuentra contextualizada.

El *Trados* muchísimo para ciertos tipos de traducción, obviamente, por ejemplo me resulta muy útil porque muchos de los textos se reciclan, es una página donde hay secciones con títulos, subtítulos, donde es fundamental trabajar con *Trados*. Me creo mi glosario y hay luego muchísimas cosas que se van a repetir. Entonces tú no puedes ponerle a una sección, si has traducido *campeonato del mundo de ralis*, si no lo tienes claro, no puedes poner luego *campeonato mundial de ralis*, o *mundial de ralis*. Entonces la idea es fijar esos términos y que luego me vayan saliendo con el programa de traducción asistida porque no es imposible acordarse de todos. Entonces, el *Trados* lo uso prácticamente a diario. (entrevista\_028: ventaja\_TAO; ventaja\_coherencia\_terminológica)

Luego hay otros textos que para traducirlos con memorias de traducción es maravilloso porque por ejemplo un informe económico de un año para otro hay variaciones, pero también hay muchos análisis que son muy parecidos y el cliente utiliza muchas veces el copiar y pegar y eso a nosotros nos facilita mucho también el trabajo, entonces nos viene muy bien tener memorias de traducción. Las usamos en los clientes principales prácticamente en todos. Y dependiendo del tipo de texto, aunque sea un cliente esporádico, por ejemplo sí que tenemos memorias de traducción para los clientes grandes, específicos para ellos, y luego tenemos memorias de traducción generales, por ejemplo de contratos. Entonces, aunque un cliente esporádico me pida un contrato, yo voy a hacerlo con esa memoria. Si otro cliente me pide un contrato, lo voy a hacer con esa memoria, y así la vamos alimentando, también viene bien. Utilizas mucho

menos porque ahí el copiar y pegar no, pero bueno, fraseología, estructuras, siempre viene bien. (entrevista\_022: ventaja\_TAO; agiliza\_trabajo\_repetitivo; ventaja\_textos\_clasificados\_área\_temática)

#### 5.2.3.4.1.3 Ventajas y desventajas observadas en los corpus

Respecto a las **desventajas de los corpus**, algunos traductores indicaron que es un recurso que no conocen o no dominan y también que las veces que han tratado de utilizarlos no han obtenido resultados satisfactorios. Esta insatisfacción se debe a que los corpus compilados son poco extensos, no suelen contener terminología especializada y no abarcan muchos campos de conocimiento. Los traductores señalaron, además, que no compilan corpus (salvo en contadas excepciones), ya que para la búsqueda de terminología puntual les resulta más rápido y útil la búsqueda en Internet como un corpus. Las **ventajas** que se observaron en los corpus son la fiabilidad de la información, que puede utilizarse para justificar la adopción de una solución de cara al cliente.

Pues la verdad es que no lo sé, porque no tengo costumbre, porque nunca los he usado y nunca me he iniciado. (entrevista\_011: descontento\_corpus; sin\_dominio)

Sobre eso vamos a entrar en una bonita polémica. Los corpus los he intentado utilizar pero nunca me han sido de utilidad porque considero que un corpus tendría que ser muy muy muy extenso para que cuando tú tienes un texto concreto delante puedas encontrar textos equivalentes o similares en los que encontrar información similar. Cuando he buscado en corpus he encontrado cosas interesantes, pero nunca han coincidido con los textos que yo tenía para traducir. Entonces, mi conclusión con respecto a los corpus es que para corpus ya está *Google*. *Google* es el corpus infinito y sabiendo buscar bien encuentras siempre los textos que te sirven realmente como textos paralelos para comparar originales o traducciones. [...] No, no me parece... o se plantea como un proyecto gigantesco de una base de datos impresionante... sería ideal si realmente se diera, pero creo que, o es demasiado complejo y, bueno, sobre todo requeriría una inversión, me imagino, importante hacer un proyecto de ese tipo. Y como no existe, por eso te digo, más vale, yo, a mí me ha servido más en mi práctica profesional aprender a utilizar bien *Google* como un corpus de textos jurídicos sabiendo que hay muchos otros textos pero sabiendo encontrar lo que yo buscaba. (entrevista\_012: descontento\_corpus; no\_ofrece\_buenos\_resultados; descontento\_poco\_extensos; Web as Corpus)

Y prepararme yo los corpus, creo que teniendo Internet y haciendo búsquedas avanzadas puedes recuperar bastante información sin necesidad de extraer esos documentos y guardártelos porque te puede servir para esa vez, pero los sitios se actualizan con bastante frecuencia, entonces me resulta más cómodo ir y hacer la búsqueda en el momento que a lo mejor prepararme todo un corpus para luego a lo mejor tener que actualizarlo regularmente. (entrevista\_024: no\_compilar\_corpus; Web as Corpus)

#### 5.2.3.4.1.4 Ventajas y desventajas observadas en Internet

**Internet** también presenta algunas **desventajas**, desde el punto de vista de los traductores. En primer lugar, el exceso de información, que no se encuentra organizada. La pérdida de tiempo que se produce en la búsqueda por el volumen de resultados y el orden de presentación de los resultados, que no siempre responde a los criterios que desearía el traductor. También, la escasa fiabilidad de algunos contenidos, como por ejemplo de

artículos de la *Wikipedia*. Finalmente, los traductores indicaron la dificultad de organizar y localizar los recursos en línea dentro de la lista de favoritos.

Es como buscar una aguja en un pajar y vas quitando paja de un sitio y de otro y al final no sabes si la aguja que encuentras es la que te vale o no, no sabes si estabas buscando esa aguja u otra, sobre todo porque lo que te comentaba, en Internet hay muchísima información, te ayuda mucho, te facilita mucho la vida, pero encontrar realmente la información fiable que tú necesitas dentro de tantísima información a veces es difícil. (entrevista\_022: descontento\_Internet; menos\_fiabilidad; descontento\_pérdida\_tiempo)

Pues para empezar que el ordenamientos de los resultados no sé a qué criterio responde y dudo mucho que sea un criterio que me interese a mí porque sé que hay criterios comerciales que a mí no me interesan para mi trabajo y pueden interferir y obviamente, como a todo el mundo, otra cosa que me perturba es que en principio no hay ninguna manera de establecer una jerarquización de fuentes, aparece todo revuelto, por eso me gusta *Google Scholar* porque te permite ir a textos más trabajados y de más autoridad, a mí lo de la autoridad en Internet me parece fundamental. (entrevista\_039: descontento\_Internet; descontento\_orden\_resultados; menos\_fiabilidad)

En cuanto a recursos en sí, la lista de favoritos está... hay tanta cosa que al final ni te acuerdas dónde está tal cosa o si lo tenías marcado. (entrevista\_029: descontento\_favoritos)

De las **ventajas de Internet**, y de los recursos en línea, los traductores mencionaron la creación colaborativa lo que posibilita a los usuarios validar la información. También indicaron que es el recurso con más información, actualizada y contextualizada, y que permite el acceso a terminología técnica o poco habitual. Igualmente permite al traductor la clasificación en listas de favoritos.

También hay recursos en línea que son colaborativos, que usuarios como yo podemos añadir traducciones. (entrevista\_041: ventaja\_recursos\_online; ventaja\_creación\_colaborativa)

Los diccionarios que están en Internet, pero no podría trabajar sin *Google*, por ejemplo, porque de verdad necesito contexto actuales, contextos que no sean fijos. (entrevista\_023: ventaja\_Internet; ventaja\_Google; Internet\_se\_actualiza\_constantemente; terminología\_contextualizada)

Y si es un término muy especializado, voy directamente a Internet. (entrevista\_020: ventaja\_Internet; terminología\_poco\_habitual)

Tengo también en favoritos una serie de recursos que según lo que estoy buscando... (entrevista\_008: favoritos\_clasificados)

Asimismo, el recurso más utilizado para la búsqueda en Internet es *Google*, según los traductores principalmente por comodidad y costumbre, pero también porque lo consideran el mejor recurso de búsqueda, en el que mejor funciona la búsqueda de imágenes y el más completo; nombraron herramientas de búsqueda específicas como *Google Imágenes*, *Google Libros* y *Google Académico*.

Sí, *Google* básicamente porque da buenos resultados y es el que utilizo normalmente en casa. También tiene muchísimas herramientas: la búsqueda de imágenes, que da mejores resultados que la de otros buscadores; el *Google Académico*, que si estás traduciendo textos médicos es muy útil, es una característica que yo no he encontrado en otros buscadores; y *Google Books*, que si necesitas por ejemplo un diccionario de un campo muy especializado que en la oficina no lo tienen y no se va a comprar, pues te metes y si el diccionario está escaneado, en ese sentido es una herramienta súper útil. (entrevista\_034: ventaja\_Google; ventaja\_mejor\_recurso; mejor\_búsqueda\_imágenes; Google imágenes; Google Scholar; Google Books)

#### 5.2.3.4.1.5 Ventajas y desventajas observadas en los foros

Respecto a las **desventajas de los foros y las listas de distribución** mencionaron la escasa fiabilidad de las respuestas o las respuestas insatisfactorias, la falta de tiempo y la recepción masiva de correos electrónicos inútiles. Aunque también los traductores indicaron **ventajas**, como que los foros resultan útiles como medio de comunicación con colegas y expertos.

Pero a ver estamos hablando de un recurso que la fiabilidad no siempre es la que debe ser. (entrevista\_006: descontento\_foro; menos fiabilidad)

Alguna vez, pero tampoco he obtenido las respuestas que deseaba. (entrevista\_012: descontento\_foro; descontento\_respuesta\_no\_satisfactoria)

No, porque me da la sensación de que no sé lo que van a tardar en contestar. Entonces, si yo necesito saberlo ahora, acabo antes buscándomelo yo que poniéndolo y a ver si alguien tarda 2 horas o 3 días en contestarme. (entrevista\_013: descontento\_foro; descontento\_pérdida\_tiempo)

Sí, estoy suscrito a dos o tres, aunque alguna de ellas me he borrado porque te llegan 80 e-mails en tres horas. (entrevista\_028: descontento\_listas\_distribución; descontento\_e-mails\_masivos)

#### 5.2.3.4.2 Demandas en los recursos

Para acabar este apartado de resultados, nos referiremos a las **demandas** de los traductores en los recursos.

Los traductores abogaron por el desarrollo de herramientas que permitan la búsqueda simultánea en recursos de distinto tipo, o que unos recursos contengan enlaces a otros recursos. En definitiva, que se apueste por la sinergia entre los recursos. Así, vimos que los traductores empiezan a utilizar herramientas de metabúsqueda, personalizables, que permiten una búsqueda desde el mismo procesador de textos, que clasifican los recursos y restringen la búsqueda por áreas temáticas. Los resultados se presentan con las palabras clave de búsqueda resaltadas y clasificadas por temas o recursos. También apuestan por facilitar el intercambio de recursos o materiales entre usuarios.

Siempre echas de menos algo. En unos sí, en otros no, lo que tiene el uno, no lo tiene el otro, pues... Sí, justo que a lo mejor miras en *Google* y dices, mira, qué bien me vendría aquí esta función que este otro buscador... Entonces, eso sí que... una mayor sinergia entre recursos, pero no me concretes qué a lo mejor porque en unos es uno y en otros es otro, pero sí, esa interrelación entre bases de datos sí. Muchas bases de datos son aisladas, por ejemplo buscas en el *DRAE* y en el *DRAE* se acaba y no puedes hacer una búsqueda relacionada por ejemplo en un diccionario que haya por ahí más técnico, que esté interrelacionado. El aislamiento en algunos casos. El que fueran multiplataforma, el que hubiera algún tipo de sinergia. Sí, básicamente es eso. (entrevista\_019: todo\_en\_uno)

En el *Yahoo groups* mío yo publiqué una serie de glosarios especiales. Justamente al preparar la base de datos de 650 sitios, yo me encontré con mucho material que estaba disponible en Internet y que otros podían aprovechar. (entrevista\_021: compartir\_material)

En general, los traductores demandaron recursos en formato electrónico y en línea, especializados, en lenguas en las que actualmente escasean los recursos (por ejemplo, reclamaron que hay una desproporción entre la cantidad de recursos en inglés respecto de otras lenguas). También requirieron la creación de recursos que no solo incluyan terminología bilingüe, sino también otra información como definiciones, contextos, imágenes, etc.

#### 5.2.3.4.2.1 Demandas en los diccionarios

Así, respecto de los diccionarios, solicitaron que sean exhaustivos y constituyan una fuente de autoridad, y que además aprovechen las ventajas que ofrece el formato electrónico y la disponibilidad en línea.

Entonces, lo que yo echo en falta son buenos diccionarios, no 50.000 diccionarios porque yo tener 50.000 diccionarios en Internet no me sirve para nada. Echo un poco en falta la visión que se tenía antes de las obras publicadas pero adaptadas a Internet. Un buen diccionario, pero con la facilidad que da Internet de actualizar, cambiar, etc. de una manera muy fácil y muy rentable económicamente. (entrevista\_039: demanda\_diccionario\_online)

Estos **diccionarios electrónicos y en línea** deberían ser especializados y abarcar diversos campos, que incluyan terminología poco habitual, por ejemplo terminología técnica, jurídica, médica y militar.

Por un lado, lo que comentábamos, diccionarios jurídicos tanto bilingües como monolingües exhaustivos y electrónicos; no existe todo lo que debería existir. (entrevista\_012: demanda\_diccionario\_electrónico; demanda\_diccionarios\_especializados)

Si yo pudiese entrar en Internet y me falta un término enológico que describa el terreno de la enología, si yo tuviese un portal en Internet, de manera que yo entro y empiezo a buscar, voy por enología, dentro me da a elegir una serie de opciones, cultivo, cata, voy a cultivo, dentro clima, terreno, regiones del mundo, por poner un ejemplo, sí que sería útil, pero eso es un trabajo, eso es como, realmente sería una especie de *Wikipedia*, pero organizado por campos porque en *Wikipedia* sí que encuentras siempre lo que buscas, pero está un poquito mal organizado porque claro es un trabajo de voluntarios, sin ánimo de lucro, pero un diccionario online en el que todos los términos estén divididos por campos y por sectores y por subsectores, sería muy útil. (entrevista\_027: demanda\_diccionario\_online; diccionario\_abarque\_bastantes\_campos)

Asimismo, los **diccionarios bilingües** deberían incluir combinaciones de lenguas menos habituales.

Then sometimes I would look for online dictionaries English-Greek, but there aren't any very good ones, so... (entrevista\_047: diccionarios\_bilingües\_combinaciones\_lenguas\_poco\_habituales)

Es necesario que el diccionario bilingüe incluya otra información, además de la terminología bilingüe. Para ello, se pidió que se tenga en cuenta la perspectiva del traductor. La principal demanda fue que los diccionarios bilingües contengan definiciones (como en el Diccionario médico de Masson), también que incluyan contextos y que

indiquen el campo temático. Otra información que puede resultar útil a los traductores, ya que la incluyen en sus glosarios, es información de uso, notas del traductor, sobre todo si no hay equivalencia, terminología prohibida, categoría gramatical y fuente; también información regional (por ejemplo, uso en distintos países de habla hispana o en distintas zonas de Estados Unidos).

Luego también más recursos con definiciones, quizás la definición para mí es la parte más complicada de un recurso, pero es la que le da más valor añadido, en un diccionario bilingüe en el que te limitas de término a término sin tener más contexto ni una explicación creo que pierdes mucha información y muchas veces no te permite, no tienes la capacidad de saber si realmente significa lo que tú piensas exactamente o no, entonces más definiciones. (entrevista\_024: demanda\_diccionario\_bilingüe; demanda\_diccionarios\_bilingües\_con\_definiciones)

Por área temática, si hubiera un diccionario, eso lo creo útil porque buscar por área temática te reduce bastante el volumen de textos que hay que buscar y tal. (entrevista\_020: demanda\_diccionario; indica\_campo\_temático)

Pues origen, destino y contexto, escribo una definición que haya podido encontrar, o sea definición y contexto, una frase a modo de ejemplo que encuentre yo especialmente clara, ah, y también palabras prohibidas. (entrevista\_038: información\_glosarios\_propios; información\_glosarios\_propios\_definiciones; información\_glosarios\_propios\_contextos; información\_glosarios\_propios\_palabras\_prohibidas)

Ponemos comentarios de los traductores mismos, como por ejemplo si indican que prefieren tal o tal traducción, también de hecho este mes empezamos a utilizar la categoría que se llama part-of-speech, que es para determinar si es un sustantivo, un verbo, un adjetivo, etc. También el contexto, sí, y siempre intentamos poner la frase que tenga más contexto o más explicación para que el traductor cuando vaya a usar el glosario pueda determinar qué palabra usar porque por ejemplo para una determinada palabra en inglés, podemos tener una, dos, tres, cuatro traducciones, esas traducciones van a depender del contexto, por eso tenemos que elegir muy bien el contexto. (entrevista\_023: información\_glosarios\_propios; información\_glosarios\_propios\_contextos; información\_glosarios\_propios\_nota\_traductor; información\_glosarios\_cat\_gram)

Hay una demanda de más **diccionarios con imágenes**, que sean bilingües y especializados, electrónicos o en línea, y de otros idiomas aparte del inglés.

Yo tengo uno en casa que es de Oxford y son multilingües y luego tenemos aquí uno lo que pasa es que este es antiguo y este lo que tiene es las imágenes en inglés y luego viene la traducción. Por ejemplo, cuerpo humano. Hay uno en inglés en la web que está muy bien pero ese es monolingüe (el de Merriam Webster). Ese es maravilloso, es una pasada, ojalá existiera algo así en castellano, si existe por favor avísame. Además es que está muy bien explicado, tiene muchísimos campos, pues es eso. Estos por imágenes son iguales. Por ejemplo, deportes, deportes de agua, el cuerpo humano, astronomía, por diferentes temas y te vas encontrando lo básico, lo bueno del Merriam Webster es que profundiza un poco más que los que suelen estar en papel, pero simplemente por una cuestión técnica, en una página web puedes meter cantidades ingentes de información, en papel, y más en este caso del que tengo yo en casa, que son 4 idiomas, si metes todo lo que tiene este electrónico, tendríamos un tocharro que sería inmanejable. Entonces, lo bueno de la web es eso, si existiera un diccionario así de exhaustivo como el de Merriam Webster y hubiera un equivalente en castellano que pudieran conectarse, sería maravilloso porque muchas cosas sí que las encuentras y además son términos muy específicos. (entrevista\_022: demanda\_diccionario\_electrónico; más\_recursos\_con\_imágenes; diccionarios\_bilingües\_imágenes)



Asimismo, se precisan más **diccionarios de colocaciones y redes de ideas**, también bilingües, especializados y electrónicos.

Yo creo que sobre el campo de la relación de términos, la colocación, y de la organización ideológica de los grupos de palabras, yo creo que ahí está todo por descubrir, creo que está todo por hacer y creo que es necesario, y además creo que hay que orientarlo bien porque no creo que lo que busquemos sea una relación absolutamente detallada de todos los términos que tengan que ver con tal cosa, con un tema o con otro, creo que eso por el volumen de la información se convertiría en un problema al mismo tiempo. Pero creo que sí hay necesidad. Hay necesidad por ejemplo los intérpretes, yo he sido intérprete, y a mí me habría venido muy bien tener grupos de palabras a los que referirme porque yo sabía que iba a interpretar sobre un tema y ya está, entonces podía leer artículos, o leer libros, pero si hubiera tenido una serie de palabras relacionadas, de conceptos, de explicaciones, sobre todo de terminología relacionada, y claramente estructurada, y bilingüe, sería una ayuda magnífica, y lo debe ser. (entrevista\_002: demanda\_diccionarios\_colocaciones\_especializados; demanda\_redes\_ideas\_especializadas; demanda\_redes\_ideas\_bilingües)

Sí, pero normalmente son en expresiones técnicas, porque si son expresiones generales, como nativo sabes ese tipo de cosas, pero si es una cosa de informática y no sabes si se dice una cosa u otra como expresión, sí que sería útil. (entrevista\_030: demanda\_diccionarios\_colocaciones\_especializados)

Los traductores también solicitaron más **diccionarios de acrónimos** en lenguas diferentes del inglés.

Lo que echo de menos son buenos diccionarios de acrónimos en otras lenguas que no sean el inglés. Los hay muy buenos en inglés. Pero para los otros idiomas... Me resultaría muy útil buenos diccionarios si hubiera. Utilizo el Oxford grande que tengo en versión electrónica también. (entrevista\_008: demanda\_diccionarios\_acrónimos\_además\_de\_inglés)

Finalmente, respecto de las **búsquedas**, se sugirió que se incorporen búsquedas sencillas pero avanzadas, diferentes a las que permite el formato papel, por ejemplo búsquedas a partir de palabras truncadas, combinación de dos o más palabras, por proximidad, búsqueda de términos dentro de la definición, que contenga enlaces a terminología relacionada y que incorporen funciones de sugerencias automáticas. Algunas de estas opciones ya se están incorporando en algunos diccionarios, pero todavía no se encuentran completamente desarrolladas.

Para mí sería cuanto más sencillo mejor (entrevista\_007: búsqueda\_sencilla)

Poder buscar con más precisión o más ampliamente, a partir de una definición o unos términos de la definición de algo y tener buenas pistas hacia cuál es el término en sí, desde *OneLook Reverse Dictionary*, pero no es muy fiable, pero estaría bien por campos, me parece que podría ser muy útil siempre que funcione bien y que abarque bastantes campos también. (entrevista\_029: demanda\_diccionario; búsqueda\_a\_partir\_términos\_definición; búsqueda\_onomasiológica\_términos\_especializados)

También se añade la idea de que los diccionarios tengan un precio razonable.

Y yo creo que si se le pone un precio, como el que se le puso al Oxford, que era un precio razonable, creo que valía 97 euros o ciento y algo, la gente lo compra porque es mucho más cómodo. (entrevista\_026: precio\_razonable)

#### 5.2.3.4.2.2 Demandas en los programas de traducción asistida

Respecto de las **demandas** en los **programas TAO** los traductores consideraron importante que las interfaces sean intuitivas, que los módulos estén integrados, que exista interoperabilidad y que permitan la traducción de múltiples formatos.

Pues el *Trados*, encuentro que *MultiTerm* queda muy separado del proyecto, tienes que abrir una nueva ventana, una nueva aplicación. En cambio en *Transit* el diccionario es mucho más visual, no te sabría decir más porque realmente *Trados* también tiene una buena herramienta para gestionar la terminología, pero para mí es más funcional *Transit*. También es más fácil al introducir los términos en la visualización, en este sentido. (entrevista\_033: demanda\_TAO; interfaz\_intuitiva; módulos\_integrados)

Yo lo centraría más en perfeccionar lo que es la gestión terminológica de los distintos programas, unificarlo también, está bien que unos trabajen con *Transit* y otros con *Trados* y otros con *Across*, pero unificarlo de forma que por ejemplo una memoria fuera compatible con todos los programas de traducción asistida. (entrevista\_033: demanda\_TAO; interoperabilidad)

Tiene una interfaz muy amigable, y lo que tiene de bueno es que trabaja con muchos formatos diferentes (entrevista\_021: demanda\_TAO; interfaz\_intuitiva; traduce\_varios\_formatos)

Asimismo, sugirieron que se diseñen herramientas libres, como por ejemplo herramientas de metabúsqueda o programas TAO.

Y quizá en el mundo en general de la traducción habría cosas, como el programa de indexación, una versión que no he visto ninguna que lo hace en software libre, creo que hay... todos queremos todo gratis, claro, pero ya que hay programas para todo, aplicaciones para el iPhone que te hacen la cama y yo qué sé, pues que alguien se dedique a hacer algo un poco más útil, pero bueno, el mundo es así. (entrevista\_029: programas\_metabuscador\_libres)

Quizá buenos programas como *Trados* o como *Déjà-Vu*, pero que sean software libre porque la verdad es que *Trados* es bastante caro y deja bastante que desear y más para el precio que vale, es mi opinión, otro a lo mejor dice, no, es Dios, pues cada uno con lo suyo. (entrevista\_029: programas\_TAO\_libres)

#### 5.2.3.4.2.3 Demandas en los corpus

También se mostró un interés por la creación de grandes **corpus especializados**, para lo cual se aportaron diversas ideas. Por ejemplo, los traductores demandaron sobre todo corpus paralelos en línea (del estilo del buscador *Linguee*) o memorias de traducción de temas especializados, que proporcionen muchas opciones de equivalencias en contexto para escoger. También se mencionó la posibilidad de crear recursos que permitan la búsqueda paralela en libros traducidos, basándose en la filosofía de *Google Books*.

Por otro lado, podríamos hablar de corpus, pero corpus enciclopédicos, por decirlo de alguna manera, no enciclopédicos porque contengan conceptos sino porque sean muy amplios y realmente tengan géneros textuales de todos los tipos que puedan aparecerme en mi trabajo profesional. (entrevista\_012: demanda\_corpus; demanda\_corpus\_grande\_representativo)

Y otra idea que sería bastante útil, pero no la he visto en ningún sitio y supongo que no existe sería poder buscar en libros, como puedes hacer en *Google Books*, cosas así, no todo el libro o lo que sea, pero tienes la opción de buscar un término y aparece en un libro publicado y lo puedes comparar, pero en las dos ediciones, una edición en versión original y en el mismo libro publicado, pero su traducción. Eso no sé cómo se podría hacer, pero por el número del ISBN o una cosa así quizá, porque si se ha publicado el Gran atlas de la química orgánica de no sé qué,

no sé cuántos de Fulano, Mengano y Butano, pues si se ha traducido al catalán por ejemplo, en inglés y catalán, si se pudieran buscar los dos libros, sería una buena manera de aprovechar la faena que ha hecho alguien de traducir ese libro, no necesariamente sería del todo fiable, pero me parece que sería una herramienta bastante útil, pero que yo sepa no hay. Sería un poco como lo que hacen en el *EUR-Lex* que salen los documentos y puedes escoger las traducciones, una idea un poco así, pero para libros publicados, no sé si habría alguna manera de poder hacer eso. Pero si en *Google* tú puedes consultar el libro en *Google Books* o lo que sea, no quiero todo el libro, simplemente que me aparezcan las dos versiones, no quiero bajarme todo el libro, no quiero piratear nada, solo quiero consultar las dos páginas 37 en las dos versiones a ver cómo se ha traducido tal. (entrevista\_029: búsqueda\_paralela\_libros\_traducidos)

#### 5.2.3.4.2.4 Demandas en Internet

En Internet, y en concreto en los **buscadores**, las demandas estuvieron orientadas a diseñar mecanismos que permitan estructurar u organizar mejor las páginas web y las imágenes en línea, con el objetivo de que se pueda realizar una búsqueda más filtrada (como permite, de alguna manera, *Google Scholar*). Por ejemplo, las páginas y las imágenes podrían clasificarse por palabras clave y permitir una búsqueda semántica, tanto de texto como de imágenes en la que se destaquen en los resultados las palabras de búsqueda en contexto. También pidieron que en Internet se pueda buscar a partir de palabras truncadas y por proximidad, como en los corpus.

Y, por último, en *Google* una búsqueda más por áreas temáticas, cosas de ese tipo. [...] Existen otros buscadores, pero siempre que los he utilizado no contienen la base que tiene *Google* y a partir de ahí, aunque te encuentran resultados más concretos, igual no está el más idóneo que sí que está en *Google* pero que es difícil de encontrar. (entrevista\_012: demanda\_Internet; demanda\_Google; demanda\_filtrar\_búsqueda)

Tú en las imágenes puedes buscar... sobre todo te devuelve las que en los pies de foto o en las palabras clave se ha puesto esos... Imagínate si las imágenes tuviesen una descripción un poquito más amplia sería más fácil que te devolviese imágenes asociadas a tu tema. Aunque una foto sea de un vestido y tenga el tipo de brocado que tú estás buscando o el tipo de costura que tú estás buscando, si no está detallado en el pie de foto o en las palabras clave, la foto no te la van a enseñar. Ese tipo de cosas... (entrevista\_013: demanda\_búsqueda\_semántica\_imágenes; demanda\_identificación\_palabras\_clave)

Cuando no tienes ni idea, la situación que me has comentado antes, sabes lo que es pero no tienes un término, no te hablo de en español, sino en inglés, la posibilidad de que tú al buscador de imágenes le pudieras meter, necesito una imagen de una tuerca redonda con tales roscas y dos hierros, por pedir, yo pediría eso, sobre todo para técnica. Bueno, la imagen o el propio buscador, por ejemplo la característica de definir de *Google*, pues algo así, pero que en vez de *define* fuera *explain*, una búsqueda un poquito más refinada [...] Claro, más que palabras sería una explicación semántica más que... o sea porque la búsqueda por palabras clave ya la tienes con *Google* y con los buscadores en general. (entrevista\_034: demanda\_Internet; demanda\_Google; demanda\_búsqueda\_semántica; demanda\_búsqueda\_semántica\_imágenes)

A mí hubo una época que *Google* te dejaba meter interrogaciones y cada interrogación era una letra. Eso era genial, pero ahora ya no te deja. Ahora puedes meter lo que se llama un wildcard, que es un asterisco y eso sustituye una cantidad indeterminada de texto, una palabra o varias. (entrevista\_016: demanda\_Google; ventaja\_palabra\_truncada)

## 5.2.4 Discusión

Los objetivos que nos proponíamos en nuestro estudio eran principalmente tres:

- En primer lugar, comprobar si los traductores tienen necesidades de tipo onomasiológico.

- En segundo lugar, recopilar y clasificar las diferentes estrategias de búsqueda que emplean los traductores profesionales y, dentro de estas, situar a las estrategias de búsqueda de tipo onomasiológico.

- En tercer lugar, identificar qué tipos de recursos consideran más útiles los traductores, sobre todo para facilitarles la búsqueda onomasiológica.

Las entrevistas que realizamos a traductores profesionales nos permitieron extraer resultados con los que pudimos obtener respuestas para cada uno de estos objetivos, y que valoramos a continuación.

Finalmente, a partir de la información proporcionada por los traductores, planteamos una serie de propuestas que pensamos deberían tenerse en cuenta en el diseño de recursos útiles para los traductores, y que faciliten a estos usuarios la realización de búsquedas onomasiológicas.

### 5.2.4.1 Discusión sobre las necesidades onomasiológicas de los traductores

Respecto al primer objetivo, concluimos que los traductores tienen **necesidades de tipo onomasiológico**, en mayor o menor medida. Esta necesidad de tipo onomasiológico surge en la fase de expresión de una traducción, es decir, son necesidades que se plantean en la LM, bien cuando el traductor conoce un concepto o significado en LO y desconoce una denominación o equivalente para designar dicho significado en LM, bien cuando el traductor desconoce la fraseología o las colocaciones típicas que debería utilizar para expresar un significado en una traducción para que el texto suene natural.

Las necesidades onomasiológicas son menores cuando se traducen textos generales o textos donde la terminología no plantea problemas porque el traductor domina el tema, la terminología y la fraseología de un campo de conocimiento. En cambio, las necesidades onomasiológicas aumentan cuando el traductor realiza una traducción para un cliente nuevo, de un área temática nueva, con terminología poco habitual (en la traducción técnica, médica y a veces en la jurídica), por ejemplo, es común la necesidad de acceder a denominaciones desconocidas de piezas técnicas o de síntomas de una enfermedad; el traductor en muchos casos entiende el significado, para qué sirve o cómo funciona una pieza, pero le falta el término exacto en la LM.

También los propios traductores han reconocido que la necesidad onomasiológica surge con mayor frecuencia en áreas en las que existen pocos recursos bilingües especializados. Por ejemplo, cuando no han podido localizar el término que desconocen en un diccionario bilingüe. No obstante, también mencionan que, aunque encuentren un candidato a equivalente en el diccionario bilingüe, no lo toman como la solución definitiva, sino que realizan búsquedas y comprobaciones adicionales de tipo conceptual en otros recursos.

#### 5.2.4.2 Discusión sobre las estrategias de búsqueda de los traductores y las estrategias de búsqueda onomasiológica

Centrándonos en las estrategias de búsqueda que han mencionado los traductores, hemos observado una serie de tendencias o preferencias en cuanto a la realización de búsquedas, lo que nos ha permitido ordenar las estrategias de búsqueda de los traductores y, dentro de esta clasificación, situar a las estrategias de búsqueda de tipo onomasiológico.

Para la sistematización de las estrategias de búsqueda hemos contrastado la propuesta preliminar de clasificación de estrategias de búsqueda en traducción que habíamos generado a partir del análisis de los resultados de estudios anteriores (y que presentábamos en el apartado 4.5) con los resultados sobre estrategias de búsqueda que realizan los traductores profesionales de nuestra entrevista (apartado 5.2.3.3).

Con ello, hemos verificado algunas estrategias de búsqueda y hemos añadido otras nuevas. Así, hemos elaborado una **nueva propuesta de clasificación de estrategias de búsqueda en traducción que incluye estrategias de búsqueda onomasiológica**. Esta clasificación la explicamos y valoramos a continuación e incluimos un resumen de la misma en el apartado 5.2.4.2.5.

##### 5.2.4.2.1 Discusión sobre las estrategias de búsqueda para la resolución de problemas de recepción (LO.a)

Comenzamos con las estrategias de búsqueda en **LO de recepción**, que hemos codificado<sup>141</sup> como LO.a.

Los traductores realizan consultas en el diccionario monolingüe de denominaciones de la LO para obtener una definición (LO.a.1.1). Además, en un acrónimo, abreviatura o sigla de la LO, el significado puede obtenerse accediendo a la forma desarrollada del

---

<sup>141</sup> La explicación de la codificación ya la hemos explicado en los resultados, en el apartado 5.2.3.3.

acrónimo en un diccionario de acrónimos. Estas estrategias de búsqueda ya se habían mencionado en otros estudios y las verificamos con nuestras entrevistas.

También verificamos que los traductores buscan una denominación (o un acrónimo, abreviatura o sigla) para acceder a su significado en contexto (LO.a.1.2), y aquí hemos visto que la búsqueda la efectúan sobre todo en Internet, a través de un buscador, introduciendo directamente la palabra en LO en un buscador, o utilizando opciones específicas de definición, como por ejemplo el filtro *define* de *Google*, también en enciclopedias en línea, como la *Wikipedia*, o en el sitio web del cliente. No obstante, cabe decir que, a pesar de que el análisis de otros estudios del capítulo 4 mostraba que los traductores accedían a menudo a corpus e incluso compilaban corpus en los que buscar el significado de las palabras de la LO, nuestro estudio muestra que los traductores profesionales apenas buscan en corpus, ni compilan corpus, lo que suelen emplear con más frecuencia es la Web como corpus para acceder al significado de las palabras.

Asimismo, los traductores acceden a imágenes para comprender el significado de una denominación (LO.a.1.3), sobre todo en buscadores de imágenes de Internet y en alguna ocasión en un diccionario visual. Sin embargo, parece que muy pocos traductores consultan diccionarios de ideas o tesauros para ver cómo se estructura el significado de un término dentro de un campo (LO.a.1.4).

Del mismo modo, se ha comprobado que los traductores sí que preguntan a especialistas sobre el significado de los términos (LO.a.2). Este es un recurso humano que se emplea cuando los recursos materiales fallan. Finalmente, el traductor pregunta el significado al cliente (LO.a.3) si no obtiene una solución por ninguna de las estrategias anteriores. Las preguntas sobre el significado al cliente son una estrategia que no habíamos recogido en nuestra propuesta preliminar de estrategias de búsqueda y que incorporamos ahora.

Finalmente, respecto a las necesidades de los traductores en LO es importante señalar que, aunque las estrategias de búsqueda de tipo LO son esencialmente estrategias de comprensión y, por ende, tienen una naturaleza de tipo semasiológico (consultar una denominación para acceder a su significado), los traductores han señalado que en ocasiones buscan el significado de un término en la LO para localizar en la definición indicios o pistas que les ayuden a realizar una búsqueda posterior para acceder a una denominación o equivalente que desconocen en LM. Por tanto, estas búsquedas constituyen también una primera fase que puede ayudar al traductor a iniciar

posteriormente una búsqueda de tipo onomasiológico en la LM, como veremos a continuación.

#### 5.2.4.2.2 Discusión sobre las estrategias de búsqueda para la resolución de problemas de transferencia (LM.b)

Las necesidades de los traductores de **transferencia en la LM** aparecen cuando el traductor conoce o comprende el significado en LO pero desconoce una denominación en LM que exprese dicho significado (LM.b). Veremos que, frente a una necesidad de este tipo pueden surgir estrategias de búsqueda de tipo onomasiológico, en las que el traductor realiza una consulta a partir de la información que conoce sobre el significado para acceder o recuperar en un recurso la denominación que expresa dicho significado.

Ante una necesidad LM.b, la primera estrategia que suele emplear un traductor es la búsqueda de terminología bilingüe, es decir, realiza una consulta en un recurso bilingüe a partir del término original para acceder a una denominación equivalente en LM (LM.b.1). En este caso, coincidimos con otros autores en que el traductor puede consultar el término original completo en un diccionario bilingüe o multilingüe (LM.b.1.1). Además, en el caso de términos compuestos, el traductor puede buscar a partir de uno de los componentes del término compuesto (LM.b.1.2). Esta es una estrategia que es posible solo en algunos diccionarios electrónicos, pero que puede resultar muy útil a los traductores, por ejemplo como hemos visto en el caso del alemán, donde los compuestos nominales son muy habituales.

Sin embargo, los traductores profesionales que hemos entrevistado señalan los inconvenientes de las listas de equivalencias entre lenguas, que no siempre representan equivalencias de significado.

Otra estrategia de búsqueda mencionada por los traductores es el acceso a contextos bilingües en un corpus paralelo (LM.b.1.3), el más nombrado es *EUR-Lex*. Asimismo, los traductores aluden a la búsqueda en memorias de traducción como corpus paralelo, los traductores crean sus propias memorias o utilizan las que le proporciona el cliente. Además, indican que en el caso de que el cliente proporcione una memoria, la memoria y el glosario del cliente son el primer recurso donde deben realizar la búsqueda, aunque estas memorias y glosarios no sean siempre de calidad.

La estrategia de búsqueda de páginas web bilingües en Internet hemos comprobado que es bastante frecuente (LM.b.1.4). Los traductores de nuestro estudio suelen utilizar la misma estrategia que proponíamos en la clasificación preliminar. Consiste en introducir la

denominación en LO en un buscador restringiendo los resultados a páginas de la lengua meta o también utilizando palabras clave en LM, como *diccionario* o *glosario*, con el fin de acceder a páginas bilingües o a diccionarios bilingües en línea donde pueda localizarse un equivalente. Los traductores, además, nos han indicado que utilizan otros filtros para limitar la búsqueda, como palabras clave del área temática, filtro de fecha, filtro de formato, filtro de extensión de la página web del cliente y también realizan búsquedas dentro de buscadores específicos, como *Google Scholar*. Esto indica, tal y como también nos han reconocido, que los traductores necesitan restringir las búsquedas en Internet porque sin las opciones de filtrado los resultados que se obtienen son demasiado amplios.

Dentro de la búsqueda bilingüe en Internet, los traductores también señalan la utilidad de buscadores que recopilan y alinean páginas bilingües que encuentra en Internet, entre las cuales se encuentran páginas bilingües de organismos internacionales, por ejemplo documentos extraídos de *EUR-Lex* o del Parlamento Europeo, entre otros. Un ejemplo de este tipo de buscadores es *Linguee*. Por tanto, una de las líneas de investigación para adaptar los recursos al traductor debería estar orientada a mejorar los buscadores para que permitieran la recuperación automática de páginas bilingües de Internet, a modo de corpus paralelo, y que además estas páginas fueran de calidad y pudieran estar restringidas a determinados campos de conocimiento, es decir, que se avanzara en la búsqueda en Internet como corpus bilingüe paralelo y especializado.

Otra estrategia de búsqueda de terminología bilingüe en recursos en línea es la búsqueda de terminología bilingüe en enciclopedias en línea, sobre todo *Wikipedia* (LM.b.1.5). En este sentido, hemos observado dos variantes de búsqueda, en una el traductor consulta a partir de una denominación en LO en *Wikipedia*; este artículo está vinculado a un artículo en LM cuya entrada seguramente se corresponde con el equivalente que el traductor necesita (LM.b.1.5.1). Asimismo, en el caso de que la entrada del artículo vinculado en LM no proporcione el equivalente que el traductor desea, otra posibilidad es leer el artículo completo en LM para tratar de localizar el equivalente en contexto (LM.b.1.5.2). Finalmente, veíamos que la estrategia de tipo LM.b.1.5 está relacionada con la estrategia LM.b.2, que trataremos a continuación, porque cuando no existe un artículo vinculado en la LM que necesita el traductor, es posible que estén disponibles artículos en otras lenguas, que también pueden ayudar al traductor a decidir o realizar otras búsquedas como puente para acceder al equivalente en la lengua que precisa.



El uso por parte de los traductores de enciclopedias en línea de tipo colaborativo, como es el caso de *Wikipedia* muestra la relevancia de los avances de la Web 2.0 para las búsquedas del traductor.

Igualmente, los traductores utilizan estrategias de búsqueda de terminología bilingüe en metabuscadores (LM.b.1.6). Esta estrategia es frecuente en traductores de organismos internacionales. Por ejemplo en la UE, los traductores utilizan el motor de búsqueda *Quest* que permite la búsqueda simultánea en distintos recursos de la UE y de otros organismos, en la base de datos *IATE*, en la legislación *EUR-Lex*, en una memoria de traducción llamada *EURAMIS*, etc., además este metabuscador puede consultarse desde el procesador de textos y permite limitar la búsqueda a los recursos preseleccionados por el traductor. Otro ejemplo lo encontramos en la ONU, donde utilizan la herramienta de metabúsqueda *dtSearch*, que también da acceso a todos los textos traducidos por este organismo. Asimismo, algunos traductores autónomos utilizan *dtSearch*, con el que indexan y organizan por campos sus glosarios, documentos traducidos, otros textos que puedan haber descargado de Internet o incluso, dependiendo de la versión y configuración del programa, es posible realizar la búsqueda también en sitios en línea. Otros traductores instalan complementos en el navegador que les permiten la selección de una serie de recursos en línea donde realizar las búsquedas. En este sentido, deducimos que los traductores dan importancia a la búsqueda simultánea en recursos de distinto tipo.

La última estrategia de búsqueda de terminología bilingüe es la utilización de terceros idiomas como puente entre la LO y la LM (LM.b.2). Esta estrategia funciona con idiomas cercanos, por ejemplo, entre las lenguas romances. Así, hemos visto que si el traductor no encuentra un equivalente inglés-español, pero encuentra un equivalente inglés-francés, el equivalente en francés le puede servir de puente para acceder al equivalente en español.

A su vez, hemos visto que el latín es un idioma que puede actuar como puente entre dos lenguas con denominaciones que tienen una correspondencia latina, como los nombres de plantas, animales, enfermedades, etc. En este caso, el traductor consulta en Internet a partir de la denominación en LO pero utilizando como restricción la palabra clave *latín* y, una vez ha accedido a un equivalente en latín, realiza una segunda consulta en la que introduce la denominación latina y restringe la búsqueda a páginas en LM, con lo que accede a un equivalente en LM.

La estrategia de tipo LM.b.2 puede ser útil cuando en la combinación de idiomas del traductor no hay suficientes recursos bilingües disponibles. Esta estrategia puede venir

sucedida de una estrategia de tipo LM.b.3.1 porque a partir del acceso a un equivalente en un idioma similar al de la LM, el traductor por analogía puede intuir cuál podría ser el equivalente en la LM y comprobarlo mediante una nueva búsqueda.

En definitiva, en la necesidad de transferencia en LM, las estrategias de búsqueda de terminología bilingüe (LM.b.1 y sus variantes y LM.b.2), que acabamos de ver, suelen ser la primera opción de búsqueda del traductor. La búsqueda de terminología bilingüe es una búsqueda (denominación-denominación), porque va de la consulta de una denominación (en LO) a la obtención de una denominación (en LM). En este sentido, pensamos que con este tipo de búsquedas los traductores evitan la búsqueda onomasiológica (significado-denominación).

Los traductores eligen la búsqueda en recursos bilingües porque es más directa y rápida. No obstante, a menudo esta búsqueda denominación-denominación no ofrece resultados, sobre todo con la terminología más especializada, donde los recursos bilingües escasean. Asimismo, los traductores han confirmado que no confían en la terminología que ofrecen los diccionarios bilingües, sobre todo si está descontextualizada. Es entonces cuando el traductor tiene que utilizar otras estrategias de búsqueda del tipo significado-denominación.

Una estrategia de búsqueda que emplea el traductor cuando falla la terminología bilingüe o para complementar esta búsqueda, es una estrategia de búsqueda por intuición (LM.b.3), es decir, el traductor trata de intuir un equivalente exacto en LM (LM.b.3.1) o al menos un componente del equivalente, por ejemplo si es un término compuesto, intuye una de sus partes (LM.b.3.2), y consulta esta intuición para comprobarla bien en un diccionario monolingüe, en un corpus, y sobre todo en Internet. Por ejemplo, se utiliza la estrategia de introducir una expresión incompleta entre comillas en *Google* sustituyendo con asteriscos aquellas palabras que el traductor desconoce.

También hemos visto que el traductor combina la consulta de la denominación en LO junto con una propuesta de equivalente en LM (LM.b.1&3).

Ahora bien, como hemos venido avanzando, si la consulta a partir de la denominación original en recurso bilingüe o la intuición del traductor no ofrecen resultados, el traductor puede verse bloqueado y entonces necesita poner en marcha otras estrategias de búsqueda más avanzadas, que partan del significado para llegar a la denominación.

Hemos visto que los traductores mencionan la búsqueda a partir de palabras relacionadas en contexto (LM.b.4). Esta estrategia consiste en que cuando el traductor

desconoce una denominación en LM para expresar un concepto de la LO, realiza la consulta a partir de información relacionada con la denominación desconocida y que, por tanto, aparecerá en el mismo contexto que ella; una vez el traductor accede a este contexto de referencia será capaz de localizar en él la denominación que desconoce.

En la bibliografía ya se sugerían este tipo de estrategias de búsqueda en las que para acceder a denominaciones desconocidas el traductor investiga otras denominaciones que sí que conoce y que deberán aparecer en un contexto próximo a aquellas que desconoce (véanse a modo de ejemplo Bowker 2000: 37-45; Sánchez-Gijón 2004a: 212-215; Wilkinson 2007: 113-120). Veíamos algunas palabras que pueden guardar una relación de proximidad con denominaciones desconocidas, por ejemplo en Bowker (2000: 37-45) se aportan ejemplos de búsqueda a partir de números, abreviaturas en inglés de uso internacional y nombres propios.

Nuestro estudio sobre estrategias de búsqueda onomasiológica en traducción nos ha aportado más datos sobre las búsquedas onomasiológicas de palabras relacionadas en contexto. Los traductores se han referido a dos tipos de relaciones que facilitan el acceso a los contextos útiles: las relaciones sintagmáticas (LM.b.4.1) y las relaciones semánticas (LM.b.4.2).

En cuanto a las estrategias de tipo LM.b.4.1, hemos comprobado que los traductores buscan a partir de relaciones sintagmáticas o colocaciones en Internet, en corpus o memorias de traducción y también en diccionarios de colocaciones, aunque en estos últimos con menos frecuencia debido a su escaso desarrollo sobre todo en lenguas diferentes al inglés. Por ejemplo, si el traductor desconoce un sustantivo, pero conoce algunos verbos típicos con los que podría combinarse o algún adjetivo típico, busca a partir de estos verbos o adjetivos para localizar en el contexto sustantivos que se combinen con ellos.

Asimismo, los traductores buscan a partir de relaciones semánticas sobre todo de sinonimia (LM.b.4.2.1), hiperonimia (LM.b.4.2.2), metonimia (LM.b.4.2.3), también a partir de las características del concepto (LM.b.4.2.4), o del área temática (LM.b.4.2.5). Los contextos en los que aparecen estas relaciones se recuperan en Internet, en sitios relacionados, en la página web del cliente, pero también en diccionarios de sinónimos y en diccionarios o redes de ideas. Las redes de ideas parecen resultar útiles a los intérpretes, pero no tanto como se entienden actualmente los diccionarios de ideas, sino más bien como listas de terminología bilingüe junto con sus relaciones organizadas por áreas especializadas de conocimiento.

Además de los contextos, las imágenes pueden resultar útiles al traductor buscando a partir de palabras clave que ayudan a recuperar una imagen donde aparece identificada la denominación que el traductor desconoce (LM.b.4.3).

De todo lo anterior deducimos que los recursos onomasiológicos deberían avanzar en la línea de facilitar a los traductores el acceso a denominaciones desconocidas a partir del significado de los términos, pero teniendo en mente una concepción más amplia de significado<sup>142</sup>.

El significado de una denominación no puede basarse únicamente en el significado denotativo o conceptual de la palabra, que es el que tiene una palabra fuera de cualquier contexto, y que es el que suele recoger el diccionario. Parte del significado de una unidad es su prosodia semántica, es decir, las unidades léxicas se ven influidas por las connotaciones que transmiten las coocurrencias habituales que acompañan a una palabra (significado pragmático o funcional)<sup>143</sup>. Por tanto, pensamos que un diccionario onomasiológico no solo debería incluir información conceptual de la terminología, sino también información sintagmática.

Nuestro estudio con traductores, como usuarios de un posible diccionario que permita la búsqueda onomasiológica, apoya la idea de Siepmann (2006: 34) de que en el diccionario la unidad básica de significado en el lenguaje no debería ser la palabra y, por tanto, los diccionarios para traductores no deberían ser listas alfabéticas de equivalentes de palabra a palabra, que no incluyan ninguna o poca información del comportamiento sintagmático de las unidades de una lengua. Los traductores, cuya unidad básica de trabajo es el texto, están acostumbrados y necesitan acceder a sus equivalentes en contexto. Por ello, un diccionario onomasiológico para traductores debería permitir acceder a las colocaciones de las palabras y además los traductores deberían poder acceder a estas colocaciones no solo de forma alfabética, sino por campos semánticos.

Asimismo, hemos visto que si todas las estrategias anteriores no ofrecen resultados, la única posibilidad que le queda al traductor es leer textos paralelos (LM.b.5) hasta que localiza la denominación que desconoce; como indican los traductores, a veces hay que leer más y a veces menos. Desde nuestro punto de vista, esta última estrategia de lectura

---

<sup>142</sup> Otros autores también señalan que la consulta onomasiológica puede entenderse en un sentido amplio (Lew 2011: 240), y ponen como ejemplo de búsqueda onomasiológica la búsqueda de palabras por similitud fonética.

<sup>143</sup> Véase también Sinclair (2004: 174) o Siepmann (2006: 34).

sigue teniendo importancia en la búsqueda del traductor porque la mayoría de recursos todavía no han incorporado mecanismos que faciliten el acceso directo a las denominaciones desconocidas, sino que es el traductor el que tiene que idear formas alternativas para llegar a estas denominaciones por medio de estrategias como las que hemos visto anteriormente. Sin embargo, si estas estrategias no funcionan, la única solución que queda es leer.

Además, los traductores de nuestro estudio se han referido, en ocasiones, a la casualidad como uno de los factores que intervienen en la recuperación de una denominación desconocida (LM.b.6).

Por otro lado, nuestro estudio ha revelado una importancia considerable de los recursos humanos en el acceso a denominaciones desconocidas. La primera estrategia de búsqueda en este sentido es el foro. Normalmente los traductores, en lugar de lanzar una pregunta a un foro, consultan las discusiones del foro en las que han intervenido otros participantes y que están relacionadas con el problema que se le plantea al traductor (LM.b.7.1). En el caso de que la duda no se encuentre ya resuelta, el traductor puede lanzar una pregunta a un foro de traductores (LM.b.7.2) o a un foro de especialistas (LM.b.7.3), también hemos visto foros en los que interactúan ambos tipos de profesionales, como por ejemplo en los foros de terminología médica, en los que intervienen traductores y redactores médicos, así como médicos especialistas.

Otra estrategia es la consulta directa a especialistas en la materia (LM.b.8.1), colegas (LM.b.8.2) y expertos terminólogos o gestores de proyectos (LM.b.8.3). Finalmente, si el traductor no recibe respuesta por las vías anteriores, consulta al cliente (LM.b.9).

Los expertos y colegas suelen ser amigos o familiares del traductor. En algunas grandes empresas, agencias o en organismos internacionales son más frecuentes las redes internas de traductores con los que se puede contactar para resolver determinados problemas. En estas empresas y organismos también es más sencillo acceder a expertos sobre la temática de la traducción, que trabajan en otras divisiones, e incluso algunas cuentan con sus propios gestores de proyectos y especialistas terminólogos que se encargan de resolver este tipo de dudas.

Normalmente la consulta con estas personas se realiza por teléfono, *Skype*, correo electrónico y también por servicios de mensajería instantánea. En los últimos años, también están cobrando popularidad las redes sociales de profesionales, donde los traductores pueden ponerse en contacto con colegas o expertos y recibir una respuesta al instante.

Esto nos ha llevado a plantearnos la importancia de los recursos humanos en las estrategias de búsqueda de los traductores. Estudios anteriores no habían profundizado en este tipo de estrategias. Sin embargo, nuestro estudio muestra que la consulta directa a personas es una de las estrategias adoptadas por los traductores cuando los recursos materiales no les aportan soluciones a sus necesidades onomasiológicas. En este caso, a los traductores a veces les resulta más sencillo explicar a un experto el concepto que tienen en mente para que este les proporcione una denominación adecuada. No en vano, Rogers y Ahmad (1998: 195) indican, por ejemplo, que el diccionario ideal para el traductor sería un experto bilingüe con acceso gratuito e ilimitado.

Ello nos ha llevado a considerar las enormes posibilidades de la Web 2.0 o social para poner en contacto a profesionales dentro de los propios recursos de búsqueda. Concretamente, pensamos que sería muy útil, por ejemplo, que en un diccionario el usuario pudiera realizar una búsqueda para acceder a la terminología que desconoce pero también, en el caso de que el diccionario no proporcione una respuesta satisfactoria, que desde el mismo diccionario se facilite al usuario el contacto con especialistas en la materia a los que poder consultar para resolver las dudas. De hecho, en esta tesis ya hemos visto algunos diccionarios que incorporan foros dentro de sus interfaces de búsqueda, aunque se trataría más bien de ampliar esta idea de foro para crear por ejemplo una gran red social de expertos, que no solo tuvieran una participación pasiva en un recurso (por ejemplo, de consulta de un diccionario), sino que los propios usuarios se conviertan en un recurso activo del diccionario.

#### 5.2.4.2.3 Discusión sobre las estrategias de búsqueda para la resolución de problemas de verificación

Las estrategias de búsqueda en LM suelen contar siempre con acciones de **verificación** del equivalente encontrado, bien en el diccionario monolingüe, en un corpus, en Internet o en la página web del cliente. También es habitual la verificación mediante la búsqueda de imágenes.

Las acciones de verificación son estrategias de búsqueda de **recepción en LM** (LM.a), ya que el traductor consulta una denominación (o equivalente en LM) en un recurso monolingüe para acceder a su significado. Con ello, comprueba que el significado del equivalente en LM coincide con el sentido que transmitía el texto en LO. El significado se obtiene accediendo a una definición en un diccionario monolingüe (LM.a.1.1), a un

contexto en un corpus o en Internet (LM.a.1.2), a una imagen (LM.a.1.3), a una estructura de conceptos (LM.a.1.4) o preguntando a un especialista (LM.a.2).

A este respecto, observamos que los traductores muestran cierta desconfianza en los recursos y tienen la necesidad de realizar comprobaciones en recursos diferentes. Por tanto, desde nuestro punto de vista, esto justifica la demanda de herramientas que permitan la búsqueda simultánea en recursos de distinto tipo.

#### 5.2.4.2.4 Discusión sobre las estrategias de búsqueda para la resolución de problemas de producción (LM.c)

Por último, observamos que los traductores, además de encontrar equivalentes satisfactorios en LM, tienen otras necesidades de **producción en LM** (LM.c). Por ejemplo, necesitan conocer la fraseología y colocaciones típicas de un campo para poder expresarse con naturalidad en la traducción. Esta búsqueda de colocaciones puede realizarse en un corpus (LM.c.1.1), pero los traductores acceden a colocaciones sobre todo en Internet (LM.c.1.2), también algunos en el diccionario monolingüe (LM.c.1.3) y en el diccionario de colocaciones (LM.c.1.4).

En la bibliografía hemos visto estudios que tratan de demostrar la utilidad para los usuarios de algunos diccionarios de colocaciones que se están desarrollando (Jousse et al. 2011). Sin embargo, en general en los estudios a los que hemos tenido acceso sobre estrategias de búsqueda en traducción no se han descrito estrategias de búsqueda de los traductores en diccionarios de colocaciones.

En nuestro estudio hemos observado que este tipo de recursos están cobrando interés entre los traductores, que ya utilizan algunos diccionarios de colocaciones, aunque también detectan carencias en los diccionarios de colocaciones que existen hasta el momento. Por ejemplo, reclaman diccionarios electrónicos bilingües especializados de colocaciones.

De ahí, que consideremos importante que los recursos faciliten a los traductores el acceso a los contextos y a las colocaciones. En este sentido, nos resultan interesantes los esfuerzos que se están realizando para la creación de diccionarios que faciliten el acceso onomasiológico a las colocaciones (Jousse et al. 2011; L'Homme y Leroyer 2009: 278; Leroyer 2006: 191; Verlinde et al. 2010: 6).

#### 5.2.4.2.5 Clasificación de estrategias de búsqueda en traducción donde se incluyen las estrategias de búsqueda onomasiológica detectadas

A continuación mostramos una tabla con el resumen de nuestra propuesta de clasificación de estrategias de búsqueda en traducción en la que se incluyen las estrategias de búsqueda onomasiológica detectadas en nuestro estudio. Consideramos que en traducción las necesidades de expresión en LM son de naturaleza onomasiológica. Los traductores utilizan estrategias de búsqueda en las que parten de la consulta a través de elementos del significado de la LO para llegar a un equivalente o denominación en la LM. También buscan fraseología y colocaciones para expresarse en LM.

Esta tabla la hemos obtenido de la comparación de resultados que hemos efectuado de la clasificación preliminar de estrategias de búsqueda generada a partir de los resultados de estudios anteriores (capítulo 4) y de nuestros propios resultados a partir de la realización de entrevistas a traductores profesionales (capítulo 5). Como ya hemos explicado a lo largo de la discusión, nuestro estudio ha verificado estrategias de búsqueda que ya presentábamos en la propuesta preliminar de estrategias de búsqueda, pero también se han observado algunas diferencias y se han añadido nuevas estrategias de búsqueda.

Por ejemplo, en lo que respecta al uso de corpus, mientras que otros estudios mostraban un gran interés por las estrategias de búsqueda en corpus (sobre todo de estudiantes de traducción), nuestro estudio indica que la compilación de corpus *ad hoc* no es una práctica tan habitual entre los profesionales de la traducción y la mayoría tampoco menciona que utilicen con frecuencia corpus ya compilados, aunque sí que todos indican que utilizan la Web como corpus.

Asimismo, de nuestro estudio se deduce la utilidad para el traductor de las herramientas de metabúsqueda, que permiten la búsqueda simultánea en distintos recursos al mismo tiempo; la búsqueda en enciclopedias bilingües en línea de tipo colaborativo como *Wikipedia*; también la utilidad que otorgan los traductores a los recursos humanos, sobre todo expertos, para la consulta de denominaciones desconocidas, y la mención de las ventajas que ofrece la Web 2.0 y las redes sociales para poner en contacto a los profesionales.

Los traductores de nuestro estudio también han mencionado diversas estrategias de búsqueda en contexto a partir de relaciones sintagmáticas y semánticas para acceder a las denominaciones desconocidas. Finalmente, hemos observado un interés de los traductores respecto de recursos que ofrezcan colocaciones.



En la Tabla 37 se sintetizan las estrategias de búsqueda en el orden en el que suelen aparecer en traducción, tal y como hemos explicado a lo largo de esta discusión. En la columna de la izquierda, se hace referencia a la **necesidad o el problema** que motiva la realización de una búsqueda (la información que el traductor conoce para realizar la búsqueda y la información que desconoce o desea obtener con la búsqueda). En la columna de la derecha se resumen las **acciones de búsqueda** que lleva a cabo el traductor (la consulta introducida, el recurso o instrumento de búsqueda y el resultado obtenido).

**Tabla 37. Resumen de nuestra clasificación de estrategias de búsqueda en traducción, que incluye a las estrategias de búsqueda de tipo onomasiológico detectadas.**

NECESIDAD o PROBLEMA - LO o LM a. Recepción b. Transferencia c. Producción * verificación	ACCIONES DE BÚSQUEDA - consulta - instrumento - resultado
<b>LO</b>	
LO.a CONOCE: denominación en LO o sigla/abreviatura/acrónimo > DESCONOCE: significado	LO.a.1 - Consulta de la denominación en LO conocida - en un recurso monolingüe - para obtener el significado desconocido
	LO.a.1.1 - Consulta de la denominación en LO/ o sigla/abreviatura/acrónimo - en un diccionario monolingüe en LO (disponible o no en línea) / o diccionario de acrónimos - para obtener definición / o también forma desarrollada de un acrónimo
	LO.a.1.2 - Consulta de la denominación en LO - en corpus en LO, en Internet, en una enciclopedia en línea, en la página web del cliente - para obtener contextos definatorios en páginas de Internet en LO, en la página web del cliente, en la entrada de una enciclopedia en línea, concordancias o colocaciones del corpus en LO
	LO.a.1.3 - Consulta de la denominación en LO - en un buscador de imágenes o videos en línea o

	<p>diccionario visual en LO</p> <ul style="list-style-type: none"> <li>- para obtener imagen o video</li> </ul>
	<p>LO.a.1.4</p> <ul style="list-style-type: none"> <li>- Consulta de la denominación en LO</li> <li>- en un diccionario específico, ideológico, tesaurus u ontología en LO/ o diccionario de sinónimos</li> <li>- para obtener estructura de conceptos en LO/ o sinónimo</li> </ul>
	<p>LO.a.2</p> <ul style="list-style-type: none"> <li>- Consulta de la denominación en LO conocida para preguntar por su significado</li> <li>- a una persona (hablante nativo o especialista)</li> <li>- para obtener respuesta del significado desconocido</li> </ul>
	<p>LO.a.3</p> <ul style="list-style-type: none"> <li>- Consulta de la denominación en LO conocida para preguntar por su significado</li> <li>- al cliente</li> <li>- para obtener respuesta del significado desconocido</li> </ul>
<b>LM</b>	
<p>LM.b</p> <p>CONOCE: término (denominación+concepto) en LO</p> <p>&gt;</p> <p>DESCONOCE: denominación equivalente en LM</p>	<p>LM.b.1</p> <ul style="list-style-type: none"> <li>- Consulta de la denominación en LO conocida</li> <li>- en un recurso bilingüe</li> <li>- para obtener la denominación equivalente en LM desconocida</li> </ul>
	<p>LM.b.1.1</p> <ul style="list-style-type: none"> <li>- Consulta de la denominación en LO completa (simple o compuesta)/ o sigla/abreviatura/acrónimo</li> <li>- en un diccionario bilingüe/ o diccionario de acrónimos bilingüe, o en los glosarios o archivos propios del traductor o del cliente</li> <li>- para obtener denominación equivalente en LM</li> </ul>
	<p>LM.b.1.2</p> <ul style="list-style-type: none"> <li>- Consulta de un componente de la denominación en LO (compuesta)</li> <li>- en un diccionario bilingüe/ o diccionario de acrónimos bilingüe</li> <li>- para obtener denominación equivalente en LM</li> </ul>
	<p>LM.b.1.3</p> <ul style="list-style-type: none"> <li>- Consulta de la denominación en LO completa (simple o compuesta) o sigla/abreviatura/acrónimo</li> <li>- en un corpus paralelo/ o en una memoria de traducción</li> </ul>

	<p>- para obtener contextos bilingües en los que aparece la denominación equivalente en LM</p> <hr/> <p>LM.b.1.4</p> <p>- Consulta de la denominación en LO completa (simple o compuesta) o sigla/abreviatura/acrónimo Y con filtro de páginas en el idioma de la LM O con alguna palabra clave como <i>diccionario</i> en LM Y con otros filtros (palabras clave área temática, extensión de la página web del cliente, fecha, formato)</p> <p>- en un buscador de Internet general como <i>Google</i>, o en un buscador más específico como <i>Google Scholar</i>, o en un buscador de páginas bilingües paralelas de Internet como <i>Linguee</i></p> <p>- para obtener páginas bilingües de Internet o diccionarios o glosarios bilingües de Internet en los que aparece la denominación equivalente en LM</p> <hr/> <p>LM.b.1.5.1</p> <p>- Consulta de la denominación en LO</p> <p>- en una enciclopedia en línea <i>Wikipedia</i></p> <p>- para obtener acceso al artículo en LO que está vinculado al artículo en LM cuya entrada se corresponde con el equivalente adecuado</p> <p>LM.b.1.5.2</p> <p>- Consulta de la denominación en LO</p> <p>- en una enciclopedia en línea <i>Wikipedia</i></p> <p>- para obtener acceso al artículo en LO y, si la entrada del artículo equivalente en LM no aporta una solución adecuada, el traductor lee dentro del artículo para localizar un equivalente adecuado</p> <hr/> <p>LM.b.1.6</p> <p>- Consulta de la denominación en LO</p> <p>- en una herramienta de metabúsqueda (búsqueda simultánea en diccionarios, corpus, memorias, etc.). Por ejemplo <i>Quest</i> o <i>dtSearch</i>.</p> <p>- para obtener alternativas de equivalente en LM en distintos recursos</p> <hr/> <p>LAS ACCIONES ANTERIORES PUEDEN TENER ACCIONES DE VERIFICACIÓN (véase LM.a.VERIFICACIÓN)</p>
<p>LM.b</p> <p>CONOCE: término en LO</p> <p>&gt;</p> <p>DESCONOCE: denominación equivalente en latín (o en otro idioma distinto a la LM)</p> <p>;</p>	<p>LM.b.2</p> <p>- Consulta a partir de una lengua intermedia como puente entre la LO y la LM (por ejemplo, el latín)</p> <p>- en Internet</p> <p>- para obtener la denominación equivalente en LM desconocida</p>

<p>CONOCE: denominación equivalente en latín encontrado (o en otro idioma)</p> <p>&gt;</p> <p>DESCONOCE: denominación equivalente en LM</p>	
<p>LM.b</p> <p>CONOCE: término original, e intuye una posible denominación equivalente en LM</p> <p>&gt;</p> <p>DESCONOCE: denominación equivalente en LM adecuada</p>	<p>LM.b.3</p> <ul style="list-style-type: none"> <li>- Consulta de la denominación equivalente en LM intuida. Intuición a partir de un conocimiento previo sobre cómo funciona una lengua</li> <li>- en un recurso monolingüe en LM</li> <li>- para seleccionar o no la denominación equivalente en LM</li> </ul>
	<p>LM.b.3.1</p> <ul style="list-style-type: none"> <li>- Consulta de la denominación equivalente completa intuida (simple o compuesta)</li> <li>- en Internet, corpus o diccionario monolingüe en LM</li> <li>- para seleccionar o no el equivalente intuido</li> </ul>
	<p>LM.b.3.2</p> <ul style="list-style-type: none"> <li>- Consulta de un componente intuido (de una denominación equivalente compuesta)</li> <li>- en Internet, corpus o diccionario monolingüe en LM</li> <li>- para seleccionar o no el equivalente intuido</li> </ul>
<p>LM.b</p> <p>CONOCE: término original, e intuye denominación equivalente en LM</p> <p>&gt;</p> <p>DESCONOCE: denominación equivalente en LM</p>	<p>LM.b.1&amp;3</p> <ul style="list-style-type: none"> <li>- Consulta del término original conocido + Consulta de una denominación equivalente en LM intuida</li> <li>- en Internet</li> <li>- para seleccionar o no la denominación equivalente en LM</li> </ul>
	<p>LAS ACCIONES ANTERIORES PUEDEN TENER ACCIONES DE VERIFICACIÓN (véase LM.a.VERIFICACIÓN)</p>
<p>LM.b</p> <p>CONOCE: término original + palabras relacionadas que supuestamente se encuentran en el contexto del equivalente que se desconoce (palabras clave que aparecen en el mismo contexto)</p> <p>&gt;</p> <p>DESCONOCE: denominación equivalente en LM</p>	<p>LM.b.4</p> <ul style="list-style-type: none"> <li>- Consulta de expresiones o palabras clave que pueden aparecer en el contexto del equivalente en LM desconocido porque tienen una relación sintagmática (colocaciones) o semántica con la denominación desconocida</li> <li>- en un recurso monolingüe en LM</li> <li>- para localizar en el contexto la denominación equivalente en LM desconocida</li> </ul>
	<p>LM.b.4.1</p> <ul style="list-style-type: none"> <li>- Consulta de una o más colocaciones, que supuestamente aparecerán en el mismo contexto que la denominación</li> </ul>

	<p>equivalente en LM que desconocemos</p> <ul style="list-style-type: none"> <li>- en un corpus comparable o corpus monolingüe en LM, Internet o diccionario de colocaciones</li> <li>- para obtener un contexto en el que localizar la denominación equivalente en LM desconocida</li> </ul>
	<p>LM.b.4.2.1</p> <ul style="list-style-type: none"> <li>- Consulta de una palabra relacionada semánticamente (sinónimo)</li> <li>- en un diccionario de sinónimos o Internet</li> <li>- para obtener alternativas de equivalente en LM (sinónimos)</li> </ul>
	<p>LM.b.4.2.2</p> <ul style="list-style-type: none"> <li>- Consulta de palabras relacionada semánticamente, (hiperónimo)</li> <li>- en Internet o redes de ideas</li> <li>- para obtener alternativas a equivalente en LM (hipónimos)</li> </ul>
	<p>LM.b.4.2.3</p> <ul style="list-style-type: none"> <li>- Consulta de palabras relacionadas semánticamente, metonimia (por ejemplo un todo)</li> <li>- en Internet o redes de ideas</li> <li>- para obtener alternativas a equivalente en LM (para acceder a una parte)</li> </ul>
	<p>LM.b.4.2.4</p> <ul style="list-style-type: none"> <li>- Consulta de palabras clave que describen el concepto, por ejemplo características</li> <li>- en Internet o redes de ideas</li> <li>- para obtener alternativas a equivalente en LM que aparecen descritos de esa manera</li> </ul>
	<p>LM.b.4.2.5</p> <ul style="list-style-type: none"> <li>- Consulta de palabras clave que indican el área temática</li> <li>- en Internet o redes de ideas</li> <li>- para obtener alternativas a equivalente en LM pertenecientes a esta área temática</li> </ul>
	<p>LM.b.4.3</p> <ul style="list-style-type: none"> <li>- Consulta de palabras clave que describen el concepto</li> <li>- en Internet o buscador de imágenes, o diccionario visual</li> <li>- para obtener imagen en la que aparece identificado el equivalente en LM</li> </ul>
<p>LM.b                  CONOCE: denominación original LO + palabras relacionadas que supuestamente se</p>	<p>LM.b.1&amp;4</p> <ul style="list-style-type: none"> <li>- Consulta del término original + Consulta de expresiones o palabras clave relacionadas que pueden aparecer en el</li> </ul>

<p>encuentran en el contexto del equivalente que se desconoce (palabras clave que aparecen en el mismo contexto)</p> <p>&gt;</p> <p>DESCONOCE: denominación equivalente en LM</p>	<p>contexto del equivalente en LM</p> <ul style="list-style-type: none"> <li>- en Internet</li> <li>- para localizar en contexto el equivalente en LM</li> </ul>
<p>LM.b</p> <p>CONOCE: denominación equivalente en LM intuida + palabras relacionadas que supuestamente se encuentran en el contexto del equivalente que se desconoce (palabras clave que aparecen en el mismo contexto)</p> <p>&gt;</p> <p>DESCONOCE: denominación equivalente en LM</p>	<p>LM.b.3&amp;4</p> <ul style="list-style-type: none"> <li>- Consulta del equivalente en LM intuido + Consulta de expresiones o palabras clave relacionadas que pueden aparecer en el contexto del equivalente en LM</li> <li>- en un corpus en LM o en Internet</li> <li>- para seleccionar o no el equivalente en LM desconocido</li> </ul>
	<p>LAS ACCIONES ANTERIORES PUEDEN TENER ACCIONES DE VERIFICACIÓN (véase LM.a.VERIFICACIÓN)</p>
<p>LM.b</p> <p>Textos paralelos</p>	<p>LM.b.5</p> <ul style="list-style-type: none"> <li>- Consulta/Lectura de textos paralelos</li> <li>- en Internet o en los propios archivos del traductor</li> <li>- para localizar denominación equivalente en LM desconocida</li> </ul>
<p>LM.b</p> <p>(casualidad, serendipity)</p>	<p>LM.b.6</p> <ul style="list-style-type: none"> <li>- Encontrar un equivalente por casualidad (serendipity), por ejemplo realizando otras búsquedas</li> </ul>
<p>LM.b</p> <p>Recursos humanos</p>	<p>LM.b.7</p> <ul style="list-style-type: none"> <li>- Consulta</li> <li>- en un foro</li> <li>- para obtener respuesta con la denominación en LM desconocida</li> </ul>
	<p>LM.b.7.1</p> <ul style="list-style-type: none"> <li>- Consulta de la denominación en LO completa (simple o compuesta) o sigla/abreviatura/acrónimo</li> <li>- base de datos de un foro de Internet de traducción</li> <li>- para obtener discusiones anteriores en las que se ha planteado el problema en un foro y que dan ideas sobre la denominación equivalente en LM</li> </ul>
	<p>LM.b.7.2</p> <ul style="list-style-type: none"> <li>- Pregunta a partir del término en LO, del contexto y de información conceptual que pueda aportar el traductor</li> <li>- en foro de traducción</li> <li>- para obtener respuesta con la denominación en LM</li> </ul>

	<p>desconocida</p> <hr/> <p>LM.b.7.3</p> <ul style="list-style-type: none"> <li>- Pregunta a partir del término en LO, del contexto y de información conceptual que pueda aportar el traductor</li> <li>- en foro de especialistas</li> <li>- para obtener respuesta con la denominación en LM desconocida</li> </ul> <hr/> <p>LM.b.8</p> <ul style="list-style-type: none"> <li>- Consulta directa</li> <li>- a profesionales especialistas</li> <li>- para obtener respuesta con la denominación en LM desconocida</li> </ul> <hr/> <p>LM.b.8.1</p> <ul style="list-style-type: none"> <li>- Consulta del término en LO, y pregunta del término en LM</li> <li>- a especialista</li> <li>- para obtener respuesta con la denominación en LM desconocida</li> </ul> <hr/> <p>LM.b.8.2</p> <ul style="list-style-type: none"> <li>- Consulta del término en LO, y pregunta del término en LM</li> <li>- a un colega traductor</li> <li>- para obtener respuesta con la denominación en LM desconocida</li> </ul>
	<p>LM.b.8.3</p> <ul style="list-style-type: none"> <li>- Consulta del término en LO, y pregunta del término en LM</li> <li>- al gestor de proyectos o grupo de terminólogos responsables.</li> <li>- para obtener respuesta con la denominación en LM desconocida</li> </ul> <hr/> <p>LM.b.9</p> <ul style="list-style-type: none"> <li>- Pregunta</li> <li>- al cliente</li> <li>- para obtener respuesta con la denominación en LM desconocida</li> </ul>
<p>LM.a.VERIFICACIÓN</p> <p>(cualquiera de las necesidades y acciones anteriores LM.b pueden tener necesidad de verificación)</p>	<p>LM.a.VERIFICACIÓN</p> <ul style="list-style-type: none"> <li>- Consulta de una denominación equivalente en LM encontrada (consulta de un equivalente encontrado o consulta de dos o más posibles equivalentes o consulta de dos o más variantes ortográficas)</li> <li>- en un recurso monolingüe en LM</li> </ul>

	<p>- para una verificación o no del equivalente en LM</p> <p>LAS ACCIONES DE VERIFICACIÓN COINCIDEN CON ACCIONES LM.a</p>
<p>LM.a</p> <p>CONOCE: equivalente en LM encontrado</p> <p>&gt;</p> <p>DESCONOCE: significado del término en LM / comprobación de significado / ejemplos de uso en contexto</p>	<p>LM.a.1</p> <p>- Consulta de una denominación equivalente en LM encontrada</p> <p>- en un recurso monolingüe en LM</p> <p>- para obtener el significado o comprobar el significado o ejemplos de uso en contexto</p>
	<p>LM.a.1.1</p> <p>- Consulta de una denominación equivalente en LM encontrada</p> <p>- en un diccionario monolingüe en LM (disponible o no en línea) / o diccionario de acrónimos</p> <p>- para obtener definición / o también forma desarrollada de un acrónimo</p>
	<p>LM.a.1.2</p> <p>- Consulta de una denominación equivalente en LM encontrada</p> <p>- en corpus en LM, en Internet, en una enciclopedia en línea, en la página web del cliente</p> <p>- para obtener contextos definitorios en páginas de Internet en LM, en la página web del cliente, en la entrada de una enciclopedia en línea, concordancias o colocaciones del corpus en LM</p>
	<p>LM.a.1.3</p> <p>- Consulta de una denominación equivalente en LM encontrada</p> <p>- en un buscador de imágenes o videos en línea o diccionario visual en LM</p> <p>- para obtener imagen o video</p>
	<p>LM.a.1.4</p> <p>- Consulta de una denominación equivalente en LM encontrada</p> <p>- en un diccionario específico, tesoro u ontología en LM</p> <p>- para obtener estructura de conceptos en LM</p>
	<p>LM.a.2</p> <p>- Consulta de la denominación equivalente en LM encontrada</p> <p>- a una persona (hablante nativo o especialista)</p> <p>- para obtener el significado desconocido o comprobación del significado respecto del término en LO</p>



LM.c CONOCE: denominación en LM > DESCONOCE: fraseología o colocaciones típicas	LM.c.1 - Consulta de una denominación en LM - en un corpus en LM, en Internet o en un diccionario de colocaciones - para obtener colocaciones que orienten al traductor sobre el uso de la denominación en contexto
	LM.c.1.1 - Consulta de una denominación en LM - en un corpus en LM - para obtener colocaciones que orienten al traductor sobre el uso de la denominación en contexto
	LM.c.1.2 - Consulta de una denominación en LM - en Internet - para obtener colocaciones que orienten al traductor sobre el uso de la denominación en contexto
	LM.c.1.3 - Consulta de una denominación en LM - en un diccionario monolingüe - para obtener colocaciones que orienten al traductor sobre el uso de la denominación en contexto
	LM.c.1.4 - Consulta de una denominación en LM - en un diccionario de colocaciones - para obtener colocaciones que orienten al traductor sobre el uso de la denominación en contexto

### 5.2.4.3 Discusión sobre percepciones y expectativas de los usuarios respecto de los recursos

En cuanto al segundo objetivo, qué es lo que esperan los traductores respecto de los recursos, hemos recopilado una serie de opciones que deberían tenerse en cuenta a la hora de diseñar recursos útiles para los traductores, según las opiniones que han manifestado los traductores de nuestro estudio.

Para empezar, respecto a la **función de los recursos**, dado que el principal objetivo de nuestro estudio era averiguar de qué manera podía adaptarse un recurso para facilitar la búsqueda onomasiológica al traductor, nos gustaría reflejar que un diccionario o un recurso cuya búsqueda fuera exclusivamente onomasiológica creemos que no resultaría completamente útil al traductor. Esto se deduce de que hemos visto que los traductores

necesitan combinar estrategias de distinto tipo, tanto onomasiológicas como semasiológicas.

Las primeras estrategias de búsqueda que pone en marcha, o por las que opta un traductor como primera opción, son de naturaleza semasiológica. Los traductores, profesionales, que se ven en todo momento presionados por el tiempo, con el fin de agilizar la búsqueda, intentan como primera opción una búsqueda de terminología bilingüe, que es una búsqueda denominación-denominación. En este sentido, necesitan diccionarios bilingües y también multilingües, ya que los traductores pueden trabajar con distintas combinaciones de idiomas y hemos visto que pueden ayudarse de las traducciones de otras lenguas que conocen para acceder al equivalente en la lengua que necesitan.

Sin embargo, nuestro estudio evidencia que el traductor precisa también de una búsqueda concepto-denominación o significado-denominación, es decir, de naturaleza onomasiológica, sobre todo cuando la búsqueda denominación-denominación falla, y también para complementarla, ya que hemos visto que la búsqueda directa de una denominación en LO a una denominación en LM, sin tener en cuenta o pasar por el significado, tampoco resulta definitiva para el traductor.

Los recursos exclusivamente bilingües, que solo incluyen listas de denominaciones equivalentes, no aportan la información que el traductor necesita a la hora de seleccionar una denominación equivalente en LM para un concepto o significado de la LO. Por tanto, es imprescindible que los recursos ofrezcan también a los traductores la posibilidad de realizar una búsqueda a partir del significado para llegar a la denominación, y para ello es necesario que los recursos contengan tanto definiciones como información sobre el comportamiento de los términos en contexto. En definitiva, es necesario habilitar opciones en los recursos por las cuales el traductor puede valerse de estrategias de búsqueda de naturaleza onomasiológica.

Asimismo, el **formato ideal de un recurso** o un diccionario destinado a traductores es el formato electrónico, todavía mejor si está disponible en línea. Además, deberán ser diccionarios exhaustivos y completos, que constituyan una fuente de autoridad. Actualmente, la mayoría de los mejores diccionarios son en formato papel.

Además, deberá darse prioridad a la elaboración de diccionarios y **recursos especializados**, que escasean actualmente. Respecto de las **lenguas**, los traductores preferirán recursos multilingües, que contengan otras lenguas aparte del inglés y también sería interesante que recogieran combinaciones de lenguas minoritarias, en las que no existen suficientes recursos.

Es imprescindible que los diccionarios, además de denominaciones equivalentes entre lenguas, contengan definiciones y contextos, así como otra **información adicional** que refleje el significado de los términos. Por tanto, los recursos podrán ofrecer una búsqueda denominación-denominación, pero con verificación del significado a través de definiciones y contextos, así como una búsqueda significado-denominación en contexto.

También en esta línea, los diseñadores de recursos deberían apostar por la elaboración de recursos visuales, que incluyan imágenes, así como recursos que faciliten al traductor el acceso a colocaciones. También sería recomendable que los recursos incluyan una nota del traductor si no hay equivalencia posible y que se tenga en cuenta la información regional.

En cuanto a los **resultados** es conveniente que se destaquen las palabras clave de la búsqueda, que se incluyan resultados con terminología bilingüe, con definiciones y contextos, así se explica la popularidad que han adquirido recursos como *Linguee* que ofrecen a los traductores equivalentes y contextos bilingües extraídos de Internet.

Respecto a las **consultas**, es conveniente que los recursos para traductores sean intuitivos, que presenten una interfaz de búsqueda simple y poco recargada. Parece que los traductores, salvo en contadas situaciones, prefieren una búsqueda directa introduciendo información en una casilla de búsqueda, en lugar de una consulta por navegación en una estructura o en un árbol. La consulta deberá resultar sencilla, por ejemplo el traductor realizará una única consulta y el recurso podrá proporcionar diversos resultados, bien organizados, para que el traductor elija el que le interesa.

Las opciones de un recurso que parecen más interesantes para el traductor son la consulta de palabras exactas, truncadas, consulta de dos o más palabras combinadas y la consulta por proximidad, así como las opciones de autocompletar y autosugerencia.

Asimismo, se recomienda que se permita la consulta simultánea en recursos de distinto tipo. En relación con esto, parece que el desarrollo de recursos de software libre podría permitir una mayor sinergia e interoperabilidad entre los recursos, aspectos que los traductores han considerado relevantes. Además, en el caso de que la herramienta se distribuya con una licencia privativa, los traductores reclaman que no sea demasiado restrictiva.

Dentro de los recursos, se demanda la búsqueda dentro de los campos de definición, la búsqueda en contexto de relaciones (tanto sintagmáticas como semánticas), y la búsqueda semántica de imágenes (clasificadas por palabras clave). En este sentido, se podría avanzar en las búsquedas semánticas en los diccionarios y también en Internet,

valiéndose de los desarrollos que ofrece actualmente tanto la inteligencia artificial como la lingüística de corpus (anotaciones tanto ontológicas como lingüísticas) (véase por ejemplo Aguado de Cea et al. 2003).

Hemos visto que los traductores tratan de utilizar cada vez más Internet a modo de corpus. Nos parece interesante la sugerencia de que al igual que *Google* ya ha incorporado la función *define* donde un usuario introduce una palabra y se obtienen definiciones de la misma (función de búsqueda semasiológica), sería interesante que se trabajara en una función denominada *describe concept* o *explain*, en la cual el usuario introdujera una explicación y se obtuvieran contextos y denominaciones que correspondieran a dicha descripción. Además de las funciones de sugerencia ortográfica, podría trabajarse en funciones de sugerencia conceptual o semántica, aplicado a ofrecer sugerencias o alternativas de búsqueda en diccionarios y en buscadores de Internet.

Por otra parte, también es interesante que los diccionarios permitan el acceso a foros moderados y fiables, de respuesta rápida, o más bien a redes sociales que reúnan a profesionales. En este sentido vemos que el **recurso humano** puede plantearse como un complemento del recurso material e incorporarse en el propio diccionario. El foro puede resultar un buen complemento del diccionario y, aunque los traductores no otorgan la máxima fiabilidad a las respuestas que se ofrecen (tal y como están organizados actualmente) sí que los utilizan para obtener pistas de palabras o términos que después pueden verificar en otros recursos. Los diseñadores también podrán aprovechar las posibilidades que nos ofrece la Web 2.0 para que sean los propios usuarios los que validen los resultados que proporciona un recurso (trabajo colaborativo), pero también permitirán poner en contacto a profesionales de una forma instantánea. Estas son algunas de las aplicaciones profesionales de la Web social que creemos pueden resultar útiles en la elaboración de recursos de búsqueda.

#### 5.2.4.4 Propuestas para la elaboración de recursos orientados a los traductores

Concluiremos esta discusión con una tabla en la que hemos tratado de responder a posibles preguntas que podrían plantearse a la hora de diseñar un recurso orientado al traductor. En función de las estrategias de búsqueda y las demandas propuestas por los traductores profesionales de nuestro estudio, hemos resumido cuáles podrían ser los intereses de los traductores en los recursos en cuanto al formato, la especialidad, las lenguas, la información que debería incluirse, la interfaz de búsqueda, las opciones de búsqueda y la licencia.

**Tabla 38. Propuesta para la elaboración de recursos orientados a los traductores.**

CRITERIOS	PREGUNTA	RESPUESTA
Usuario	¿Quién es el usuario?	- Un traductor
Formato	¿Papel o electrónico?	- Electrónico, mejor si está disponible en línea.
Especialidad	¿General o especializado?	- Especializado
Lenguas	¿Monolingüe, bilingüe o multilingüe?	- Multilingüe
	¿Qué lenguas?	- Se precisan recursos en otras lenguas aparte de inglés. - Se precisan recursos en lenguas minoritarias.
Información que se puede obtener	¿Solo equivalencias u otra información?	- Equivalencias - Definiciones - Contextos - Colocaciones - Imágenes - Información regional - Nota del traductor si no hay equivalencias
Interfaz de búsqueda	¿Simple o compleja?	- Cuanto más simple mejor
Tipo de búsqueda	¿Búsqueda directa o por navegación?	- Principalmente búsqueda directa. Con posibilidad de navegación en las relaciones entre palabras.
	¿Consulta simple o compleja?	- Cuanto más simple mejor - Consulta de palabras truncadas - Consulta de dos o más palabras combinadas - Consulta por proximidad - Funciones de autocompletar - Funciones de autosugerencia
	¿En qué recurso?	- Búsqueda simultánea en más de un tipo de recurso. - Elección personalizada de los recursos.

	¿Qué resultados?	- Resultados bien organizados, por criterios de búsqueda, no por criterios alfabéticos. Deberían resaltarse los criterios de búsqueda.
Funciones de búsqueda	¿Semasiológico u onomasiológico?	- Ambas, de la denominación al significado y del significado a la denominación.
	¿Búsqueda semántica?	- Búsqueda semántica en los campos de definición - Búsqueda semántica en contexto (de relaciones sintagmáticas y semánticas). El análisis semántico debe ofrecer el significado más apropiado al contexto. - Búsqueda semántica de imágenes - Funciones de autosugerencia semántica
	¿Redes sociales?	- Ventajas de la Web 2.0 - Aprovechar las redes sociales para la creación, intercambio y evaluación de los recursos por parte de los usuarios. - Aprovechar las ventajas de las redes sociales para poner en contacto a los expertos. Evolución de los foros tradicionales.
Licencia	¿Software libre o privativo?	- La licencia libre permitiría la sinergia e interoperabilidad. - Licencia privativa poco restrictiva.

Con estas propuestas podemos contribuir a las preguntas y respuestas que se plantean distintos autores respecto de qué propiedades deberían tener los diccionarios que se adapten a un determinado tipo de usuario, en nuestro caso, los traductores.

Por ejemplo, respecto a los tipos de funciones, y en relación con las propuestas de la teoría funcional de la lexicografía (véase Bergholtz 2011; Bergholtz y Tarp 2003; Tarp 2007, 2008a, entre otros), sobre las situaciones o necesidades de los usuarios que podrían cubrirse en un recurso lexicográfico electrónico, diremos que los traductores de nuestro estudio han mostrado necesidades tanto *cognitivas* como *comunicativas*.

Dentro de lo que estos autores consideran necesidades comunicativas, los traductores de nuestro estudio tienen una necesidad de recepción, es decir, a partir de la consulta de una palabra en el diccionario pretenden recuperar su significado; también una necesidad de traducción, a partir de la introducción de una palabra precisan recuperar un equivalente,

pero no un equivalente descontextualizado, sino con información sobre el significado y el contexto; finalmente, una necesidad de producción, en las dos funciones propuestas por Bergenholtz (2011: 44-52). Primero, la función *I am writing a text and am looking for an expression with a specific meaning*, en este caso el texto que el traductor está escribiendo es el texto en LM y necesita una expresión en LM que refleje un significado determinado del texto en LO. Segundo, la función *I am writing a text with a specific expression*, donde el traductor precisa conocer la fraseología y colocaciones típicas para poder expresarse con naturalidad en la traducción. Asimismo, los traductores señalan también necesidades de tipo cognitivo, cuando lo que necesitan es información amplia para documentarse sobre una temática.

Finalmente, nos gustaría añadir que muchas de nuestras propuestas pueden satisfacerse con tecnologías que ya se encuentran en desarrollo y se están aplicando, en mayor o menor medida, en algunos recursos, como han sugerido también otros autores. En este sentido, de lo que se trata es de saber aplicar y combinar correctamente estas tecnologías teniendo en mente las necesidades que han manifestado los traductores, aunque también será necesario continuar avanzando en las tecnologías que ya existen y diseñar otras nuevas.

Por consiguiente, pensamos que algunos de los avances que se están produciendo en la investigación se encuentran en la dirección correcta y que, orientados a satisfacer las demandas que hemos señalado a lo largo de este estudio, podrían mejorar considerablemente las búsquedas de los traductores.





## 6 CONCLUSIÓN

Nuestro trabajo ha constituido una primera aproximación a las estrategias de búsqueda onomasiológica en traducción. Inicialmente, en la primera parte de antecedentes de este trabajo, hemos visto que la idea de elaborar diccionarios onomasiológicos no es nueva, ya que tradicionalmente se ha mostrado un interés por avanzar en la creación de diccionarios que permitieran realizar una búsqueda del significado a la denominación, por oposición a los clásicos diccionarios que permiten la búsqueda a partir de una palabra para obtener una definición. A modo de ejemplo pueden tomarse las iniciativas de diccionarios y proyectos que hemos presentado en el primer capítulo.

Sin embargo, aunque el avance de las tecnologías está revolucionando los formatos y todo parece indicar que se están dando pasos importantes que podrían favorecer el desarrollo de recursos que ofrezcan a los usuarios búsquedas a partir del significado para obtener una denominación (búsqueda onomasiológica), la realidad es que se observa todavía en el mercado una preponderancia de los diccionarios tradicionales, de búsqueda semasiológica y que, además, apenas aprovechan las ventajas del formato electrónico; también numerosos diccionarios electrónicos disponibles son simples adaptaciones del formato impreso y todavía muchos diccionarios se encuentran disponibles solo en papel.

Asimismo, hemos observado que los traductores en su actividad diaria tienen una necesidad de terminología, que precisan encontrar en los recursos y, por tanto, los traductores son usuarios potenciales de los recursos de búsqueda, conclusión a la que llegábamos en el capítulo 2.

De ahí que hayamos planteado un estudio dividido en dos grandes partes: las *técnicas de búsqueda* en los recursos (parte 2) y las *estrategias de búsqueda* de los traductores (parte 3).

Por un lado, en la parte 2, hemos realizado un estudio de los recursos que se encuentran disponibles en la actualidad y que podría utilizar un traductor. Este primer acercamiento se ha realizado desde el punto de vista de los recursos y, por ello, tiene una perspectiva técnica. La idea era conocer en primer lugar qué opciones, qué funciones, es decir, qué *técnicas de búsqueda*, se han desarrollado hasta el momento en los recursos. Los resultados sobre técnicas de búsqueda se han obtenido a partir de las funciones de búsqueda descritas en la investigación y de nuestro propio análisis de los recursos.

En primer lugar, hemos visto que el análisis de las funciones de búsqueda detalladas en una serie de estudios ha contribuido a la descripción de las características y las

posibilidades en materia de búsquedas, donde últimamente los autores tratan de centrarse sobre todo en las facilidades que aporta el medio electrónico. Sin embargo, también hemos observado que en la descripción de los recursos se han adoptado diferentes criterios de clasificación y puntos de vista, lo que no nos ha permitido vislumbrar una ordenación clara de las posibilidades de búsqueda que ofrecen los recursos.

Por tanto, hemos detectado una necesidad de profundizar respecto a cómo son las formas de búsqueda en los recursos y cómo se pueden organizar. Así, en un segundo estadio del estudio sobre técnicas de búsqueda, hemos efectuado el análisis de un conjunto de recursos: diccionarios (Pastor y Alcina 2010, en prensa), corpus (Pastor y Alcina 2009) e Internet (Pastor y Alcina 2011), a partir del cual hemos sintetizado las técnicas de búsqueda en función de tres elementos (la consulta, el instrumento y el resultado). Esta clasificación de técnicas de búsqueda ha resultado válida para ser aplicada a la descripción y evaluación de cualquiera de los recursos examinados. Además, esta clasificación está abierta a la incorporación de los futuros avances que se produzcan en los recursos. Todo nuestro análisis sobre técnicas de búsqueda en los recursos junto con nuestra propuesta de clasificación de técnicas de búsqueda en diccionarios, corpus e Internet la hemos presentado en el capítulo 3.

Por otro lado, la parte 3 de nuestra investigación ha consistido en un acercamiento desde el punto de vista de los usuarios. Los diseñadores de recursos han reconocido la importancia de tener en cuenta a los usuarios a la hora de elaborar un recurso. De hecho, hemos comprobado que existe una importante corriente investigadora que aboga por la elaboración de recursos centrados en el usuario.

Así, hemos realizado un análisis de estudios previos que han abordado las *estrategias de búsqueda* de los traductores, y que se han centrado en las formas de uso por parte de estos usuarios de diccionarios, corpus e Internet. Los resultados aportados por estos autores en sus investigaciones nos han servido como materia prima de un análisis en el que hemos mantenido un primer contacto con las estrategias de búsqueda de los traductores, en el cual hemos detectado una serie de tendencias o patrones que aparentemente se repetían en traducción. A partir de esta información, hemos elaborado una propuesta preliminar de clasificación de estrategias de búsqueda en traducción, que hemos presentado al final del capítulo 4.

No obstante, este análisis también nos ha permitido observar que las estrategias de búsqueda onomasiológica de los traductores apenas se habían tratado en la bibliografía, aunque sí que se sugería su importancia (Bowker 2000; Mackintosh 1998; Rogers y

Ahmad 1998). Igualmente, hemos percibido que la mayoría de estudios examinados sobre estrategias de búsqueda utilizan como sujetos a estudiantes de traducción y, por tanto, no nos ha sido posible extrapolar directamente a la traducción profesional los resultados obtenidos a partir de este primer análisis.

Esta situación justificaba nuestra investigación empírica sobre estrategias de búsqueda onomasiológica en traducción, en el capítulo 5, donde hemos realizado dos estudios. En un primer estudio piloto (Pastor 2008a, 2008b, 2010) distribuimos una encuesta a 15 traductores profesionales con el fin de evaluar las preguntas que nos podían resultar útiles para recabar la información que precisábamos y también de valorar el instrumento de medida más adecuado; asimismo, obtuvimos algunos resultados con los que realizamos un primer acercamiento a las estrategias de búsqueda onomasiológica mencionadas por estos traductores, lo que nos permitió corroborar el posible interés o la necesidad de un estudio sobre estrategias de búsqueda onomasiológica en traducción.

Después, planteamos nuestro estudio principal en el que realizamos una entrevista cualitativa en la que participaron un grupo de 48 traductores profesionales que nos informaron sobre las estrategias de búsqueda utilizadas y donde se centraron en las estrategias de búsqueda de tipo onomasiológico que les resultan útiles para acceder a denominaciones desconocidas. Además, aportaron sugerencias o demandas que podrían incluirse en los recursos para que se adapten mejor a estas necesidades.

A partir del análisis cualitativo de estas conversaciones con traductores, obtuvimos nuestros propios resultados sobre estrategias de búsqueda en traducción y, concretamente, sobre las estrategias de búsqueda onomasiológica. Posteriormente, la comparación de los resultados extraídos del análisis de nuestras entrevistas a traductores profesionales con los del análisis preliminar de los estudios anteriores nos ha permitido generar una nueva clasificación más completa de estrategias de búsqueda en traducción, que incorpora y sitúa a las estrategias de búsqueda onomasiológica en el conjunto de estrategias de búsqueda que utilizan los traductores. Asimismo, hemos podido verificar o rechazar algunas estrategias de búsqueda que aparentemente los traductores profesionales no utilizan con tanta frecuencia y hemos añadido otras que sí que han resultado relevantes y que pensamos que no han recibido suficiente trato en la bibliografía, como las búsquedas a partir del significado y de las relaciones tanto semánticas como sintagmáticas, y las consultas a recursos humanos, es decir, a personas.

## **6.1 Conclusión sobre estrategias de búsqueda onomasiológica en traducción**

Del análisis efectuado a lo largo de este trabajo resultaría arriesgado afirmar que hayamos obtenido una clasificación de estrategias de búsqueda en traducción universal, definitiva y válida para todos los traductores, en cualquier ámbito de conocimiento, cualquier combinación de lenguas, etc., precisamente porque, como hemos podido deducir de este mismo estudio, las estrategias de búsqueda de los traductores no siempre son fijas y pueden variar dependiendo de la situación, el tipo de texto, los problemas o necesidades concretos, la experiencia en traducción y los conocimientos específicos en un determinado campo de conocimiento, entre otros factores.

No obstante, sí que hemos identificado una serie de problemas o necesidades que propician la búsqueda del traductor en un recurso y son principalmente de dos tipos, de comprensión y de expresión. Por un lado, en las necesidades de comprensión nos ha resultado más sencillo establecer un patrón o un sistema de búsqueda reiterativo. Se consulta una palabra cuyo significado se desconoce, bien en un diccionario, en un corpus, en Internet, o en otro recurso, y se accede a una definición, un sinónimo o un contexto, que ayuda al traductor a comprender su significado.

Por otro lado, las necesidades o problemas de expresión hemos visto que en traducción ocurren en LM, cuando el traductor se enfrenta a un concepto o un significado de la LO que conoce para el cual tiene que ofrecer una denominación en la LM que desconoce, o cuando necesita expresarse en la LM una vez ha encontrado un equivalente pero desconoce la fraseología o las colocaciones típicas que debería utilizar para que el texto suene natural. Ante esta necesidad de expresión del traductor, las diferentes estrategias de búsqueda que entran en juego no están tan claras ni siguen un patrón tan definido, dependen en gran medida del tipo de información de la que dispone el traductor para efectuar la búsqueda y de los recursos con los que cuente, principalmente, pero también de los conocimientos previos del traductor y de su habilidad para efectuar búsquedas e interpretar los resultados. En este sentido, coincidimos con otros estudios en que las estrategias de búsqueda pueden llegar a ser flexibles (Bernardini et al. 2003: 10); sobre todo en esta necesidad de expresión es donde hemos visto que surgen estrategias de búsqueda de tipo onomasiológico.

Si bien pensamos que esta flexibilidad del traductor en las estrategias de búsqueda onomasiológica viene motivada seguramente porque los recursos no están específicamente

diseñados para el traductor y para las búsquedas que precisa realizar ante una necesidad de expresión que, en la mayoría de los casos, el traductor no es capaz de resolver con un solo recurso. Pensamos que el acceso onomasiológico a la terminología no ha avanzado todo lo que cabría esperar. Aunque de manera parcial sea posible realizar una serie de búsquedas que podrían asemejarse a la recuperación onomasiológica de terminología, la realidad es que las herramientas no ponen a disposición de los traductores facilidades para realizar búsquedas onomasiológicas. Las estrategias de búsqueda están dispersas porque cada sistema recupera la información de forma diferente. Cada herramienta aporta unas ventajas frente a las demás y ninguna por separado permite realizar una búsqueda onomasiológica completa.

Sin ánimo de resolver por completo la incógnita sobre qué es lo que necesita un traductor para acceder a las denominaciones que desconoce o que no recuerda a partir del significado, sí que hemos aportado una serie de pautas o de tendencias que hemos observado y que seguramente podrán despejar posibles caminos en el desarrollo de recursos para adaptarlos a estas necesidades onomasiológicas que tienen los traductores.

Los resultados más relevantes que creemos que aporta esta investigación son, por un lado, el tipo de información de la que suele partir un traductor cuando tienen una necesidad de expresión en LM, es decir, a partir de qué información consulta el traductor en un recurso cuando va a efectuar una consulta, y también en qué tipo de recurso podría interesar al traductor introducir la consulta para obtener como resultado la denominación que desconoce.

Hemos visto que, por regla general, para acceder a una denominación desconocida en LM, la primera opción de consulta de los traductores es a partir del término de la LO con el fin de encontrar en un recurso bilingüe una posible equivalencia en LM, es decir, los traductores realizan una **búsqueda de terminología bilingüe**. Para ello, las estrategias de búsqueda han consistido en la introducción del término original en un diccionario bilingüe, en un corpus bilingüe o en una memoria de traducción, el acceso a páginas bilingües en Internet, en una enciclopedia en línea o en un metabuscador, que podría englobar la búsqueda simultánea en cualquiera de los recursos anteriores. No obstante, hemos observado que la búsqueda aislada de terminología bilingüe tampoco resulta útil al traductor, sino que precisa que se incluya información sobre el significado que le permita seleccionar el equivalente más adecuado.

Si la búsqueda de terminología bilingüe falla, el traductor realiza una búsqueda normalmente en contexto, sobre todo en Internet, primero basada en la intuición, es decir, genera una **intuición de un posible equivalente** y lo comprueba en contexto.

Lo siguiente, en el caso de que la intuición no haya funcionado, el traductor trata de localizar el equivalente desconocido a partir de la información que obtiene del contexto y que está relacionada con el **significado de la denominación que desconoce**. En este sentido, el traductor consulta en un recurso a partir de palabras relacionadas con aquella que el traductor desconoce, con lo que recupera contextos en los que probablemente podrá localizar más fácilmente la palabra desconocida, ya que aparecerá en combinación con las palabras relacionadas que ha buscado.

Hemos determinado que una búsqueda onomasiológica es la búsqueda de una palabra o denominación a partir de su significado. Si tenemos en cuenta dos tipos de significados encontramos, por un lado, el significado denotativo, que se corresponde con el significado universal de una palabra, es el significado que tradicionalmente hemos podido encontrar en el artículo de un diccionario; sin embargo, también hemos visto que existe otro significado que viene determinado por la combinación en contexto de las palabras, la prosodia semántica de las palabras, las colocaciones (Bowker 2000: 32; Siepmann 2006: 34; Sinclair 2004: 174) y, en este sentido, una palabra puede definirse también por el uso que se hace de ella en un contexto en el que se combina con otras palabras.

En este segundo sentido es en el que entendemos la búsqueda de colocaciones o de relaciones sintagmáticas como una posible estrategia de búsqueda onomasiológica. Si el traductor busca a partir de palabras clave que describen el concepto (características y relaciones semánticas con otros conceptos) estaríamos hablando de una búsqueda a partir del significado conceptual o denotativo para llegar a una denominación desconocida, mientras que si busca a partir de colocaciones o palabras que se combinan en el plano sintagmático para obtener la palabra que se desconoce, entonces el traductor igualmente estaría buscando a partir del significado. Por consiguiente, ambos tipos de estrategias podrían considerarse estrategias de búsqueda onomasiológica porque van del significado a la denominación.

Finalmente, si el traductor no ha sido capaz de localizar la denominación desconocida mediante la búsqueda en recursos materiales, accede a los **recursos humanos**, consulta a un compañero, un especialista en la materia, o al propio cliente. Estos recursos humanos son los últimos a los que recurre un traductor en la búsqueda de un equivalente

desconocido, pero son a su vez los máspreciados porque suelen ofrecer siempre resultados satisfactorios.

Además de equivalentes en LM, los traductores necesitan conocer cuál es la forma natural de expresión en un determinado campo, necesitan acceder a la fraseología y las colocaciones típicas de la LM. Hemos visto que esta información la obtienen en los corpus, en Internet y, últimamente, también en diccionarios de colocaciones.

## **6.2 Conclusión sobre la elaboración de recursos orientados a los traductores**

En relación con lo que hemos visto sobre las necesidades variadas de los traductores y también por las demandas de estos profesionales en los recursos, hemos concluido que precisan **más de una función en los recursos** o diccionarios. Por tanto, un diccionario o un recurso que solo ofrezca la función onomasiológica no resultará completamente útil al traductor, como tampoco lo será uno que solo permita una búsqueda semasiológica. Hemos visto que los traductores necesitan **combinar estrategias** de distinto tipo en los recursos, **tanto semasiológicas como onomasiológicas**, de una denominación para obtener un significado y de un significado para obtener una denominación.

En este sentido, la elaboración de varios recursos monofuncionales diferentes a partir de una misma base de datos o recursos multifuncionales adaptados a las necesidades individuales de los usuarios, idea que sugieren los autores (Bergenholtz 2011; Tarp 2011, entre otros), puede resultar de gran interés para la elaboración de recursos orientados a los traductores que precisan funciones de distinto tipo, tanto cognitivas como comunicativas y, dentro de las comunicativas, tanto de recepción, como de transferencia y de producción. Para la elaboración de este tipo de recursos será necesario el aprovechamiento de las tecnologías más avanzadas que ofrece el formato electrónico.

El inconveniente principal señalado por los traductores respecto de los diccionarios exclusivamente bilingües es que los términos se encuentran en la mayoría de los casos descontextualizados, lo que provoca una indecisión o incluso una desconfianza acerca de si los términos propuestos por el diccionario bilingüe son los adecuados para expresar un significado determinado. Los traductores también señalan la deficiencia de que los diccionarios bilingües son demasiado generales y no son todo lo completos que deberían, reclaman diccionarios bilingües especializados que contengan terminología exhaustiva e información adicional como definiciones y contextos, no solo terminología bilingüe.

En este sentido, creemos que el formato electrónico ofrece enormes posibilidades en la elaboración de diccionarios bilingües que incorporen no solo mayor cantidad de información, sino que la presenten a los usuarios de una forma ordenada y sencilla. Los equivalentes pueden presentarse clasificados por campos temáticos, significados y contextos de uso. Además deberían incluir contextos de uso reales que permitan al traductor seleccionar el equivalente más adecuado (como por ejemplo en diccionarios como *FrameNet* o *DiCoInfo*, que incluyen contextos extraídos de un corpus), o bien podrían permitir la consulta directa de cualquier equivalente en un corpus complementario al diccionario, de forma similar al proyecto *EOHS-Term*, donde las entradas de las fichas del diccionario están vinculadas a un corpus en el que puede accederse a los contextos.

Tradicionalmente los diccionarios se han centrado en recoger el significado denotativo de las palabras. En los últimos años se ha cobrado interés por incorporar en los diccionarios otro tipo de información basada en el contexto que complementa el significado de las palabras y han surgido los llamados diccionarios de colocaciones.

Por medio de nuestro estudio hemos observado que el traductor necesita acceder a las colocaciones. Sin embargo, este tipo de búsqueda, que los traductores están realizando hasta ahora en algunos corpus, o sobre todo en textos paralelos que obtienen en Internet, o en memorias de traducción, está resultando complicada ya que estos recursos no incorporan los mecanismos necesarios para que el traductor pueda acceder de forma directa a las colocaciones de una palabra, sino que es más bien al contrario, el traductor debe desarrollar estrategias o métodos que de alguna manera le faciliten este tipo de búsqueda.

Hemos visto que los traductores alegaban la dificultad de buscar a partir de colocaciones en Internet porque la Web es demasiado amplia y los resultados que devuelve un buscador ante este tipo de búsquedas son infinitos y no tienen por qué estar relacionados con el contexto que el traductor necesita. Resulta necesaria la aplicación de filtros de búsqueda.

Por ello, consideramos de gran relevancia, como futuros recursos para traductores, las iniciativas en el desarrollo de diccionarios de colocaciones, sobre todo que cubran ámbitos especializados y que ofrezcan colocaciones en otras lenguas aparte del inglés, y en combinaciones de lenguas poco habituales. En este sentido, reflejamos los esfuerzos de proyectos como *DiCoInfo* para la creación de diccionarios de colocaciones especializados y multilingües, por ejemplo un diccionario multilingüe de colocaciones del ámbito de la informática, que se está elaborando en francés, inglés y español.



Es necesario que los traductores puedan acceder de forma onomasiológica a las colocaciones de una palabra (Jousse et al. 2011: 134; L'Homme y Leroyer 2009: 278), pero sin necesidad de conocer la teoría o la metodología con la que se ha elaborado el diccionario, por ejemplo en los diccionarios basados en la LEC, sin necesidad de tener conocimientos del metalenguaje de las funciones léxicas (como también se sugiere en Almela et al. 2011: 10; Verlinde et al. 2010: 4). Es imprescindible que se diferencie bien una búsqueda para administradores y una búsqueda para usuarios.

Asimismo, la búsqueda en las memorias de traducción es principalmente de terminología bilingüe, aunque añade un componente contextual, ya que las memorias de traducción seleccionan los equivalentes más probables en función del contexto anterior y posterior que los rodea, por lo que van un paso más allá de la mera búsqueda término-término que ofrece el diccionario bilingüe tradicional. Son muchos los traductores que emplean los programas de traducción asistida para la búsqueda en memorias de traducción a modo de corpus paralelos. En este sentido, a los traductores les resulta útil realizar la búsqueda y la selección de los equivalentes en función del contexto que los rodea.

Últimamente los traductores también están incorporando herramientas de metabúsqueda, como *dtSearch* o *Quest* que permiten la consulta simultánea en distintos tipos de documentación y recursos que se encuentren almacenados en el ordenador del traductor o en un servidor. Estas últimas herramientas que permiten la consulta personalizada de varios recursos en uno se han valorado de forma muy positiva por parte de los traductores.

Finalmente, Internet se está convirtiendo cada vez más en la principal herramienta de búsqueda, dada la desconfianza actual de los traductores respecto de los diccionarios y la escasez de corpus paralelos y memorias de traducción lo suficientemente grandes como para resolver las necesidades de expresión de los traductores.

Por ejemplo, hemos observado que los corpus, pese a que incorporan técnicas de búsqueda que pueden resultar muy útiles para mejorar las búsquedas, los traductores profesionales apenas utilizan estos recursos, sobre todo si lo comparamos con la cantidad de estudios que investigan o fomentan su uso entre los estudiantes de traducción (véase el apartado 4.3 y subapartados de este trabajo).

En este sentido, nuestros resultados coinciden con los de otras investigaciones (Jääskeläinen y Mauranen 2004: 50; Wheatley 2006). Los traductores profesionales de nuestro estudio argumentan que los corpus compilados que hay disponibles no les sirven en su trabajo porque son demasiado generales, y algunos están desfasados o son demasiado

caros. Según estos profesionales no han encontrado corpus suficientemente grandes o representativos de lenguaje especializado. Estas deficiencias de los corpus también se habían identificado en otros trabajos (Bernardini y Zanettin 2000; Bowker y Pearson 2002; Corpas Pastor 2001a, 2004; Corpas Pastor y Seghiri 2009; Pearson 1998; Sánchez-Gijón 2003a, 2003b; Wilkinson 2006b; Zanettin 2002b, entre otros). Por ello, estos autores recomiendan que los traductores compilen sus propios corpus (lo que se ha denominado corpus *ad hoc*) y han tratado de demostrar la utilidad de esta estrategia para la búsqueda de terminología en traducción. Sin embargo, los estudios sobre el uso de corpus *ad hoc* que hemos recopilado en este trabajo muestran el uso de este tipo de corpus por parte de estudiantes de traducción, en entornos de enseñanza, no de traducción real. Además, en la mayoría, los corpus *ad hoc* los facilita el propio profesor o investigador, no son los estudiantes de traducción los que los compilan.

Aunque coincidimos con la idea de que los corpus podrían ser recursos de gran utilidad para las búsquedas en traducción y, en concreto, para las búsquedas de naturaleza onomasiológica (como se sugiere en Bowker 2000), nuestro estudio con traductores profesionales revela que los traductores no suelen compilar sus propios corpus *ad hoc* para la actividad de traducción diaria; indican que, aunque probablemente pudieran resultar de utilidad, no disponen del tiempo necesario para compilar un corpus suficientemente completo, que la compilación de corpus es una tarea muy costosa y que en la práctica de traducción profesional no se realiza. Los traductores autónomos, especialmente, señalan que los temas que traducen pueden llegar a ser de lo más variado y que no les resultaría rentable compilar un corpus para un solo encargo de traducción, que además quedará desfasado rápidamente, en comparación con la búsqueda actualizada que ofrece la red.

La mayoría de traductores explica que, con la infinidad de información disponible en Internet, tratan de utilizar la Web como un corpus y *Google* como la principal herramienta de consulta. En este sentido, creemos que se debería trabajar en facilitar a los traductores la búsqueda en Internet como corpus especializado tanto monolingüe, como sobre todo multilingüe.

Por tanto, seguramente uno de los esfuerzos de la investigación debería encaminarse a incorporar las posibilidades de búsqueda que ofrece la lingüística de corpus, y que ya se reflejan en las técnicas de búsqueda que ofrecen algunos corpus, en herramientas de metabúsqueda, programas TAO y sobre todo en los buscadores de Internet para que estos avances se pongan al alcance de los traductores y, por ende, mejoren las búsquedas en estos recursos y verdaderamente permitan la consulta de la Web como si fuera un corpus.

En esta línea, los traductores apuntan algunas sugerencias, como por ejemplo la consulta de palabras tanto exactas como truncadas y la búsqueda combinada de palabras con criterios de proximidad, que no se encuentran suficientemente desarrolladas en el buscador *Google*. Algunos de nuestros traductores han señalado esta deficiencia y han expresado la ventaja que supondría que se pudieran realizar este tipo de búsquedas.

Asimismo, pensamos que el tratamiento y la presentación de los resultados por parte de los corpus pueden beneficiar a la búsqueda en Internet. Los traductores evidencian la arbitrariedad en la ordenación de los resultados de los buscadores, que no atienden a criterios lingüísticos que podrían interesar al traductor, sino más bien a otro tipo de factores como los comerciales. Por eso, señalan la dificultad y el tiempo que supone localizar la información relevante y fiable de entre los miles de resultados inservibles que aparecen. Algunos traductores comentan las ventajas que supondrían proporcionar mecanismos que permitieran filtrar esa información, o incluso la creación de dos redes paralelas, una de ocio y una profesional, por poner un ejemplo.

A su vez, los traductores señalan la relevancia de que *Google* resalte las palabras de búsqueda en un párrafo por cada resultado que recupera, sin embargo, este sistema tampoco satisface del todo el análisis lingüístico que puede precisar un traductor y, de hecho, los traductores han reconocido que tardan más tiempo del que les gustaría en localizar las fuentes que les interesan. La incorporación de técnicas de corpus de recuperación de concordancias KWIC, que resaltan las palabras clave en contexto, y la implementación de técnicas de obtención de colocaciones y de agrupaciones de palabras podría agilizar considerablemente el tratamiento de los datos de Internet. En este sentido, los esfuerzos realizados en la línea de la investigación de la Web as Corpus (WaC)<sup>144</sup> pueden favorecer las búsquedas de los traductores y los desarrolladores de estos recursos deberían tener en mente a los traductores como potenciales usuarios de estas herramientas y diseñarlas acorde a sus necesidades.

También debería avanzarse en la limitación de los resultados a páginas web de un determinado idioma o país, o a páginas web bilingües, de manera que se pudiera consultar Internet como un corpus bilingüe. A este respecto hemos visto que los avances proporcionados por recursos como *Linguee* han despertado el interés de los traductores.

---

<sup>144</sup> Para más información sobre los avances de la Web as Corpus, véanse en este trabajo las descripciones de *KWiCFinder* (Fletcher 2007a), *Web Concordancer* (Fletcher 2007b), *WebCONC* (Hüning 2001-2009) y *WebCorp* (Renouf et al. 2007), entre otros.

Un gran avance de las búsquedas en Internet sería que el traductor pudiera restringir la búsqueda a páginas de una determinada temática (como si con un solo clic pudiera compilar y buscar en un corpus *ad hoc* especializado), esta búsqueda estaría en la línea de las posibilidades que ofrecen buscadores específicos del tipo de *Google Scholar*, que en este caso recupera artículos de investigación, pero con la posibilidad de que se pudieran acotar los campos.

Por último, la búsqueda en Internet, como han reclamado los traductores de nuestro estudio, debería encaminarse a posibilitar a los traductores estrategias de búsqueda de tipo onomasiológico, búsquedas semánticas. Como sugería una de las traductoras de nuestro estudio, sería interesante que *Google* incorporara una herramienta similar a la de *define* pero que podría denominarse *explain and find a word*. Con *define* el traductor introduce una palabra y obtiene definiciones de Internet, con *explain* podría introducir una serie de palabras clave que describieran o que estuvieran relacionadas con un determinado significado y obtendría denominaciones que cumplieran estos requisitos en los contextos de Internet. Los traductores también solicitan estrategias de búsqueda semántica de imágenes en línea, para lo cual las imágenes estarían etiquetadas con una serie de palabras clave que describieran su contenido y el traductor, al introducir dichas palabras clave u otras relacionadas en un buscador, se recuperarían las imágenes que cumplieran estos requisitos y que, con gran probabilidad, resultarían útiles al traductor. En este sentido, los avances que se produzcan en materia de la Web semántica o Web 3.0 deberían tener en cuenta a los traductores e intentar satisfacer estas demandas.

A su vez, la Web 2.0 puede favorecer en el sentido de la elaboración de recursos disponibles en línea teniendo en cuenta las necesidades y valoraciones de los usuarios; en los recursos colaborativos los traductores pueden validar o descartar los resultados obtenidos y pueden añadir información que consideren relevante, aunque deberán acreditar la calidad de esta información, y esta deberá pasar por un control de los administradores. Asimismo, creemos que merece la pena investigar sobre cómo incorporar la tecnología de las redes sociales para poner en contacto a profesionales y expertos, es decir, para dar acceso a los recursos humanos cuando el recurso material no ha resultado satisfactorio. En este sentido, la Web social puede aportar novedades a la idea de los foros tradicionales.

Finalmente, de todas las ideas que hemos tratado en este trabajo extraemos dos recomendaciones para profesionales, docentes e investigadores en traducción, que creemos que deberán tener en cuenta en los próximos años. Por un lado, los usuarios, en este caso los traductores, deberán continuar mejorando las estrategias de búsqueda que utilizan en

los recursos para obtener el máximo provecho de los mismos, y esto debe reflejarse en la enseñanza que se imparte a los nuevos traductores que debe incorporar necesariamente la búsqueda en los recursos. Por otro lado, también hemos visto que se hace necesaria una reflexión sobre cómo pueden mejorarse los recursos y aprovechar las enormes ventajas que puede ofrecer en la actualidad el medio electrónico para incorporar nuevas opciones o técnicas de búsqueda que faciliten la consulta a los usuarios, en nuestro caso los traductores, entre las cuales deberá examinarse la búsqueda onomasiológica.

### **6.3 Trabajos futuros**

Este estudio ha supuesto un acercamiento a las estrategias de búsqueda onomasiológica en traducción. La decisión de optar por un análisis de tipo cualitativo y la elección de la entrevista como principal instrumento de obtención de datos han estado determinadas por el objetivo que nos hemos planteado en nuestra investigación, realizar una aproximación general a las estrategias de búsqueda onomasiológica en el proceso de traducción, un tema casi inexplorado hasta el momento.

Los resultados de este estudio podrán tomarse como punto de partida o como ideas en trabajos futuros sobre las estrategias de búsqueda onomasiológica de los traductores. La metodología que hemos utilizado puede adaptarse para observar por ejemplo las necesidades onomasiológicas de los traductores en situaciones concretas, especialidades concretas y combinaciones de idiomas específicos. Por ejemplo, para el desarrollo de un recurso o de un diccionario de una determinada especialidad o combinación de lenguas quizá pueda interesar conocer las opiniones de los traductores que se dediquen específicamente a traducir sobre ese campo y en ese par de lenguas, ya que podrán aportar sugerencias más concretas.

Asimismo, las proposiciones y demandas de los traductores en este estudio podremos aplicarlas en el diseño de recursos orientados a los traductores que permitan realizar búsquedas onomasiológicas. A partir de ahí, será necesario verificar si estas adaptaciones suponen realmente una mejora en las búsquedas de los traductores.

En definitiva, nuestro trabajo contribuye a la investigación sobre la adaptación de los recursos a un tipo de usuario (el traductor) y aporta ideas sobre la satisfacción de las necesidades individuales de los usuarios en situaciones de tipo onomasiológico, una cuestión que es de especial relevancia para el presente y pensamos que también lo será para el futuro de la investigación sobre la elaboración de recursos.

## 6 CONCLUSION

Our thesis is a first approach to onomasiological search strategies in translation. Initially, in the first part of this paper (the background), we discover that the idea of developing onomasiological dictionaries is certainly not new; traditionally, interest has been shown in advancing in the creation of dictionaries that allow searches from the meaning to find the word, as opposed to classical dictionaries that allow searches to be made from a word to retrieve a definition. We can find examples of these dictionary initiatives and projects in Chapter one.

Despite the fact that progress in technology revolutionizes formats, and that important steps are apparently being taken to promote the development of new resources which can offer users searches from the meaning to find the word (onomasiological search), the truth is that traditional dictionaries which provide semasiological searches still prevail in the market, and what is more, many of them do not take advantage of the electronic format. Many electronic dictionaries are simple adaptations of paper dictionaries, and still lots of dictionaries are only available in a printed form.

In addition, we observe that translators require terminology in their daily activities, which they need to find in the resources. Hence, translators are potential users of search resources. We reached this conclusion in Chapter two.

Consequently, we propose a study that is divided into two main parts: resources' *search techniques* (Part 2) and translators' *search strategies* (Part 3).

On the one hand, we examined in Part 2 the resources that are available today and which translators could use. This first approach was made from the resources viewpoint and has, therefore, a technical perspective. The idea was to firstly know what options, what functions, that is, what *search techniques*, have been developed to date in the resources. The search techniques results were obtained from not only the search functions described in the literature, but also from our own analysis of the available resources.

Firstly, we see that the analysis of search functions in some studies has contributed to the description of features and search possibilities, and that authors have recently attempted to focus mainly on the features provided by electronic resources. However, we also observe that in their description of the resources, some authors have adopted different classification criteria and points of view, which do not allow us to discern a clear categorization of the search possibilities offered by the resources.

Consequently, we observe the need to further investigate what the search possibilities in the resources are like and how they can be organized. Thus, the second stage of this study into search techniques analyzes a set of resources: dictionaries (Pastor and Alcina 2010, forthcoming), corpora (Pastor and Alcina 2009) and the Internet (Pastor and Alcina 2011) from which we synthesized their search techniques based on three elements (query, resource and result). This search technique classification is valid to be applied to the description and assessment of any examined resource. Furthermore, this classification is open to the incorporation of future advances into resources. The whole analysis of the search techniques in the resources and our proposed classification of the search techniques in dictionaries, corpora and the Internet are presented in Chapter three.

Furthermore, Part 3 of our research is based on an approach taken from the users' viewpoints. Resource designers have acknowledged the importance of taking users into account when developing a resource. In fact, we found that many studies encourage the development of user-oriented resources.

Therefore, we analyzed previous studies which address translators' search strategies, and which focus on the ways these users use dictionaries, corpora and the Internet. The results reported by the authors of these works were used as raw material in an analysis in which we established a first contact with translators' search strategies. Here we observed a number of search trends or patterns that are apparently repeated in translation. From this information, we developed a preliminary proposal for the classification of search strategies in translation, which we present at the end of Chapter four.

However, this analysis also allowed us to observe that the literature has virtually not addressed translators' onomasiological search strategies, although some authors suggest their importance (Bowker 2000b; Mackintosh 1998; Rogers and Ahmad 1998). We also noticed that most studies have examined translation students' search strategies. Therefore, it is not possible to directly extrapolate the results from this first analysis to professional translation.

This situation justifies our empirical research on onomasiological search strategies in translation in Chapter five, where we conducted two studies. In a first pilot study (Pastor 2008a, 2008b, 2010), we handed out a questionnaire to 15 professional translators in order to evaluate those questions that might prove useful to acquire the information we need and to also choose the most appropriate measuring instrument. We also obtained some results for a first approach to the onomasiological search strategies mentioned by these translators,

which allowed us to confirm the potential interest in or need for a study on onomasiological search strategies in translation.

Subsequently we proposed our main study, where we conducted a qualitative interview with 48 professional translators who informed us about their search strategies by focusing on those onomasiological search strategies that they find useful to retrieve unknown words. They also provided suggestions or indicated demands to be included in the resources in order to better suit these needs.

From the qualitative analysis of these conversations with translators, we obtained our own results on search strategies in translation, specifically on their onomasiological search strategies. Afterwards, the comparison made of the results of analyzing our interviews with professional translators with the preliminary analysis of previous studies allowed us to create a new more complete classification of search strategies in translation, which incorporates and places onomasiological search strategies among all the search strategies used by translators. We were also able to verify or reject some search strategies that professional translators apparently do not use that often. We also added other strategies that we found relevant and had not received enough attention in the literature, such as searches from semantic and syntagmatic relations and meanings, and human resources consultation; i.e., people.

### ***6.1 Conclusion on onomasiological search strategies in translation***

From the analysis conducted throughout this work, it is risky to affirm that we have obtained a universal classification of search strategies in translation that is definitive and valid for all translators in any field of knowledge, any language combination, etc. As deduced from this same study, this is precisely because translators' search strategies are not always fixed, but may vary depending on the situation, type of text, translators' problems or needs, experience and specific knowledge in a given field, among other factors.

However, we identify a number of problems or needs that encourage translators' search in a resource, which are mainly of two types: comprehension and expression. On the one hand, for comprehension needs, it was easier to establish a pattern or an iterative search system. Translators query a word whose meaning is unknown in a dictionary, corpus or in an Internet tool, or in another resource, and they access a definition, synonym or context which helps them understand its meaning.



On the other hand, we note that expression needs or problems in translation occur in the target language (TL) when translators face a known concept or meaning in the source language (SL) to which they must provide a word in the TL, which is unknown to them. Alternatively, having found the unknown equivalent, translators need to express in the TL, but they do not know the phraseology and the most typical collocations that should be used for the text to sound natural. Given translators' expression needs, the different search strategies that emerge are not that obvious or do not follow a clearly defined pattern, but depend mainly on the type of information and resources available to translators for searching purposes, and also on translators' prior knowledge and ability to perform searches and to interpret results. In this sense, we agree with other studies that search strategies can be flexible (Bernardini et al. 2003: 10), especially for this expression need where we have seen that onomasiological search strategies are used.

Nonetheless, we believe that such translator flexibility in onomasiological search strategies is probably motivated by the resources not being specifically designed for translators, and for also solving translators' expression needs which, in most cases, translators are unable to solve with a single resource. We believe that onomasiological access to terminology has not progressed as desired. Although it is partially possible to perform some search strategies that resemble the onomasiological retrieval of terminology, the truth is that tools do not provide translators with onomasiological search options. Search strategies spread because each system retrieves information differently. Each tool offers some advantages over others, but none allows a full onomasiological search on its own.

Although our aim is not to completely decipher what translators need to access an unknown or forgotten word from the meaning, we identify several patterns or trends that can surely guide the development of resources adapted to translators' onomasiological needs.

The most relevant results obtained from this research are as follows: firstly, the information type normally employed by translators when they have an expression need in the TL; i.e., which information is used by translators to query a resource; secondly, what type of resource is interesting to query in order to retrieve an unknown word.

We see that, in general, in order to access an unknown word in the TL, translators' first option is to query from the term in the SL to retrieve a potential equivalence in the TL in a bilingual resource; that is, translators perform a **search for bilingual terminology**. The search strategies that do this include the query of the source term in a bilingual

dictionary, a bilingual corpus or a translation memory, bilingual web pages on the Internet, an online encyclopedia or a metasearch engine, which could comprise simultaneous searches in any of the former resources. However, we have observed that finding isolated bilingual terminology is not useful for translators; they also need information about the meaning to be included to help them select the most appropriate equivalent.

If the search for bilingual terminology fails, translators normally perform a search in the context, especially on the Internet, firstly based on intuition; i.e., they generate the **intuition of a possible equivalent** and check it in the context.

If intuition does not work, then translators try locating the unknown equivalent from the context information that is related to the **meaning of the unknown word**. Hence translators query a resource from the words related to the unknown word to retrieve contexts where they can probably and more easily locate the unknown word as it may be combined with the related words that they have searched for.

We determine that an onomasiological search consists of finding a word from its meaning. If we consider two types of meanings we have, on the one hand, the denotative meaning, which corresponds to the universal meaning of a word; it is the meaning that can be traditionally found in the article of a dictionary. However, we also observe another meaning determined by the combination of words in the context, the semantic prosody of words, collocations (Bowker 2000: 32; Siepmann 2006: 34; Sinclair 2004: 174) and, in this sense, a word's meaning can be also defined by its use in the context combined with other words.

Thus, we identify the search for collocations or syntagmatic relations between words as a potential onomasiological search strategy. If translators search from the keywords describing the concept (semantic properties and relations with other concepts), then we can talk about a search from a conceptual or denotative meaning to find an unknown word. Yet if they query from the word's collocates, which is a syntagmatic relation of the unknown word, then translators also search from the meaning. Therefore, both types of strategies can be considered to be onomasiological search strategies because they start a query from the meaning to obtain the word.

Finally, if translators are unable to locate an unknown word by searching in material resources, they access **human resources**, they ask a colleague, an expert in the field, or even the customer. These human resources are the last resources used by translators in their search for unknown equivalents. Notwithstanding, they are, at the same time, the most valuable because they usually provide satisfactory results.

Besides TL equivalents, translators need to express themselves naturally in a particular field; they need to know the typical phraseology and collocations in the TL. We have seen that this information is obtained in corpora, on the Internet and, more recently, in collocations dictionaries.

## **6.2 Conclusion on the development of translator-oriented resources**

Based on what we have seen about translators' diverse needs, and also about the demands of these professionals in relation to the resources, we conclude that **more than one function in the resources** and dictionaries is required. Therefore, a dictionary or a resource that only offers the onomasiological function will neither prove completely useful to translators, nor will it be the one that allows only a semasiological search. We have seen that translators need to **combine different types of strategies** in the resources, these being **both semasiological and onomasiological strategies**, by not only querying from a word to retrieve the meaning, but by also querying from a meaning to retrieve the word.

In light of this, the development of several different monofunctional resources from the same database or polyfunctional resources tailored to users' individual needs, as suggested by some authors (Bergenholtz 2011; Tarp 2011, among others), may be of much interest to help develop translator-oriented resources requiring different kinds of cognitive and communicative functions, and within communicative functions (reception, production and transfer). In order to develop these resources, we require the use of the latest technologies offered by electronic formats.

The main drawback noted by translators in bilingual dictionaries is the decontextualization of terms, causing translators to hesitate, or even doubt, about whether the terms proposed by the bilingual dictionary are suitable for expressing a particular meaning. Translators also indicate the deficiency of bilingual dictionaries because they are too general and not as complete as they should be. Consequently, they demand specialized dictionaries containing exhaustive terminology and additional information, such as definitions and contexts, and not only bilingual terminology.

In line with this, we believe that the electronic format offers numerous possibilities to develop bilingual dictionaries which include not only more information, but also present it to users in an organized and easy fashion. Equivalents could be accessed by being classified into thematic areas, meanings and use contexts. Furthermore, real contexts should be included to allow translators to select the most appropriate equivalent (for

example, the same as in the *FrameNet* or *DiCoInfo* dictionaries, which include contexts extracted from corpora), or which allow direct queries in a complementary dictionary corpus, similarly to the *EOHS-Term* project where dictionary articles are linked to a corpus in which contexts can be consulted.

Dictionaries have traditionally focused on collecting the denotative meaning of words. In recent years, growing interest has been shown in other types of dictionaries which include context information that complements the meaning of words, and collocations dictionaries are also emerging.

In our study, we observe that translators need to access collocations. To date, however, this type of search which translators perform in corpora, especially in parallel texts on the Internet or in translation memories, is proving difficult because these resources do not include the necessary mechanisms to provide translators with direct access to collocations. Indeed, quite the opposite is true as translators must develop strategies or methods which somehow allow this type of search in the resources.

We observe that translators claim that it is difficult to look for collocations on the Internet because the scope of the Web is too broad. Thus, the results retrieved by a search engine for such searches may be endless and may not always be related to the context required by translators. So, filters must be employed.

Therefore, we find that initiatives in developing collocations dictionaries are of much importance as future resources for translators, particularly in specialized areas. They should offer collocations in other languages other than English, along with unusual language combinations. In this sense, we sincerely acknowledge the efforts of projects such as *DiCoInfo* to create multilingual specialized collocations dictionaries; for example, a multilingual collocations dictionary in the field of computing, which is being currently prepared in French, English and Spanish.

It is necessary to provide translators onomasiological access to collocations (Jousse et al. 2011: 134; L'Homme and Leroyer 2009: 278) without them having to know the theory or methodology used to develop the dictionary; for instance, in dictionaries based on ECL, without users having to know the metalanguage of lexical functions (as also stated by Almela et al. 2011: 10; Verlinde et al. 2010: 4). It is essential to differentiate a query for administrators and a query for users.

Furthermore, in a translation memory, translators search mainly for bilingual terminology. However, these resources add a contextual component as translation memories select the most likely equivalents in accordance with the context surrounding

them. Hence, the mere “term to term” search offered by traditional bilingual dictionaries takes another step forward. Many translators use CAT tools to query translation memories as parallel corpora. Therefore, translators find it useful to search and select equivalents based on the context surrounding them.

Recently, translators have also incorporated metasearch tools, such as *Quest* or *dtSearch*, which allow a simultaneous query in different types of documents and resources that may be stored on translators’ computers or in a server. These newest tools to allow a customized query in one or several resources are highly valued by translators.

Finally, the Internet is increasingly becoming the main search tool for translators given the current distrust of translators in dictionaries, and the scarcity of parallel corpora and translation memories that are large enough to meet translators’ expression needs.

For example, we notice that professional translators hardly ever use corpora despite them incorporating search techniques that can be most useful for improving searches, especially when compared to the number of studies investigating or promoting their use among translation students (see Section 4.3 and subsections of this thesis).

In this sense, our results are consistent with other studies (Jääskeläinen and Mauranen 2004: 50; Wheatley 2006). The professional translators who participated in our study argue that available compiled corpora are not useful for their work because they are too general, and that some are outdated, or simply too expensive. These professionals state that they have not found large enough or representative corpora of specialized language. Other studies have also reported these disadvantages of corpora (Bernardini and Zanettin 2000; Bowker and Pearson 2002; Corpas Pastor 2001a, 2004; Corpas Pastor and Seghiri 2009; Pearson 1998; Sánchez-Gijón 2003a, 2003b; Wilkinson 2006b; Zanettin 2002b, among others). Therefore, these authors recommend that translators compile their own corpora (called *ad hoc* corpora) and try to demonstrate the value of this strategy in the search for terminology in translation. However, there are studies on the use of *ad hoc* corpora which have been used by translation students in learning environments, but not in real translation. Moreover, the *ad hoc* corpus in most of these research works is provided by the teacher or researcher, and translation students do not compile the corpus.

Although we agree that corpora can be useful resources for translation searches, particularly for onomasiological searches (as suggested in Bowker 2000b), our study reveals that professional translators rarely compile their own *ad hoc* corpora for their daily translation activity. They state that, although corpora are probably useful, they do not have the time to compile a large enough corpus; corpus compilation is a very complex task and

is not that usual in professional translation practice. It is interesting to note that freelance translators state that translation text subjects may vastly vary, so it is not very useful to compile a corpus for a single translation project because it quickly goes out-of-date as compared with the updated searches available on the Web.

Most translators explain that given the endless information available on the Internet, they try to use the Web as a corpus, and *Google* is their main search tool. Therefore, we believe that we should work to allow translators to search on the Internet as a specialized monolingual, and especially multilingual, corpus.

So surely one of the research efforts should aim to incorporate the search capabilities offered by corpus linguistics (already reflected in some corpora search tools) into metasearch tools, CAT tools, and especially online search engines, for these advances to become available to translators, which would improve searches in these resources, and would genuinely allow to query the Web as a corpus.

In line with this, the surveyed translators have made some suggestions; for example, the possibility of querying both exact and truncated words; querying two or more words with distance restrictions. These options are not sufficiently developed in *Google*, and some translators have noted these deficiencies and have expressed the advantage of incorporating this kind of search.

We also believe that the processing and arrangement of the results by corpora could be useful on online search engines. Translators report the randomness by which search engines arrange the results, which do not take into account language criteria that may be of interest for translators, but consider others, like commercial criteria. For this reason, they voice the difficulty and time spent in locating relevant and reliable information from the thousands of results obtained, which actually prove useless. Some translators discuss the advantages that including mechanisms to filter information, or even the creation of two parallel Webs (i.e., leisure and professional), would imply.

Moreover, translators indicate the usefulness of highlighting query words in a paragraph for each result retrieved by *Google*; however, this system does not entirely satisfy the linguistic analysis, which may require translators; indeed, translators recognize that they spend more time locating interesting results than expected. The management of Internet data can be speeded up by incorporating corpus search techniques, such as the retrieval of KWIC concordances, highlighting the keywords in the context, as well as the

retrieval of collocations and clusters. In line with this, Web as Corpus (WAC)<sup>145</sup> research efforts can improve translators' search possibilities, and developers of these resources should consider translators as potential users of these tools to design them according to their needs.

In addition, progress should be made in limiting the results on the web pages of a particular language or country, or bilingual web sites, in order to allow searches to be made on the Internet as a bilingual corpus. By bearing this in mind, we have seen that the evolution of resources such as *Linguee* have aroused translators' interest.

Furthermore for translators, a breakthrough in Internet searches would be the possibility of restricting queries to pages to a specific thematic (that is, translators could compile and search a specialized *ad hoc* corpus with one click). This kind of search would be in line with the possibilities provided by specific search engines, such as *Google Scholar*, which recovers research articles, but would offer the possibility of limiting searches to more specific fields.

Finally, search possibilities on the Internet, as claimed by the translators who participated in our study, should aim to enable translators' onomasiological search strategies, semantic searches. As suggested by one of the translators of our study, it would be interesting if *Google* incorporated a tool similar to *define*, but that could be called *explain and find a word*. With *define*, translators introduce a word and retrieve definitions on the Internet. With *explain* translators, could introduce a sequence of keywords to describe a particular meaning, or one related to it, and they would retrieve words defined in that way in Internet contexts. Translators also claim semantic search strategies to retrieve images online. To this end, images could be labelled with a set of keywords describing their content and, if translators introduced these or other related keywords into a search engine, images meeting these requirements would be retrieved, which would most probably prove useful to translators. Thus any progress made in the Semantic Web or Web 3.0 areas should contemplate translators in an attempt to satisfy their demands.

Besides Web 2.0 can support the development of online resources by taking into account users' needs and ratings; in collaborative resources, translators can validate or discard the results, and can add the information that they consider relevant. However, they

---

<sup>145</sup> For more information on the progress of the Web as Corpus, see in this work the description of *KWiCFinder* (Fletcher 2007a), *Web Concordancer* (Fletcher 2007b), *WebCONC* (Hüning 2001-2009) and *WebCorp* (Renouf et al. 2007), among others.

should prove the quality of this information, and it should be under the control of administrators. We also believe that it is worth investigating how to incorporate social networking technology to bring together professionals and experts; that is, to provide access to human resources when material resources are not helpful. In this sense, the social Web can bring innovations to traditional forums.

In short, all the ideas discussed in this paper allow us to offer two recommendations for translation professionals, teachers and researchers, which we believe ought to be considered in forthcoming years. The first is that users, in our case, translators, should continue to improve the search strategies they use in the resources in order to make the best of them, and this should be reflected in the education received by new translators, who should necessarily incorporate search into the resources. The second is that we have also observed that there is a need to reflect on how to improve the resources and how to harness the enormous advantages that the electronic medium offers in order to incorporate new options or search techniques that facilitate users' queries, in our case, translators' queries, among which onomasiological searches are a must.

### **6.3 Future Work**

This study offers an approach to onomasiological search strategies in translation. The choice of a qualitative analysis and interviews as our measurement instruments to collect data was motivated by our research objectives; we aimed to provide a general approach to onomasiological search strategies in the translation process, a practically unexplored subject to date.

The results obtained with this study can be used as a starting point or as ideas for future work on translators' onomasiological search strategies. Our methodology can be adapted to observe translators' onomasiological needs in specific situations, specialties and language combinations. For example, in order to develop a resource or a dictionary in a particular field or language combination, it may be interesting to know the opinions of those translators who specialize in that field and in those languages since they may provide more specific suggestions.

In addition, the proposals and demands of the translators in this study may be applied to help design translation-oriented resources by allowing them to perform onomasiological search strategies. Subsequently, it will be necessary to verify whether these adaptations really improve translators' searches.



Ultimately, our work contributes to the research on adapting resources to a specific user (translators) and provides ideas to help satisfy these users' individual needs in onomasiological situations. This issue we believe is particularly relevant for present and future research into resources development.



## BIBLIOGRAFÍA<sup>146</sup>

- Abadal, E. (2004). "Control terminológico en la recuperación de la información". En C. Gonzalo García y V. García Yebra (eds.), *Manual de documentación y terminología para la traducción especializada*. Madrid: Arco/Libros, 409-420.
- Abaitua, J. (2002). "Tratamiento de corpora bilingües". En M. A. Martí y J. Llisterri (eds.), *Tratamiento del lenguaje natural*. Barcelona: Edicions Universitat de Barcelona, 61-90.
- Abate, F. R. (1985). "Dictionaries Past & Future: Issues and Prospects". *Dictionaries*, 7, 270-283.
- Aguado de Cea, G., Álvarez de Mon y Rego, I. y Pareja Lora, A. (2003). "Primeras aproximaciones a la anotación lingüístico-ontológica de documentos de la Web Semántica: OntoTag". *Revista Iberoamericana de Inteligencia Artificial*, 7(18), 37-50. Disponible en <http://journal.iberamia.org/> [Consulta: 8 de mayo de 2012].
- Ahmad, K. y Fulford, H. (1992). *Knowledge Processing 4. Semantic Relations and their Use in Elaborating Terminology* (Computing Sciences Technical Report CS-92-07). Guildford: University of Surrey.
- Ahmad, K. y Rogers, M. (1997). "Corpus Linguistics and Terminology Extraction". En S. E. Wright y G. Budin (eds.), *Handbook of Terminology Management. Application-Oriented Terminology Management* (Vol. 2). Amsterdam/Philadelphia: John Benjamins, 725-760.
- Alanen, A. (1996). "The Translator and the Current Services of the Internet ". *Proseminar paper presented at University of Tampere, Finland, Department of Translation/Carl Wieck*. Disponible en <http://www.uta.fi/~tranuk/prosemc.htm> [Consulta: 8 de mayo de 2012].
- Alarcón, R. (2009). *Extracción automática de contextos definatorios en corpus especializados*. (Tesis doctoral, Ciencias del Lenguaje y Lingüística Aplicada), Universidad Pompeu Fabra.
- Alarcón, R., Bach, C. y Sierra, G. (2008). "Extracción de contextos definatorios en corpus especializados: Hacia una elaboración de una herramienta de ayuda terminográfica". *Revista Española de Lingüística*, 37, 247-278.
- Alcaraz Varó, E. y Martínez Linares, M. A. (1997) *Diccionario de lingüística moderna*. Barcelona: Ariel.
- Alcina, A. (1997). "Soportes de almacenamiento y formas de difusión de datos terminológicos: las fuentes de información en terminología". *Revista Española de Lingüística Aplicada*, 12, 221-235.
- Alcina, A. (2003). "La programación de objetivos didácticos en Terminótica atendiendo a las nuevas herramientas y recursos". En N. Gallardo San Salvador (ed.), *Terminología y traducción: un bosquejo de su evolución*. Granada: Atrio, 79-90.
- Alcina, A. (2008). "Translation Technologies Scope, Tools and Resources". *Target*, 20(1), 80-103.
- Alcina, A. (2009a). "Metodología y tecnologías para la elaboración de diccionarios terminológicos onomasiológicos". En A. Alcina, E. Valero y E. Rambla (eds.), *Terminología y sociedad del conocimiento*. Bern: Peter Lang, 33-58.

---

<sup>146</sup> Hemos citado con dos apellidos las obras de aquellos autores que habitualmente firman sus publicaciones de dicha manera, así como las de aquellos en los que el uso de un solo apellido pudiera resultar ambiguo.

- Alcina, A. (2009b). "Teaching and Learning Terminology: New Strategies and Methods". *Terminology*, 15(1), 1-9.
- Alcina, A., Soler, V. y Estellés, A. (2005). "Internet como instrumento para la documentación en terminología y traducción. Hacia las plataformas de recursos electrónicos para el traductor especializado". En D. Sales (ed.), *La Biblioteca de Babel: Documentarse para traducir*. Granada: Comares, 221-241.
- Almela, M., Cantos, P. y Sánchez, A. (2011). "Towards a Dynamic Combinatorial Dictionary: A Proposal for Introducing Interactions between Collocations in an Electronic Dictionary of English Word Combinations". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 1-11. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Alonso, A., Millon, C. y Williams, G. (2011). "Collocational Networks and their Application to an E-Advanced Learner's Dictionary of Verbs in Science (DicSci)". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 12-22. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Alonso Ramos, M. (2009). "Hacia un nuevo recurso léxico: ¿fusión entre corpus y diccionario?" En P. Cantos Gómez y A. Sánchez Pérez (eds.), *A Survey of Corpus-based Research. Panorama de investigaciones basadas en corpus*. Murcia: AELINCO, 1191-1207.
- Alvar Ezquerro, M. (1998) *Diccionario ideológico de la lengua española*. Barcelona: VOX-Bibliograf.
- Alves, F. (ed.). (2003). *Triangulating Translation*. Amsterdam/Philadelphia: John Benjamins.
- Alves, F. y Liparini Campos, T. (2009). "Translation technology in time: investigating the impact of translation memory systems and time pressure on types of internal and external support". En S. Göpferich, A. L. Jakobsen y I. M. Mees (eds.), *Behind the mind. Methods, models and results in translation process research*. Copenhagen: Samfundslitteratur, 191-218.
- Anaya Revuelta, I. (2005). "La función onomasiológica de los diccionarios: el diccionario como herramienta para aprender nuevas voces o recuperar la palabra olvidada". *Hesperia, Anuario de Filología Hispánica*, VIII, 7-26.
- Andersen, B. y Almind, R. (2011). "The Technical Realization of Three Monofunctional Phrasal Verb Dictionaries". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 208-229.
- Andersen, B. y Nielsen, S. (2009). "Ten Key Issues in Lexicography for the Future". En H. Bergenholtz, S. Nielsen y S. Tarp (eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 355-365.
- Arntz, R. (1993). "Terminological Equivalence and Translation". En H. Sonneveld y K. Loening (eds.), *Terminology: Applications in Interdisciplinary Communication*. Amsterdam/Philadelphia: John Benjamins, 5-19.
- Asadi, P. y Séguinot, C. (2005). "Shortcuts, Strategies and General Patterns in a Process Study of Nine Professionals". *Meta*, 50(2), 522-547. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Aston, G. (1999). "Corpus use and learning to translate". *Textus*, 12, 289-314.

- Aston, G. (2000). "I corpora come risorse per la traduzione e per l'apprendimento". En S. Bernardini y F. Zanettin (eds.), *I corpora nella didattica della traduzione - Corpus use and learning to translate*. Bologna: CLUEB, 21-29.
- Aston, G. (2009). "Foreword". En A. Beeby, P. Rodríguez Inés y P. Sánchez-Gijón (eds.), *Corpus Use and Translating*. Amsterdam/Philadelphia: John Benjamins, IX-X.
- Atkins, B. T. S. (1996). "Bilingual Dictionaries: Past, Present and Future". En M. Gellerstam, J. Järborg, S.-G. Malmgren, K. Norén, L. Rogström y C. R. Pappmehl (eds.), *7th EURALEX International Congress (Proceedings Part II)*, Göteborg, 13-18 de agosto de 1996, 515-546. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex1996\\_2/](http://www.euralex.org/elx_proceedings/Euralex1996_2/) [Consulta: 8 de mayo de 2012].
- Atkins, B. T. S. y Knowles, F. E. (1988). "Interim Report on the EURALEX / AILA Research Project into Dictionary Use". En I. Magey y J. Zigany (eds.), *3rd International EURALEX Congress (BudaLEX 88 Proceedings)*, Budapest, 4-9 de septiembre de 1988, 381-392. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex1988/](http://www.euralex.org/elx_proceedings/Euralex1988/) [Consulta: 8 de mayo de 2012].
- Atkins, B. T. S., Lewis, H., Summers, D. y Whitcut, J. (1987). "A Research Project into the Use of Learners' Dictionaries". En A. P. Cowie (ed.), *The Dictionary and the Language Learner. Papers of the EURALEX Seminar at the University of Leeds, 1-3 April 1985*. Tübingen: Niemeyer, 29-43.
- Atkins, B. T. S. y Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Atkins, B. T. S. y Varantola, K. (1997). "Monitoring Dictionary Use". *International Journal of Lexicography*, 10(1), 1-45.
- Atkins, B. T. S. y Varantola, K. (1998a). "Language Learners Using Dictionaries: The Final Report on the EURALEX/AILA Research Project on Dictionary Use". En B. T. S. Atkins (ed.), *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Niemeyer, 21-81.
- Atkins, B. T. S. y Varantola, K. (1998b). "Monitoring Dictionary Use". En B. T. S. Atkins (ed.), *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Niemeyer, 83-122.
- Austermühl, F. (2001). *Electronic Tools for Translators*. Manchester/Northampton: St. Jerome.
- Bach, C. (2005). "Los marcadores de reformulación como localizadores de zonas discursivas relevantes en el discurso especializado". *Revista Debate Terminológico*, 1. Disponible en <http://ricoterm.iula.upf.edu/docums/bach05.pdf> [Consulta: 8 de mayo de 2012].
- Baeza-Yates, R. y Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. New York: ACM Press.
- Baker, M. (1993). "Corpus Linguistics and Translation Studies: Implications and Applications". En M. Baker, G. Francis y E. Tognini-Bonelli (eds.), *Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins, 233-252.
- Barnhart, C. L. (1962). "Problems in Editing Commercial Monolingual Dictionaries". En F. W. Householder y S. Saporta (eds.), *Problems in Lexicography*. Bloomington: Indiana University Research Center for Language and Semiotic Studies, 161-181.
- Baroni, M. y Bernardini, S. (2004). "BootCaT: Bootstrapping Corpora and Terms from the Web". En M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa y R. Silva (eds.), *4th International Conference on Language Resources and Evaluation (LREC 2004)*,

- Portugal, 1313-1316. Disponible en <http://www.lrec-conf.org/proceedings/lrec2004/> [Consulta: 8 de mayo de 2012].
- Baroni, M., Kilgarriff, A., Pomikálek, J. y Rychlý, P. (2006). "WebBootCaT: Instant Domain-Specific Corpora to Support Human Translators". *11th Annual Conference of the European Association for Machine Translation (EAMT 2006)*, Oslo, 247-252.
- Barrière, C. (2009). "Finding domain specific collocations and concordances on the Web". En I. Ilisei, V. Pekar y S. Bernardini (eds.), *International Workshop on Natural Language Processing Methods and Corpora in Translation, Lexicography and Language Learning*, Borovets (Bulgaria), 1-8.
- Battelle, J. (2006). *The Search. How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*. New York: Portfolio.
- Baxter, J. (1980). "The Dictionary and Vocabulary Behaviour: a Single Word or a Handful?" *TESOL Quarterly*, 14(3), 325-336.
- Beeby, A., Rodríguez Inés, P. y Sánchez-Gijón, P. (2009). "Introduction". En A. Beeby, P. Rodríguez Inés y P. Sánchez-Gijón (eds.), *Corpus Use and Translating*. Amsterdam/Philadelphia: John Benjamins, 1-8.
- Beeferman, D. (1998). "Lexical discovery with an enriched semantic network". En S. Harabagiu (ed.), *Proceedings of the Workshop on Applications of WordNet in Natural Language Processing Systems, ACL/COLING*, Montréal, 135-141.
- Béjoint, H. (1981). "The Foreign Student's Use of Monolingual English Dictionaries: A Study of Language Needs and Reference Skills". *Applied Linguistics*, 2(3), 207-221.
- Béjoint, H. (1989a). "The Teaching of Dictionary Use: Present State and Future Tasks". En F. J. Hausmann, O. Reichmann, E. Wiegand y L. Zgusta (eds.), *Wörterbücher/Dictionaries/Dictionnaires, Ein internationales Handbuch zur Lexikographie/An International Encyclopedia of Lexicography/Encyclopédie internationale de lexicographie* (Vol. 1). Berlin: de Gruyter, 208-215.
- Béjoint, H. (1989b). *Traditions et innovations dans les dictionnaires monolingues generaux de langue anglaise depuis 1960*. (Tesis doctoral, Doctorat d'Etat Es-Lettres et Sciences Humaines), Université de Franche-Comte.
- Béjoint, H. (2010). *The Lexicography of English*. Oxford: Oxford University Press.
- Bell, R. T. (1991). *Translation and Translating. Theory and Practice*. New York: Longman.
- Benito-Ruiz, E. (2009). "Infoxication 2.0". En M. Thomas (ed.), *Handbook of Research on Web 2.0 and Second Language Learning*. Pennsylvania: IGI-InfoSci, 60-79.
- Bensoussan, M., Sim, D. y Weiss, R. (1984). "The Effect of Dictionary Usage on EFL Test Performance Compared with Student and Teacher Attitudes and Expectations". *Reading in a Foreign Language*, 2(2), 262-276.
- Berber Sardinha, T. (2004). *Linguística de Corpus*. Barueri (Brasil): Manole.
- Bergen, D. (2006). "Translation strategies and the student of translation". En J. Tommola (ed.), *Kieli ja kulttuuri kääntäjän työvälineinä*. Turku: University of Turku, 109-126.
- Bergenholtz, H. (2003). "User-oriented Understanding of Descriptive, Proscriptive and Prescriptive Lexicography". *Lexikos*, 13, 1-16.
- Bergenholtz, H. (2005). "Den usynlige elektroniske produktions- og korrekturordbog". *LexicoNordica*, 12, 19-40.
- Bergenholtz, H. (2011). "Access to and Presentation of Needs-Adapted Data in Monofunctional Internet Dictionaries". En P. A. Fuertes-Olivera y H. Bergenholtz



- (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 30-53.
- Bergenholtz, H. y Bergenholtz, I. (2011). "A Dictionary is a Tool, a Good Dictionary is a Monofunctional Tool". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 187-207.
- Bergenholtz, H., Bothma, T. J. y Gouws, R. H. (2011). "A model for integrated dictionaries of fixed expressions". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 34-42. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Bergenholtz, H. y Johnsen, M. (2005). "Log Files as a Tool for Improving Internet Dictionaries". *Hermes, Journal of Linguistics*, 34, 117-142. Disponible en <http://www.regnskabsordbogen.dk/PDF/Hermes/Hermes-34-6-HB-Johnsen.pdf> [Consulta: 8 de mayo de 2012].
- Bergenholtz, H. y Nielsen, S. (2006). "Subject-field components as integrated parts of LSP dictionaries". *Terminology*, 12(2), 281-303.
- Bergenholtz, H., Nielsen, S. y Tarp, S. (eds.). (2009). *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang.
- Bergenholtz, H. y Tarp, S. (2003). "Two opposing theories: On H.E. Wiegand's recent discovery of lexicographic functions". *Hermes, Journal of Linguistics*, 31, 171-196. Disponible en [http://download2.hermes.asb.dk/archive/download/H31\\_11.pdf](http://download2.hermes.asb.dk/archive/download/H31_11.pdf) [Consulta: 8 de mayo de 2012].
- Bergenholtz, H. y Tarp, S. (2010). "LSP Lexicography or Terminography? The Lexicographer's Point of View". En P. A. Fuertes-Olivera (ed.), *Specialized Dictionaries for Learners*. Berlin: De Gruyter, 27-38.
- Bergeron, M. y Van Steenberg, M. (2000). "The Internet Search Tutorial for Translators". Disponible en <http://www.mabercom.com/websearch/index.html> [Consulta: 15 de marzo de 2008].
- Bergman, M. K. (2001). "White Paper: The Deep Web: Surfacing Hidden Value". *The Journal of Electronic Publishing*, 7(1). Disponible en <http://www.journalofelectronicpublishing.org/> [Consulta: 8 de mayo de 2012].
- Bergman, M. K. (2004). "Guide to Effective Searching of the Internet". Disponible en <http://www.brightplanet.com/images/stories/pdf/searchengin tutorial.pdf> [Consulta: 3 de septiembre de 2008].
- Bernardini, S. (2000a). *Competence, Capacity, Corpora: A Study in Corpus-Aided Language Learning*. Bologna: CLUEB.
- Bernardini, S. (2000b). "I corpora nella didattica della traduzione: dall'addestramento alla formazione". En S. Bernardini y F. Zanettin (eds.), *I corpora nella didattica della traduzione - Corpus use and learning to translate*. Bologna: CLUEB, 81-103.
- Bernardini, S., Stewart, D. y Zanettin, F. (2003). "Corpora in Translator Education. An Introduction". En F. Zanettin, S. Bernardini y D. Stewart (eds.), *Corpora in Translator Education*. Manchester: St. Jerome, 1-14.
- Bernardini, S. y Zanettin, F. (2000). "Introduzione". En S. Bernardini y F. Zanettin (eds.), *I corpora nella didattica della traduzione - Corpus use and learning to translate*. Bologna: CLUEB, 13-20.
- Berners-Lee, T., Hendler, H. y Lassila, L. (2001). "The Semantic Web". *Scientific American*, 29-37.

- Bertaccini, F. y Aston, G. (2001). "Going to the Clochemerle: exploring cultural connotations through ad hoc corpora". En G. Aston (ed.), *Learning with Corpora*. Bologna: CLUEB, 198-219.
- Bogaards, P. (1993). "Models of Dictionary Use" *Toegepaste Taalwetenschap in Artikelen 46/47 (Dutch Contributions to AILA '93. Selected in Honour of Johan Matter)*. Amsterdam: Free University, 17-28.
- Bogaards, P. (1998). "What type of words do language learners look up?" En B. T. S. Atkins (ed.), *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Niemeyer, 151-157.
- Bogaards, P. (1999). "Research on Dictionary Use: An Overview". En R. R. K. Hartmann (ed.), *Dictionaries in Language Learning. Recommendations, National Reports and Thematic Reports from the TNP Sub-Project 9: Dictionaries*. Berlin: Freie Universität Berlin, 32-34.
- Bogaards, P. (2003). "Uses and users of dictionaries". En P. Van Sterkenburg (ed.), *A Practical Guide to Lexicography*. Amsterdam/Philadelphia: John Benjamins, 26-33.
- Bonet Heras, J. (2004). "La necesaria especialización del traductor técnico". En C. Gonzalo García y V. García Yebra (eds.), *Manual de documentación y terminología para la traducción especializada*. Madrid: Arcos/Libros, 37-48.
- Bosque, I. (2004) *Redes. Diccionario combinatorio del español contemporáneo*. Madrid: SM.
- Bosque, I. (2006) *Diccionario combinatorio práctico del español contemporáneo*. Madrid: SM.
- Bothma, T. J. (2011). "Filtering and Adapting Data and Information in an Online Environment in Response to User Needs". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 71-102.
- Bowden, P. R., Halstead, P. y Rose, T. G. (1996). "Extracting Conceptual Knowledge from Text Using Explicit Relation Markers". En N. Shadbolt, K. O'Hara y G. Schreiber (eds.), *Advances in Knowledge Acquisition. 9th European Knowledge Acquisition Workshop, EKAW '96, Nottingham, United Kingdom, May 1996. Proceedings*. Berlin: Springer-Verlag, 147-162.
- Bowker, L. (1998a). "Exploitation de corpus pour la recherche terminologique ponctuelle". *Terminologies nouvelles*, 18, 22-27.
- Bowker, L. (1998b). "Using Specialized Monolingual Native-Language Corpora as a Translation Resource: A Pilot Study". *Meta*, 43(4), 631-651. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Bowker, L. (2000). "Towards a Methodology for Exploiting Specialized Target Language Corpora as Translation Resources". *International Journal of Corpus Linguistics*, 5(1), 17-52.
- Bowker, L. (2002). "Working Together: A Collaborative Approach to DIY Corpora". En E. Yuste (ed.), *1st International Workshop on Language Resources for Translation Work and Research (LREC 2002 Workshop Proceedings)*, Las Palmas de Gran Canaria, 29-32.
- Bowker, L. (2003a). "Corpus-based Applications for Translator Training: Exploring the Possibilities". En S. Granger, J. Lerot y S. Petch-Tyson (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam/New York: Rodopi, 169-184.



- Bowker, L. (2003b). "Terminology Tools for Translators". En H. Somers (ed.), *Computers and Translation. A Translator's Guide*. Amsterdam/Philadelphia: John Benjamins, 49-65.
- Bowker, L. (2006). "Translation Memory and "Text"". En L. Bowker (ed.), *Lexicography, Terminology, and Translation. Text-based Studies in Honour of Ingrid Meyer*. Ottawa: University of Ottawa Press, 175-187.
- Bowker, L. y Bennison, P. (2003). "Student Translation Archive. Design, Development and Application". En F. Zanettin, S. Bernardini y D. Stewart (eds.), *Corpora in Translator Education*. Manchester: St. Jerome, 103-118.
- Bowker, L. y Marshman, E. (2009). "Better integration for better preparation. Bringing terminology and technology more fully into translator training using the CERTT approach". *Terminology*, 15(1), 60-87.
- Bowker, L. y Pearson, J. (2002). *Working with Specialized Language. A practical guide to using corpora*. London/New York: Routledge.
- Breslin, J. G., Passant, A. y Decker, S. (2009). *The Social Semantic Web*. Heidelberg/New York: Springer.
- Budin, G. y Mörth, K. (2011). "Hooking up to the corpus: the *Viennese Lexicographic Editor's* corpus interface". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 52-59. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Budin, G. y Wright, S. E. (1997). "Terminology Management Resources on the Internet". En S. E. Wright y G. Budin (eds.), *Handbook of Terminology Management. Application-Oriented Terminology Management (Vol. 2)*. Amsterdam/Philadelphia: John Benjamins, 845-860.
- Burdach, A. M., Diéguez, M. I. y Lazo, R. M. (1990). "Creación de un banco de datos terminológicos: algunas perspectivas teóricas". *Actas del II Simposio de RITerm*. Disponible en <http://www.riterm.net/actes/2simposio/burdach.htm> [Consulta: 8 de mayo de 2012].
- Cabré, M. T. (1993). *La terminología: teoría, metodología, aplicaciones*. Barcelona: Antártida.
- Cabré, M. T. (1999). "Fuentes de información terminológica para el traductor". En M. Pinto y J. A. Cordón (eds.), *Técnicas documentales aplicadas a la traducción*. Madrid: Síntesis, 19-39.
- Cabré, M. T. (2000). "El traductor y la terminología: necesidad y compromiso". *Panacea*, 1(2), 2-4.
- Cabré, M. T. (2004). "La terminología en la traducción especializada". En C. Gonzalo García y V. García Yebra (eds.), *Manual de documentación y terminología para la traducción especializada*. Madrid: Arcos/Libros, 89-125.
- Cabré, M. T. (2006). "From Terminological Data Banks to Knowledge Databases: The Text as the Starting Point". En L. Bowker (ed.), *Lexicography, Terminology, and Translation. Text-based Studies in Honour of Ingrid Meyer*. Ottawa: University of Ottawa Press, 93-106.
- Cabré, M. T. (2007). "Términos y palabras en los diccionarios". En M. Emsel y J. Cuartero Otal (eds.), *Vernetzungen: Bedeutung in Wort, Satz und Text. Festschrift für Gerd Wotjak zum 65. Geburtstag*. Frankfurt: Peter Lang, 71-84.
- Cabré, M. T., Bach, C., Estopà, R., Feliu, J., Martínez, G. y Vivaldi, J. (2004a). "The GENOMA-KB project: towards the integration of concepts, terms, textual corpora and entities". En M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa y R. Silva (eds.),

- 4th International Conference on Language Resources and Evaluation (LREC 2004)*, Lisboa, European Languages Resources Association, 87-90. Disponible en <http://www.lrec-conf.org/proceedings/lrec2004/> [Consulta: 8 de mayo de 2012].
- Cabré, M. T., Estopà, R. y Feliu, J. (2004b). "A Specialized Knowledge Base: From Distributed Information to the Specialized Dictionary Construction". En G. Williams y S. Vessier (eds.), *11th EURALEX International Congress*, Lorient, 6-10 de julio de 2004, 867-872. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2004/](http://www.euralex.org/elx_proceedings/Euralex2004/) [Consulta: 8 de mayo de 2012].
- Cabré, M. T., Estopà, R., Freixa, J., Lorente, M. y Tebé, C. (2002). "Les necessitats terminològiques del traductor científic". En J. Chabás, R. Gaser y J. Rey (eds.), *2nd International Conference on Specialized Translation*, Universitat Pompeu Fabra, Barcelona, 165-174.
- Cabré, M. T. y Tebé, C. (1998). "Hacia un nuevo modelo de bancos de datos terminológicos". *Actas del VI Simposio de RITerm*. Disponible en <http://www.upf.edu/pdi/df/teresa.cabre/docums/ca02hacia.pdf> [Consulta: 8 de mayo de 2012].
- Campoy Cubillo, M. C. (2002). "Dictionary Use and Dictionary Needs of ESP Students: An Experimental Approach". *International Journal of Lexicography*, 15(3), 206-228.
- Carr, M. (1997). "Internet Dictionaries and Lexicography". *International Journal of Lexicography*, 10(3), 209-230.
- Casares, J. (1941). "Nuevo concepto del diccionario de la lengua". En J. Casares (ed.), *Nuevo concepto del diccionario de la lengua y otros problemas de lexicografía y gramática*. Madrid: Espasa-Calpe.
- Casares, J. (1942) *Diccionario ideológico de la lengua española* (2.<sup>a</sup> ed.). Barcelona: Gustavo Gili. (Reprinted from: 1999).
- Castagnoli, S. (2008). "Corpus et bases de données terminologiques: l'interpretation au service des usagers". En F. Maniez, P. Dury, N. Arlin y C. Rougemont (eds.), *Corpus et dictionnaires de langues de spécialité*. Bresson: Presses Universitaires de Grenoble, 213-230.
- Castrillón Cardona, E. R., Patiño García, P. y Plested Álvarez, M. C. (2005). "La informática al servicio de la traducción". *Lenguaje*, 33, 343-358.
- Clark, R., Rothwell, A. y Shuttleworth, M. (2002). "Integrating Language Technology into a Postgraduate Translation Programme". En B. Maia, J. Haller y M. Ulyrch (eds.), *Training the Language Services Provider for the New Millennium*. Porto: Universidade do Porto, 63-71.
- Codina, L. (1995). "La ciber-ingenuidad: algunos mitos relacionados con la sociedad de la información". *El profesional de la información*, 35. Disponible en <http://www.elprofesionaldelainformacion.com/contenidos/1995/junio/index.html> [Consulta: 8 de mayo de 2012].
- Coffey, S. (2002). "Using a Source Language Corpus in Translator Training. A resource for teaching and testing activities". *InTRAlinea*, (Special Issue CULT2K). Disponible en [http://www.intralea.it/specials/cult2k/eng\\_open.php](http://www.intralea.it/specials/cult2k/eng_open.php) [Consulta: 8 de mayo de 2012].
- Colominas, C. (2004). "Los corpora como herramientas de traducción". En M. A. García Peinado y E. Ortega Arjonilla (eds.), *Panorama actual de la investigación en Traducción e Interpretación*. Granada: Atrio, 362-372.
- Condamines, A. y Rebeyrolle, J. (2000). "Construction d'une base de connaissances terminologiques à partir de textes : expérimentation et définition d'une méthode".

- En J. Charlet, M. Zacklad, G. Kassel y D. Bourigault (eds.), *Ingénierie des connaissances, évolutions récentes et nouveaux défis*. Paris: Eyrolles, 225-241.
- Corpas Pastor, G. (2001a). "Compilación de un corpus ad hoc para la enseñanza de la traducción inversa especializada". *TRANS. Revista de traductología*, 5, 155-184.
- Corpas Pastor, G. (2001b). "En torno al concepto de colocación". *Euskera*, XLVI, 89-108. Disponible en <http://www.euskaltzaindia.net/dok/euskera/11643.pdf> [Consulta: 8 de mayo de 2012].
- Corpas Pastor, G. (2002). "Traducir con corpus: de la teoría a la práctica". En J. García Palacios y M. T. Fuentes Morán (eds.), *Texto, terminología y traducción*. Salamanca: Almar, 189-226.
- Corpas Pastor, G. (2004). "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.), *Manual de documentación y terminología para la traducción especializada*. Madrid: Arcos/Libros, 223-274.
- Corpas Pastor, G. (2008). *Investigar con corpus en traducción: los retos de un nuevo paradigma*. Frankfurt: Peter Lang.
- Corpas Pastor, G., 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. Ayala Castro (ed.), *Diccionarios y enseñanza*. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá, 239-273.
- Corpas Pastor, G. y Seghiri, 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*. Amsterdam/Philadelphia: John Benjamins, 75-108.
- Corripio, F. (1983) *Diccionario de ideas afines*. Barcelona: Herder.
- Corris, M., Manning, C., Poetsch, S. y Simpson, J. (2000). "Bilingual Dictionaries for Australian Languages: User studies on the place of paper and electronic dictionaries". En U. Heid, S. Evert, E. Lehmann y C. Rohrer (eds.), *9th EURALEX International Congress*, Stuttgart, 8-12 de agosto de 2000, 169-181. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2000/](http://www.euralex.org/elx_proceedings/Euralex2000/) [Consulta: 8 de mayo de 2012].
- Cowie, A. P. (1999). *English Dictionaries for Foreign Learners - A History*. Oxford: Clarendon Press.
- Cruse, D. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Champe, G. G. (1997). "Project-Integrated Terminology Management for Technical Writing and Translation". En S. E. Wright y G. Budin (eds.), *Handbook of Terminology Management. Application-Oriented Terminology Management* (Vol. 2). Amsterdam/Philadelphia: John Benjamins, 503-515.
- Charmaz, K. (2006). *Constructing Grounded Theory. A Practical Guide Through Qualitative Analysis*. London: Sage Publications.
- Christ, O., Schulze, B. M., Hofmann, A. y König, E. (1999). "The IMS Corpus Workbench: Corpus Query Processor (CQP). User's Manual". Disponible en <http://www.ims.uni-stuttgart.de/projekte/CorpusWorkbench/CQPUserManual/HTML/> [Consulta: 8 de mayo de 2012].
- Church, K. W. (2008). "Approximate Lexicography and Web Search". *International Journal of Lexicography*, 21(3), 325-336.
- Danielsson, P. y Ridings, D. (1996). "Corpus and Terminology: Software for the Translation Program at Göteborgs Universitet or Getting Students to do the Work". En S. Botley, J. Glass, T. McEnery y A. Wilson (eds.), *Proceedings of Teaching*

- and Language Corpora (TALC 1996)* (Vol. Technical Papers 9), University Centre for Computer Corpus Research on Language (UCREL), 65-72.
- Davies, M. (2005). "The advantage of using relational databases for large corpora. Speed, advanced queries, and unlimited annotation". *International Journal of Corpus Linguistics*, 10(3), 307-334.
- De la Rosa Martín, R. (2003). "Los estudiantes de traducción y el uso de los diccionarios". *Interlingüística*, 14, 289-304.
- De Schryver, G.-M. (2003). "Lexicographers' Dreams in the Electronic-Dictionary Age". *International Journal of Lexicography*, 16(2), 143-199.
- De Schryver, G.-M. (2009). "State-of-the-Art Software to Support Intelligent Lexicography". En R. Zhu (ed.), *Proceedings of the International Seminar on Kangxi Dictionary & Lexicology*, Beijing Normal University. Disponible en <http://www.hcxf.cn/read.asp?id=570> [Consulta: 8 de mayo de 2012].
- De Schryver, G.-M. y Joffe, D. (2004). "On How Electronic Dictionaries are Really Used". En G. Williams y S. Vessier (eds.), *11th EURALEX International Congress*, Lorient, 6-10 de julio de 2004, 187-196. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2004/](http://www.euralex.org/elx_proceedings/Euralex2004/) [Consulta: 8 de mayo de 2012].
- De Schryver, G.-M. y Prinsloo, D. J. (2000). "Dictionary-Making Process with 'Simultaneous Feedback' from the Target Users to the Compilers". En U. Heid, S. Evert, E. Lehmann y C. Rohrer (eds.), *9th EURALEX International Congress*, Stuttgart, 8-12 de agosto de 2000, 197-209. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2000/](http://www.euralex.org/elx_proceedings/Euralex2000/) [Consulta: 8 de mayo de 2012].
- Delahaut, P. y Demaecker, C. (2002). "La traduction scientifique et technique: ses particularités et sa didactique". En B. Maia, J. Haller y M. Ulrych (eds.), *Training the Language Services Provider for the New Millennium*. Porto: Universidade do Porto, 293-304.
- Delavigne, V. (2008). "Construire un dictionnaire d'oncologie pour les patients: aspects méthodologiques". En F. Maniez, P. Dury, N. Arlin y C. Rougemont (eds.), *Corpus et dictionnaires de langues de spécialité*. Bresson: Presses Universitaires de Grenoble.
- Delisle, J. (1980). *L'Analyse du discours comme méthode de traduction : initiation à la traduction française de textes pragmatiques anglais : théorie et pratique*. Ottawa: Éditions de l'Université d'Ottawa.
- Désilets, A., Barrière, C. y Quirion, J. (2007). "Making Wikimedia Resources more Useful for Translators". *NRC Publications Archive*. Disponible en <http://nparc.cisti-icist.nrc-cnrc.gc.ca/npsi/ctrl?lang=en> [Consulta: 8 de mayo de 2012].
- Diab, T. A. y Hamdan, J. M. (1999). "Interacting with Words and Dictionaries: The Case of Jordanian EFL Learners". *International Journal of Lexicography*, 12(4), 281-305.
- Dickens, A. y Salkie, R. (1996). "Comparing Bilingual Dictionaries with a Parallel Corpus". En M. Gellerstam, J. Järborg, S.-G. Malmgren, K. Norén, L. Rogström y C. R. Pappmehl (eds.), *7th EURALEX International Congress (Proceedings Part II)*, Göteborg, 13-18 de agosto de 1996, 551-559. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex1996\\_2/](http://www.euralex.org/elx_proceedings/Euralex1996_2/) [Consulta: 8 de mayo de 2012].
- Díez Orzas, P. (2004). "Integración de modelos léxicos en lingüística computacional y traducción". En C. Gonzalo García y V. García Yebra (eds.), *Manual de*



- documentación y terminología para la traducción especializada*. Madrid: Arcos/Libros, 421-444.
- Dodd, W. S. (1989). "Lexicomputing and the dictionary of the future". En G. James (ed.), *Lexicographers and Their Works. (Exeter Linguistic Studies 14.)*. Exeter: Exeter University Press, 83-93.
- Dornseiff, F. (1933) *Der deutsche Wortschatz nach Sachgruppen*. Berlin: De Gruyter.
- Dung, J. (2009). "Online Dictionaries in a Web 2.0 Environment". En H. Bergenholtz, S. Nielsen y S. Tarp (eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 135-164.
- Enríquez Raído, V. (2011). *Investigating the Web Search Behaviors of Translation Students: An Exploratory and Multiple-Case Study*. (Tesis doctoral, Doctor of Philosophy), Universitat Ramon Llull.
- Ericsson, K. A. y Simon, H. A. (1984). *Protocol Analysis: Verbal Reports as Data*. Cambridge: MIT Press.
- Estella Clota, M. (2010). "La traducció xarxa: utopies i distopies". *Tradumàtica*, 8. Disponible en <http://www.fti.uab.es/tradumatica/revista/> [Consulta: 8 de mayo de 2012].
- 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, Programa de doctorado Lingüística Aplicada (lèxic y discurs)), Universitat Pompeu Fabra.
- Faber, P. (2002a). "Investigar en terminología". En P. Faber y C. Jiménez Hurtado (eds.), *Investigar en terminología*. Granada: Comares, 3-25.
- Faber, P. (2002b). "Oncoterm: Sistema bilingüe de información y recursos oncológicos". En A. Alcina y S. Gamero (eds.), *La traducción científico-técnica y la terminología en la sociedad de la información*. Castellón de la Plana: Publicaciones de la Universitat Jaume I, 177-188.
- Faber, P. (2003). "Terminological Competence and Enhanced Knowledge Acquisition". *Research in Language*, 1, 95-117. Disponible en <http://lexicon.ugr.es/pdf/faber2003.pdf> [Consulta: 8 de mayo de 2012].
- Faber, P. (2005). "La traducción del discurso científico y su terminología". *Lección inaugural. Apertura del curso académico 2005-2006*, Granada, 1-38. Disponible en <http://www.ugr.es/discurso-faber.pdf> [Consulta: 8 de mayo de 2012].
- Faber, P. (2009). "Terminología multilingüe y ontologías". En A. Alcina, E. Valero y E. Rambla (eds.), *Terminología y sociedad del conocimiento*. Bern: Peter Lang, 289-308.
- Faber, P. (2010). "Terminología, traducción especializada y adquisición de conocimiento". En E. Alarcón Navío (ed.), *La traducción en contextos especializados. Propuestas didácticas*. Granada: Atrio, 87-96.
- Faber, P., León Araúz, P. y Pérez Hernández, C. (2010). "A Wine Dictionary for Non-Drinkers: LSP Dictionary Functions". En M. Ibáñez Rodríguez (ed.), *Vino, lengua y traducción* (Vol. 2). Valladolid: Universidad de Valladolid, Secretariado de Publicaciones e Intercambio Científico, 109-119.
- Faber, P., León Araúz, P. y Prieto Velasco, J. A. (2009). "Semantic Relations, Dynamicity, and Terminological Knowledge Bases". *Current Issues in Language Studies*, 1(1), 1-23. Disponible en <http://www.academicpress.us/> [Consulta: 8 de mayo de 2012].
- Faber, P., 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(1), 39-65.

- Faber, P., López Rodríguez, C. I. y Tercedor, M. (2001). "Utilización de técnicas de corpus en la representación del conocimiento médico". *Terminology*, 7(2), 167-197.
- Faber, P., Márquez Linares, C. y Vega Expósito, M. (2005). "Framing Terminology: A Process-Oriented Approach". *Meta*, 50(4). Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Faber, P., Montero Martínez, S., Castro Prieto, M. R., Senso Ruiz, J., Prieto Velasco, J. A., León Araúz, P., Márquez Linares, C. y Vega Expósito, M. (2006). "Process-oriented terminology management in the domain of Coastal Engineering". *Terminology*, 12(2), 189-213.
- Fata, I. (2010). "The Bilingual Specialised Translation Dictionary for Learners". En P. A. Fuertes-Olivera (ed.), *Specialized Dictionaries for Learners*. Berlin: De Gruyter, 83-100.
- Feliu, J. (2004). *Relacions conceptuals i terminologia: anàlisi i proposta de detecció semiautomàtica*. (Tesis doctoral, Programa de doctorado Lingüística Aplicada (lèxic)), Universitat Pompeu Fabra.
- Feliu, J., Vivaldi, J. y Cabré, M. T. (2002). "Towards an Ontology for a Human Genome Knowledge Base". En E. Yuste (ed.), *3rd International Conference on Language Resources and Evaluation (LREC 2002)*, Las Palmas de Gran Canaria, 1885-1890. Disponible en <http://www.lrec-conf.org/proceedings/lrec2002/> [Consulta: 8 de mayo de 2012].
- Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press.
- Fenner, A. (1989). "Techniques, Presentation and Specifications". En C. Picken (ed.), *The Translator's Handbook*. London: Aslib, 43-58.
- Fernández-Pampillón Cesteros, A. y Matesanz del Barrio, M. (2006). "Los diccionarios electrónicos: hacia un nuevo concepto de diccionario". *Estudios de Lingüística del Español (ELiEs)*, 24(Nuevos géneros discursivos: los textos electrónicos). Disponible en <http://elies.rediris.es/elies24/> [Consulta: 8 de mayo de 2012].
- Ferraresi, A. (2009). "Google and Beyond: Web-As-Corpus Methodologies for Translators". *Tradumàtica*, 7, 1-8. Disponible en <http://www.fti.uab.cat/tradumatica/revista/num7/sumari.htm> [Consulta: 8 de mayo de 2012].
- Fillmore, C. J. (1985). "Frames and the Semantics of Understanding". *Quaderni di Semantica*, 6, 222-253.
- Fillmore, C. J. y Atkins, B. T. S. (1998). "FrameNet and Lexicographic Relevance". En A. Rubio, N. Gallardo, R. Castro y A. Tejada (eds.), *Proceedings of the ELRA Conference on Linguistic Resources*, Granada, 417-423.
- Fillmore, C. J., Baker, C. F. y Sato, H. (2002). "The FrameNet Database and Software Tools". En M. González Rodríguez y C. P. Suárez Araujo (eds.), *3rd International Conference on Language Resources and Evaluation (LREC 2002)*, Las Palmas de Gran Canaria, 1157-1160. Disponible en <http://www.lrec-conf.org/proceedings/lrec2002/> [Consulta: 8 de mayo de 2012].
- Fillmore, C. J., Johnson, C. R. y Petruck, M. R. L. (2003). "Background to FrameNet". *International Journal of Lexicography*, 16(3), 235-250.
- Fletcher, W. H. (2007a). "Concordancing the Web: Promise and Problems, Tools and Techniques". En M. Hundt, N. Nesselhauf y C. Biewer (eds.), *Corpus Linguistics and the Web*. Amsterdam: Rodopi, 25-46.
- Fletcher, W. H. (2007b). "Implementing a BNC-Compare-able Web Corpus". En C. Fairon, H. Naets, A. Kilgarriff y G.-M. De Schryver (eds.), *Building and Exploring Web Corpora. Proceedings of the 3rd Web As Corpus Workshop, Incorporating Cleaneval. Cahiers du Cental 4*, Louvain-la-Neuve, 43-56.

- Flórez, S. y Alcina, A. (2011). "Free and Open Source Software for the Translation Classroom: a Catalogue of Available Tools". *The Interpreter and Translator Trainer*, 5(2), 325-357.
- Forget, N. (1999). *Les dictionnaires électroniques dans l'optique de la traduction*. (Trabajo de investigación, École de traduction et d'interprétation), Université d'Ottawa, Ottawa. Disponible en <http://aix1.uottawa.ca/~etithese/nforg/> [Consulta: 8 de mayo de 2012].
- Frankenberg-Garcia, A. (2002). "COMPARA - language learning and translation training". En B. Maia, J. Haller y M. Ulrych (eds.), *Training the Language Services Provider for the New Millennium*. Porto: Universidade do Porto, 187-199.
- Frankenberg-Garcia, A. y Santos, D. (2003). "Introducing *Compara* the Portuguese-English Parallel Corpus". En F. Zanettin, S. Bernardini y D. Stewart (eds.), *Corpora in Translator Education*. Manchester: St. Jerome, 71-88.
- Fraser, J. (1993). "Public Accounts: Using Verbal Protocols to Investigate Community Translation". *Applied Linguistics*, 14(4), 325-343.
- Fraser, J. (1994). "Translating Practice into Theory: A Practical Study of Quality in Translator Training". En C. Picken (ed.), *ITI Conference 7 Proceedings*. London: ITI, 130-142.
- Fraser, J. (1999). "The Translator and the Word: The Pros and Cons of Dictionaries in Translation". En G. Anderman y M. Rogers (eds.), *Word, Text, Translation. Liber Amicorum for Peter Newmark*. Great Britain: Multilingual Matters, 25-34.
- Friedbichler, I. y Friedbichler, M. (2000). "The potential of domain-specific target-language corpora for the translator's workbench". En S. Bernardini y F. Zanettin (eds.), *I corpora nella didattica della traduzione - Corpus use and learning to translate*. Bologna: CLUEB, 107-116.
- Fuertes-Olivera, P. A. (2009a). "The Function Theory of Lexicography and Electronic Dictionaries: WIKTIONARY as a Prototype of Collective Free Multiple-Language Internet Dictionary" *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 99-134.
- Fuertes-Olivera, P. A. (2009b). "Systematic introductions in specialised dictionaries: Some proposals in relation to accounting dictionaries". En S. Nielsen y S. Tarp (eds.), *Lexicography in the 21st Century. In honour of Henning Bergenholtz*. Amsterdam/Philadelphia: John Benjamins, 161-178.
- Fuertes-Olivera, P. A. (2010). "Lexicography for the Third Millennium: Free Institutional Internet Terminological Dictionaries for Learners". En P. A. Fuertes-Olivera (ed.), *Specialized Dictionaries for Learners*. Berlin: De Gruyter, 193-210.
- Fuertes-Olivera, P. A. y Arribas-Baño, A. (2008). *Pedagogical Specialized Lexicography*. Amsterdam/Philadelphia: John Benjamins.
- Fuertes-Olivera, P. A. y Bergenholtz, H. (2011). "Introduction: The Construction of Internet Dictionaries". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lxicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 1-16.
- Fuertes-Olivera, P. A. y Nielsen, S. (2012). "Online Dictionaries for Assisting Translators of Lsp Texts: The Accounting Dictionaries". *International Journal of Lexicography*, 25(2), 191-215.
- Fuertes-Olivera, P. A. y Niño-Amo, M. (2011). "Internet Dictionaries for Communicative and Cognitive Functions: *El Diccionario Inglés-Español de Contabilidad*". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lxicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 168-170.

- Fulford, H. (2001). "Exploring terms and their linguistic environment in text". *Terminology*, 7(2), 259-279.
- Gala, N. y Lafourcade, M. (2011). "NLP lexicons: innovative constructions and usages for machines and humans". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 89-95. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Galisson, R. (1983). "Image et usage du dictionnaire chez des étudiants (en langue) de niveau avancé". *Études de Linguistique Appliquée*, 49, 5-88.
- Gallardo San Salvador, N. y de Irazazábal, A. (2002). "Elaboración de un vocabulario multilingüe del campo temático de la siderurgia". En A. Alcina y S. Gamero (eds.), *La traducción científico-técnica y la terminología en la sociedad de la información*. Castellón de la Plana: Publicaciones de la Universitat Jaume I, 189-198.
- Gambín, J. (2004). "A Practical Approach to Terminology Work". *ATA Seminar: Translating for the Pharmaceutical Industry*, San Juan, Puerto Rico. Disponible en <http://www.abroadlink.com/Articles/A%20Practical%20Approach%20to%20Terminology%20Work.pdf> [Consulta: 8 de mayo de 2012].
- Gamero, S. (2001). *La traducción de textos técnicos*. Barcelona: Ariel Lenguas Modernas.
- García Figuerola, C. (1999). "La recuperación de la información en colecciones documentales multilingües". En M. Pinto y J. A. Cordon (eds.), *Técnicas documentales aplicadas a la traducción*. Madrid: Síntesis, 129-142.
- García Yebra, V. (1984). *Teoría y práctica de la traducción*. Madrid: Gredos.
- Gasiglia, N. y Paulussen, H. (2008). "De la création d'un corpus bilingue du tourisme à partir du Web à son exploration avec ParaSearch et Unitex pour la documentation des lexicographes". En F. Maniez, P. Dury, N. Arlin y C. Rougemont (eds.), *Corpus et dictionnaires de langues de spécialité*. Bresson: Presses Universitaires de Grenoble, 231-270.
- Gavioli, L. y Zanettin, F. (2000). "I corpora bilingui nell'apprendimento della traduzione. Riflessioni su un'esperienza pedagogica". En S. Bernardini y F. Zanettin (eds.), *I corpora nella didattica della traduzione - Corpus use and learning to translate*. Bologna: CLUEB, 61-80.
- Geeraerts, D. (2000). "Adding Electronic Value. The electronic version of the Grote Van Dale". En U. Heid, S. Evert, E. Lehmann y C. Rohrer (eds.), *9th EURALEX International Congress*, Stuttgart, 8-12 de agosto de 2000, 75-84. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2000/](http://www.euralex.org/elx_proceedings/Euralex2000/) [Consulta: 8 de mayo de 2012].
- Gerloff, P. (1986). "Second Language Learners' Reports on the Interpretive Process: Talk-aloud Protocols of Translation". En J. House y S. Blum-Kulka (eds.), *Interlingual and Intercultural Communication. Discourse and Cognition in Translation and Second Language Acquisition Studies*. Tübingen: Gunter Narr, 243-262.
- Gerzymisch-Arbogast, H. (1996). *Termini im Kontext. Verfahren zur Erschließung und Übersetzung der textspezifischen Bedeutung von fachsprachlichen Ausdrücken*. Tübingen: Gunter Narr.
- Gietz, R. A. (1990). "La información en la traducción técnico científica. Perspectivas de la T.A.C. y de la T.A." *Actas del II Simposio de RITerm*. Disponible en <http://www.riterm.net/actes/2simposio/gietz.htm#2> [Consulta: 8 de mayo de 2012].
- Gile, D. (2004). "Integrated Problem and Decision Reporting as a Translator Training Tool". *JoSTrans. The Journal of Specialised Translation*, (2). Disponible en [http://www.jostrans.org/issue02/art\\_gile.php](http://www.jostrans.org/issue02/art_gile.php) [Consulta: 8 de mayo de 2012].



- Glaser, B. G. (2001). *The Grounded Theory Perspective III: Theoretical Coding*. California: Sociology Press.
- Glaser, B. G. y Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Publishing Company.
- Gómez González-Jover, A. (2005). *Terminografía, lenguajes profesionales y mediación interlingüística. Aplicación metodológica al léxico especializado del sector industrial del calzado y de las industrias afines*. (Tesis doctoral), Universidad de Alicante. Disponible en <http://descargas.cervantesvirtual.com/servlet/SirveObras/24605018878048721976613/022151.pdf> [Consulta: 8 de mayo de 2012].
- Gómez, R. (1999). "Internet en la traducción. Ideas sobre el uso de Internet en la traducción del inglés al español". *Punto y coma*, 57. Disponible en <http://ec.europa.eu/translation/bulletins/puntoycoma/57/pyc577.htm> [Consulta: 8 de mayo de 2012].
- Gonzalo García, C. (2004). "Fuentes de información en línea para la traducción especializada". En C. Gonzalo García y V. García Yebra (eds.), *Manual de documentación y terminología para la traducción especializada*. Madrid: Arcos/Libros, 275-307.
- Gonzalo García, C. 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.), *Manual de documentación y terminología para la traducción especializada*. Madrid: Arcos/Libros, 335-360.
- Göpferich, S. (2009). "Towards a model of translation competence and its acquisition: the longitudinal study *Transcomp*". En S. Göpferich, A. L. Jakobsen y I. M. Mees (eds.), *Behind the mind. Methods, models and results in translation process research*. Copenhagen: Samfundslitteratur, 11-38.
- Göpferich, S. y Jääskeläinen, R. (2009). "Process research into the development of translation competence: Where are we, and where do we need to go?" *Across Languages and Cultures*, 10(2), 169-191. Disponible en [http://www.uni-saarland.de/fileadmin/user\\_upload/Professoren/fr46\\_ProfGerzymisch-Arbogast/AS/G%C3%B6pferich\\_J%C3%A4skel%C3%A4inen.pdf](http://www.uni-saarland.de/fileadmin/user_upload/Professoren/fr46_ProfGerzymisch-Arbogast/AS/G%C3%B6pferich_J%C3%A4skel%C3%A4inen.pdf) [Consulta: 8 de mayo de 2012].
- Gorden, R. L. (1975). *Interviewing: Strategy, Techniques, and Tactics*. Homewood (Illinois): Dorsey Press.
- Gouws, R. H. (2009). "Dictionaries as Innovative Tools in a New Perspective on Standardisation". En H. Bergenholtz, S. Nielsen y S. Tarp (eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 265-283.
- Gouws, R. H. (2010). "The Monolingual Specialised Dictionary for Learners". En P. A. Fuertes-Olivera (ed.), *Specialized Dictionaries for Learners*. Berlin: De Gruyter, 55-68.
- Gouws, R. H. (2011). "Learning, Unlearning and Innovation in the Planning of Electronic Dictionaries". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lxicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 17-29.
- Granger, S. (2003). "The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies?" En S. Granger, J. Lerot y S. Petch-Tyson (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam/New York: Rodopi, 17-30.
- Greenbaum, S., Meyer, C. F. y Taylor, J. (1984). "The image of the dictionary for American college students". *Dictionaries*, 6, 31-52.

- Grefenstette, G. (1998). "The Future of Linguistics and Lexicographers: Will there be Lexicographers in the Year 3000?" En T. Fontenelle, P. Hilgsmann, A. Michiels, A. Moulin y S. Theissen (eds.), *8th EURALEX International Congress (Proceedings Part I)*, Liège, 4-8 de agosto de 1998, 25-41. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex1998\\_1/](http://www.euralex.org/elx_proceedings/Euralex1998_1/) [Consulta: 8 de mayo de 2012].
- Grönroos, M. (2005). "Bringing Intelligence to Translation Memory Technology". *Translating and the Computer*, 27. Disponible en <http://www.mt-archive.info/Aslib-2005-Gronroos.pdf> [Consulta: 8 de mayo de 2012].
- Gross, G. (1997). "La grammaire, les dictionnaires et l'informatique". En J. Pruvost (ed.), *Les dictionnaires de la langue française et l'informatique. Actes du colloque la Journée des dictionnaires*, Cergy-Pontoise, 55-64.
- Hadouche, F., L'Homme, M.-C., Lapalme, G. y Le Serrec, A. (2009). "Intégration d'informations syntactico-sémantiques dans les bases de données terminologiques : méthodologie d'annotation et perspectives d'automatisation". En A. Alcina y M.-C. L'Homme (eds.), *1st International Workshop on Terminology and Lexical Semantics (TLS 2009)*, Université de Montréal, 22-31. Disponible en <http://olst.ling.umontreal.ca/pdf/ProceedingsTLS09.pdf> [Consulta: 8 de mayo de 2012].
- Halskov, J. y Barrière, C. (2008). "Web-based extraction of semantic relation instances for terminology work". *Terminology*, 14(1), 20-44.
- Hamon, T. y Nazarenko, A. (2001). "Detection of synonymy links between terms. Experiment and results". En D. Bourigault, C. Jacquemin y M.-C. L'Homme (eds.), *Recent Advances in Computational Terminology*. Amsterdam/Philadelphia: John Benjamins, 185-208.
- Hansen, G. (1999). *Probing the Process in Translation: Methods and Results*. Copenhagen: Samfundslitteratur.
- Hansen, G. (2002). *Empirical Translation Studies: Process and Product*. Copenhagen: Samfundslitteratur.
- Hansen, S. y Teich, E. (2002). "The creation and exploitation of a translation reference corpus". En E. Yuste (ed.), *1st International Workshop on Language Resources for Translation Work and Research (LREC 2002 Workshop Proceedings)*, Las Palmas de Gran Canaria, 1-4.
- Harley, A. (2000). "Software Demonstration: Cambridge Dictionaries Online". En U. Heid, S. Evert, E. Lehmann y C. Rohrer (eds.), *9th EURALEX International Congress*, Stuttgart, 8-12 de agosto de 2000, 85-88. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2000/](http://www.euralex.org/elx_proceedings/Euralex2000/) [Consulta: 8 de mayo de 2012].
- Harman, N. (2002). "Beyond Paper Dictionaries Mining the Web for Technical Terminology in Chinese". *Paper prepared for the Translation Teacher's Certificate of the Consortium for the Training of Translation Teachers*, Imperial College, University of London.
- Hartmann, R. R. K. (1987). "Four Perspectives on Dictionary Use: A Critical Review of Research Methods". En A. P. Cowie (ed.), *The Dictionary and the Language Learner. Papers of the EURALEX Seminar at the University of Leeds, 1-3 April 1985*. Tübingen: Niemeyer, 11-28.
- Hartmann, R. R. K. (1989a). "The Dictionary as an Aid to Foreign-Language Teaching". En F. J. Hausmann, O. Reichmann, E. Wiegand y L. Zgusta (eds.), *Wörterbücher/Dictionaries/Dictionnaires, Ein internationales Handbuch zur*

- Lexikographie/An International Encyclopedia of Lexicography/Encyclopédie internationale de lexicographie*. Berlin: de Gruyter, 181-189.
- Hartmann, R. R. K. (1989b). "Lexicography, Translation and the So-Called Language Barrier". En M. Snell-Hornby y E. Pöhl (eds.), *Translation and Lexicography. Papers read at the Euralex Colloquium held at Innsbruck 2-5 July 1987*. Amsterdam/Philadelphia: John Benjamins, 9-20.
- Hartmann, R. R. K. (1989c). "Sociology of the Dictionary User: Hypotheses and Empirical Studies". En F. J. Hausmann, O. Reichmann, E. Wiegand y L. Zgusta (eds.), *Wörterbücher/Dictionaries/Dictionnaires, Ein internationales Handbuch zur Lexikographie/An International Encyclopedia of Lexicography/Encyclopédie internationale de lexicographie* (Vol. 1). Berlin: de Gruyter, 102-111.
- Hartmann, R. R. K. (1999). "Thematic Report 2. Case Study: The Exeter University Survey of Dictionary Use". En R. R. K. Hartmann (ed.), *Dictionaries in Language Learning. Recommendations, National Reports and Thematic Reports from the TNP Sub-Project 9: Dictionaries*. Berlin: Freie Universität Berlin, 36-52.
- Hartmann, R. R. K. (2005). "Onomasiological Dictionaries in 20<sup>th</sup>-Century Europe". *Lexicographica*, 21(Thematic Part Onomasiological dictionaries of European languages in the past and present), 6-19.
- Harvey Ciampi, T. (2004). "WWW Search Interfaces for Translators. The Internet - A powerful terminology search tool for translators". *Hieronymus*, 3. Disponible en <http://www.tradulex.org/> [Consulta: 8 de mayo de 2012].
- Heid, U. (2011). "Electronic Dictionaries as Tools: Toward an Assessment of Usability". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 287-304.
- Hull, D. A. (2001). "Software tools to support the construction of bilingual terminology lexicons". En D. Bourigault, C. Jacquemin y M.-C. L'Homme (eds.), *Recent Advances in Computational Terminology*. Amsterdam/Philadelphia: John Benjamins, 225-244.
- Hüllen, W. (2005). "Thematic Part Onomasiological dictionaries of European languages in the past and present. Introduction". *Lexicographica*, 21, 1-5.
- Hurskainen, A. (2008). *SALAMA Dictionary Compiler - A Method for Corpus-Based Dictionary Compilation* (Technical Reports in Language Technology. Report No 2). Disponible en <http://www.njas.helsinki.fi/salama/salama-dictionary-compiler.pdf> [Consulta: 8 de mayo de 2012].
- Hurtado Albir, A. (2001). *Traducción y traductología. Introducción a la traductología*. Madrid: Cátedra.
- Ide, K. (1993). "A Catalogue of Electronic Dictionaries". *Language*, 22(5), 42-49.
- Ivanova, V. y Vilar Sánchez, K. (2008). "El análisis cualitativo de datos aplicado al estudio lingüístico contrastivo: recursos para indicar la fuente normativa en contratos de trabajo". En E. Tabares Plasencia, V. Ivanova y E. Krüger (eds.), *Análisis lingüístico contrastivo de textos especializados en español y alemán*. Berlin: Frank & Timme, 87-104.
- Jääskeläinen, R. (1989). "The Role of Reference Material in Professional vs. Non-Professional Translation: A Think-Aloud Protocol Study". En S. Tirkkonen-Condit y S. Condit (eds.), *Empirical Studies in Translation and Linguistics*. Joensuu: University of Joensuu, 175-198.
- Jääskeläinen, R. (1993). "Investigating Translation Strategies". En S. Tirkkonen-Condit y J. Laffling (eds.), *Recent Trends in Empirical Translation Research*. Joensuu: University of Joensuu, 99-120.

- Jääskeläinen, R. (1996). "Hard Work Will Bear Beautiful Fruit. A Comparison of Two Think-Aloud Protocol Studies". *Meta*, 41(1), 60-74. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Jääskeläinen, R. y Mauranen, A. (2004). "Translators at work: a case study of electronic tools used by translators in industry". En G. Barnbrook, P. Danielsson y M. Mahlberg (eds.), *Meaningful Texts: The Extraction of Semantic Information from Monolingual and Multilingual Corpora*. London: Continuum, 48-53.
- Jääskeläinen, R. y Tirkkonen-Condit, S. (1991). "Automatised Processes in Professional vs. Non-professional Translation: A Thinking-Aloud Protocol Study". En S. Tirkkonen-Condit (ed.), *Empirical Research in Translation and Intercultural Studies*. Tübingen: Gunter Narr, 89-110.
- Jacquet-Pfau, C. (2002). "Les dictionnaires du français sur cédérom". *International Journal of Lexicography*, 15(1), 89-104.
- Jakubíček, M., Kilgarriff, A., McCarthy, D. y Rychlý, P. (2010). "Fast Syntactic Searching in Very Large Corpora for Many Languages". En R. Otoguro, K. Ishikawa, H. Umemoto, K. Yoshimoto y Y. Harada (eds.), *Proceedings of Workshop on Advanced Corpus Solutions* (Vol. PACLIC 24), Tohoku University (Japan), 741-748. Disponible en <http://www.dianamccarthy.co.uk/papers/PACLIC2010corpWS.pdf> [Consulta: 8 de mayo de 2012].
- Jensen, A. (1999). "Time Pressure in Translation". En G. Hansen (ed.), *Probing the Process in Translation: Methods and Results*. Copenhagen: Samfundslitteratur, 103-119.
- Jensen, A. y Jakobsen, A. L. (2000). "Translating under Time Pressure: An Empirical Investigation of Problem-solving Activity and Translation Strategies by Non-professional and Professional Translators". En A. Chesterman, N. Gallardo San Salvador y Y. Gambier (eds.), *Translation in Context*. Amsterdam/Philadelphia: John Benjamins, 105-116.
- Jousse, A.-L., L'Homme, M.-C., Leroyer, P. y Robichaud, B. (2011). "Presenting Collocates in a Dictionary of Computing and the Internet according to User Needs". En I. Boguslavsky y L. Wanner (eds.), *Actes de la 5ème Conférence internationale sur la Théorie Sens-Texte (MTT 2011)*, Barcelona, 134-144. Disponible en [http://olst.ling.umontreal.ca/pdf/Jousse\\_et\\_al\\_2011.pdf](http://olst.ling.umontreal.ca/pdf/Jousse_et_al_2011.pdf) [Consulta: 8 de mayo de 2012].
- Jousse, A.-L. y Polguère, A. (2005). *LE DICO ET SA VERSION DICOUEBE. Document descriptif et manuel d'utilisation. Version du rapport 1.0*. Université de Montréal. Disponible en <http://idefix.ling.umontreal.ca/dicouebe/DiCoDOC.pdf> [Consulta: 8 de mayo de 2012].
- Jousse, A.-L., Polguère, A. y Tremblay, O. (2008). "Du dictionnaire au site lexical pour l'enseignement/apprentissage du vocabulaire". En F. Grossmann y S. Plane (eds.), *Lexique et production verbale. Vers une meilleure intégration des apprentissages lexicaux* (Vol. Éducation et didactiques). Villeneuve d'Ascq: Presses Universitaires du Septentrion, 141-157.
- Kaalep, H.-J. y Mikk, J. (2008). "Creating Specialised Dictionaries for Foreign Language Learners: A Case Study". *International Journal of Lexicography*, 21(4), 369-394.
- Kalliokuusi, V. y Varantola, K. (1998). "From general dictionaries to terminological glossaries. User expectations vs editorial aims". En T. Fontenelle, P. Hiligsmann, A. Michiels, A. Moulin y S. Theissen (eds.), *8th EURALEX International Congress (Proceedings Part II)*, Liège, 601-610. Disponible en



- [http://www.euralex.org/elx\\_proceedings/Euralex1998\\_2/](http://www.euralex.org/elx_proceedings/Euralex1998_2/) [Consulta: 8 de mayo de 2012].
- Kashyap, V., Bussler, C. y Moran, M. (2008). *The Semantic Web: Semantics for Data and Services on the Web*. Berlin: Springer.
- Kay, M. (1984). "The Dictionary Server". *10th International Conference on Computational Linguistics (COLING 1984)*, Stanford, 2-6 de julio de 1984, 461-461. Disponible en <http://aclweb.org/anthology-new/> [Consulta: 8 de mayo de 2012].
- Kharma, N. N. (1985). "Wanted: A Brand-New Type of Learners' Dictionary". *Multilingua*, 4(2), 85-90.
- Kilgarriff, A., P.V.S, A. y Pomikálek, J. (2011). "Comparable Corpora BootCat". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 122-128. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Kilgarriff, A., Rychly, P., Smrz, P. y Tugwell, D. (2004). "The Sketch Engine". En G. Williams y S. Vessier (eds.), *11th EURALEX International Congress*, Lorient, 6-10 de julio de 2004, 105-116. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2004/](http://www.euralex.org/elx_proceedings/Euralex2004/) [Consulta: 8 de mayo de 2012].
- King, P. (2003). "Parallel concordancing and its applications". En S. Granger, J. Lerot y S. Petch-Tyson (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam/New York: Rodopi, 157-168.
- Kipfer, B. A. (1987). "Dictionaries and the Intermediate Student: Communicative Needs and the Development of User Reference Skills". En A. P. Cowie (ed.), *The Dictionary and the Language Learner. Papers of the EURALEX Seminar at the University of Leeds, 1-3 April 1985*. Tübingen: Niemeyer, 44-54.
- Kiraly, D. C. (1995). *Pathways to Translation. Pedagogy and Process*. Kent (Ohio): The Kent State University Press.
- Kiraly, D. C. (2000). *A Social Constructivist Approach to Translator Education. Empowerment from Theory to Practice*. Manchester: St. Jerome.
- Knowles, F. E. (1989). "The Computer in Lexicography". En F. J. Hausmann, O. Reichmann, E. Wiegand y L. Zgusta (eds.), *Wörterbücher/Dictionaries/Dictionnaires, Ein internationales Handbuch zur Lexikographie/An International Encyclopedia of Lexicography/Encyclopédie internationale de lexicographie* (Vol. 1). Berlin: de Gruyter, 1645-1672.
- Koplenig, A. (2011). "Understanding How Users Evaluate Innovative Features of Online Dictionaries - An Experimental Approach". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 147-150. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Kosem, I. y Kosem, K. (2011). "Foreword". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 10-12 de noviembre de 2011, IV. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Kraif, O. (2008). "Extraction automatique de lexique bilingue: application pour la recherche d'exemples en lexicographie". En F. Maniez, P. Dury, N. Arlin y C.

- Rougemont (eds.), *Corpus et dictionnaires de langues de spécialité*. Bresson: Presses Universitaires de Grenoble, 67-87.
- Krings, H. P. (1986a). "Translation Problems and Translation Strategies of Advanced German Learners of French (L2)". En J. House y S. Blum-Kulka (eds.), *Interlingual and Intercultural Communication. Discourse and Cognition in Translation and Second Language Acquisition Studies*. Tübingen: Gunter Narr, 263-276.
- Krings, H. P. (1986b). *Was in den Köpfen von Übersetzern vorgeht. Eine empirische Untersuchung der Struktur des Übersetzungsprozesses an Fortgeschrittenen*. Tübingen: Gunter Narr.
- Kübler, N. (2003). "Corpora and LSP Translation". En F. Zanettin, S. Bernardini y D. Stewart (eds.), *Corpora in Translator Education*. Manchester: St. Jerome, 25-42.
- Kugler, M., Ahmad, K. y Thurmair, G. (1995). *Translator's Workbench. Tools and Terminology for Translation and Text Processing*. Berlin: Springer.
- Künzli, A. (2001). "Experts versus novices : l'utilisation de sources d'information pendant le processus de traduction". *Meta*, 46(3), 507-523. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Kussmaul, P. (1995). *Training the Translator*. Amsterdam/Philadelphia: John Benjamins.
- Kussmaul, P. y Tirkkonen-Condit, S. (1995). "Think-Aloud Protocol Analysis in Translation Studies". *TTR : traduction, terminologie, rédaction*, 8(1), 177-199. Disponible en <http://www.erudit.org/revue/TTR/1995/v8/n1/037201ar.pdf> [Consulta: 8 de mayo de 2012].
- Kvale, S. (1996). *Interviews: An Introduction to Qualitative Research Interviewing*. California: Sage Publications.
- Kvale, S. y Brinkmann, S. (2009). *Interviews. Learning the Craft of Qualitative Research Interviewing*. London: Sage Publications.
- L'Homme, M.-C. (2006). "The processing of terms in dictionaries: New models and techniques. A state of the art". *Terminology*, 12(2), 181-188.
- L'Homme, M.-C. (2008). "Le DiCoInfo. Méthodologie pour une nouvelle génération de dictionnaires spécialisés". *Traduire*, 217, 78-103.
- L'Homme, M.-C. y Leroyer, P. (2009). "Combining the semantics of collocations with situation-driven search paths in specialized dictionaries". *Terminology*, 15(2), 258-283.
- L'Homme, M.-C., Leroyer, P. y Robichaud, B. (2010). "Advanced Encoding for Multilingual Access in a Terminological Data Base – A Matter of Balance". *Terminology and Knowledge Engineering (TKE 2010)*, Dublin City University. Disponible en [http://olst.ling.umontreal.ca/pdf/TKE2010-lho-ler-rob%20\\_final.pdf](http://olst.ling.umontreal.ca/pdf/TKE2010-lho-ler-rob%20_final.pdf) [Consulta: 8 de mayo de 2012].
- Lagoudaki, E. (2006). *Translation Memory systems: Enlightening users' perspective*. Disponible en <http://www3.imperial.ac.uk/pls/portallive/docs/1/7307707.PDF> [Consulta: 15 de marzo de 2012].
- Larsen, L. L. (1997). "Information Literacy: The Web is not an Encyclopedia". *ACM*, 89-93. Disponible en <http://dl.acm.org/citation.cfm?id=262079> [Consulta: 8 de febrero de 2012].
- Laviosa, S. (1998). "The Corpus-Based Approach: A New Paradigm in Translation Studies". *Meta*, 43(4), 474-479. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Laviosa, S. (2002). *Corpus-based Translation Studies: Theory, Findings, Applications*. Amsterdam/New York: Rodopi.

- Laviosa, S. (2003a). "Corpora and the Translator". En H. Somers (ed.), *Computers and Translation. A Translator's Guide*. Amsterdam/Philadelphia: John Benjamins, 89-117.
- Laviosa, S. (2003b). "Corpora and Translation Studies". En S. Granger, J. Lerot y S. Petch-Tyson (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam/New York: Rodopi, 45-54.
- Le Poder, M.-E. (2004). "La Terminología en la Licenciatura en Traducción e Interpretación". En M. A. García Peinado y E. Ortega Arjonilla (eds.), *Panorama actual de la investigación en Traducción e Interpretación*. Granada: Atrio, 23-28.
- Lehr, A. (1996). "Electronic Dictionaries". *Lexicographica*, 12, 310-317.
- Leroyer, P. (2006). "Dealing with phraseology in business dictionaries: focus on dictionary functions – not phrases". *Linguistik online*, 27(2), 183-194. Disponible en [http://www.linguistik-online.de/27\\_06/leroyer.pdf](http://www.linguistik-online.de/27_06/leroyer.pdf) [Consulta: 8 de mayo de 2012].
- Leroyer, P. (2008). "La correction de la traduction L2>L1 comme fonction du dictionnaire spécialisé". En F. Maniez, P. Dury, N. Arlin y C. Rougemont (eds.), *Corpus et dictionnaires de langues de spécialité*. Bresson: Presses Universitaires de Grenoble, 111-133.
- Leroyer, P. (2009a). "Balancing the tools. The functional transformation of lexicographic tools for tourists". En S. Nielsen y S. Tarp (eds.), *Lexicography in the 21st Century. In honour of Henning Bergenholtz*. Amsterdam/Philadelphia: John Benjamins, 103-122.
- Leroyer, P. (2009b). "Lexicography Hits the Road: New Information Tools for Tourists". En H. Bergenholtz, S. Nielsen y S. Tarp (eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 285-310.
- Leroyer, P. (2011). "Change of Paradigm: From Linguistics to Information Science and from Dictionaries to Lexicographic Information Tools". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 121-140.
- Lew, R. (2011). "Online Dictionaries of English". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 230-250.
- Lew, R. y Mitton, R. (2011). "Not the Word I Wanted? How Online English Learners' Dictionaries Deal with Misspelled Words". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 165-174. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Lewis, R. B. (1997). "ATLAS/ti and NUD-IST: A Comparative Review of Two Leading Qualitative Data Analysis Packages". *Cultural Anthropology Methods*, 10(3), 41-47. Disponible en <http://www.umass.edu/accela/llc/794d/pdf/Atlas%20and%20Nudist%20comparison.pdf> [Consulta: 8 de mayo de 2012].
- Livbjerg, I. y Mees, I. M. (1999). "A Study of the Use of Dictionaries in Danish-English Translation". En G. Hansen (ed.), *Probing the Process in Translation: Methods and Results*. Copenhagen: Samfundlitteratur 135-149.
- Livbjerg, I. y Mees, I. M. (2003). "Patterns of dictionary use in non-domain-specific translation". En F. Alves (ed.), *Triangulating Translation*. Amsterdam/Philadelphia: John Benjamins, 123-136.

- López Rodríguez, C. I., Faber, P. y Tercedor, M. (2006a). "Terminología basada en el conocimiento para la traducción y la divulgación médicas: el caso de Oncoterm". *Panacea*, VII(24), 228-240. Disponible en [http://www.medtrad.org/panacea/IndiceGeneral/n24\\_tradyterm-l.rodriguez.etal.pdf](http://www.medtrad.org/panacea/IndiceGeneral/n24_tradyterm-l.rodriguez.etal.pdf) [Consulta: 8 de mayo de 2012].
- López Rodríguez, C. I. y Tercedor, M. (2008). "Corpora and Students' Autonomy in Scientific and Technical Translation Training". *JoSTrans. The Journal of Specialised Translation*, 9. Disponible en [http://www.jostrans.org/issue09/art\\_lopez\\_tercedor.pdf](http://www.jostrans.org/issue09/art_lopez_tercedor.pdf) [Consulta: 8 de mayo de 2012].
- López Rodríguez, C. I., Tercedor, M. y Faber, P. (2006b). "Gestión terminológica basada en el conocimiento y generación de recursos de información sobre el cáncer: el proyecto ONCOTERM". *Revistaesalud*, 2(8). Disponible en <http://www.revistaesalud.com/index.php/revistaesalud/article/view/127/325> [Consulta: 8 de mayo de 2012].
- Lörscher, W. (1986). "Linguistic Aspects of Translation Processes: Towards an Analysis of Translation Performance". En J. House y S. Blum-Kulka (eds.), *Interlingual and Intercultural Communication. Discourse and Cognition in Translation and Second Language Acquisition Studies*. Tübingen: Gunter Narr, 275-292.
- Lörscher, W. (1991). *Translation Performance, Translation Process, and Translation Strategies. A Psycholinguistic Investigation*. Tübingen: Gunter Narr.
- Lörscher, W. (1992). "Investigating the Translation Process". *Meta*, 37(3), 426-439. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Lowe, J. B., Baker, C. F. y Fillmore, C. J. (1997). "A Frame-Semantic Approach to Semantic Annotation". *Proceedings of the SIGLEX Workshop on Tagging Text with Lexical Semantics: Why, What, and How?*, Washington, 18-24. Disponible en <http://www.aclweb.org/anthology-new/W/W97/W97-0204.pdf> [Consulta: 8 de mayo de 2012].
- Lyons, J. (1977). *Semantics* (Vol. 1). Cambridge: Cambridge University Press.
- Mackintosh, K. (1998). "An empirical study of dictionary use in L2-L1 translation". En B. T. S. Atkins (ed.), *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Niemeyer, 123-149.
- Mahesh, K. (1996). *Ontology Development for Machine Translation: Ideology and Methodology* (MCCS-96-292). Computing Research Laboratory, New Mexico State University.
- Mahesh, K. y Nirenburg, S. (1995a). "Semantic Classification for Practical Natural Language Processing". En R. P. Schwartz, C. Beghtol, E. K. Jacob, B. H. Kwasnik y P. J. Smith (eds.), *Advances in Classification Research: Proceedings of the 6th ASIS SIG/CR Classification Research Workshop* (Vol. 6), Chicago, 79-94. Disponible en <http://journals.lib.washington.edu/index.php/acro/article/view/12663/11170> [Consulta: 8 de mayo de 2012].
- Mahesh, K. y Nirenburg, S. (1995b). "A Situated Ontology for Practical NLP". En D. Suce (ed.), *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing. Conference on Artificial Intelligence (IJCAI 1995)*, Montréal, 19-21. Disponible en <http://ilit.umbc.edu/SergeiPub/SituatOntology95.pdf> [Consulta: 8 de mayo de 2012].
- Maia, B. (1997). "Do-it-yourself corpora... with a little bit of help from your friends!" En B. Lewandowska-Tomaszczyk y P. J. Melia (eds.), *Practical applications in*



- Language Corpora (PALC 1997)*, University of Lodz, 403-410. Disponible en <http://ler.letras.up.pt/uploads/ficheiros/5828.pdf> [Consulta: 8 de mayo de 2012].
- Maia, B. (2000). "Making Corpora: A Learning Process". En S. Bernardini y F. Zanettin (eds.), *I corpora nella didattica della traduzione - Corpus use and learning to translate*. Bologna: CLUEB, 47-60.
- Maia, B. (2003). "Training Translators in Terminology and Information Retrieval using Comparable and Parallel Corpora". En F. Zanettin, S. Bernardini y D. Stewart (eds.), *Corpora in Translator Education*. Manchester: St. Jerome, 43-54.
- Malmkjær, K. (1998). "Love thy Neighbour: Will Parallel Corpora Endear Linguists to Translators?" *Meta*, 43(4), 474-660. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Malmkjær, K. (2003). "On a Pseudo-subversive Use of Corpora in Translator Training". En F. Zanettin, S. Bernardini y D. Stewart (eds.), *Corpora in Translator Education*. Manchester: St. Jerome, 119-134.
- Maniez, F. (2007). "Using the Web and computer corpora as language resources for the translation of complex noun phrases in medical research articles". *Panacea*, 8(26), 162-167. Disponible en [http://medtrad.org/panacea/IndiceGeneral/n26\\_tribuna-Maniez.pdf](http://medtrad.org/panacea/IndiceGeneral/n26_tribuna-Maniez.pdf) [Consulta: 8 de mayo de 2012].
- Marshman, E. y L'Homme, M.-C. (2008). "Portabilité des marqueurs de la relation causale: étude sur deux corpus spécialisés". En F. Maniez, P. Dury, N. Arlin y C. Rougemont (eds.), *Corpus et dictionnaires de langues de spécialité*. Bresson: Presses Universitaires de Grenoble.
- Marshman, E., Morgan, T. y Meyer, I. (2002). "French patterns for expressing concept relations". *Termonology*, 8(1), 1-29.
- Martín Mingorance, L. (1998). "La lexicografía onomasiológica". En A. Marín Rubiales (ed.), *El modelo lexemático-funcional. El legado lingüístico de Leocadio Martín Mingorance*. Granada: Universidad de Granada, 261-278.
- Martin, W. y Al, B. P. F. (1988). "User-Oriented in Dictionaries: 9 Propositions". *3rd International EURALEX Congress (BudaLEX 88 Proceedings)*, Budapest, 4-9 de septiembre de 1988, 393-399. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex1988/](http://www.euralex.org/elx_proceedings/Euralex1988/) [Consulta: 8 de mayo de 2012].
- Martínez Miguélez, M. (2001). "Uso del Programa Computacional Atlas.ti de Thomas Mühr (Univ. de Berlín) en la estructuración teórica de "datos" cualitativos". *ARGOS*, 34, 139-156. Disponible en <http://prof.usb.ve/miguelm/estructuracionteorica%203.html> [Consulta: 8 de mayo de 2012].
- Massey, G. y Ehrensberger-Dow, M. (2010a). "Investigating demands on language professionals: methodological challenges in exploring translation competence". *Bulletin suisse de linguistique appliquée*, 1, 127-141. Disponible en [http://www.uni-saarland.de/fileadmin/user\\_upload/Professoren/fr46\\_ProfGerzymisch-Arbogast/AS/Massey\\_Dow\\_Bulletin\\_VALS\\_ASILA.pdf](http://www.uni-saarland.de/fileadmin/user_upload/Professoren/fr46_ProfGerzymisch-Arbogast/AS/Massey_Dow_Bulletin_VALS_ASILA.pdf) [Consulta: 8 de mayo de 2012].
- Massey, G. y Ehrensberger-Dow, M. (2010b). "Technical and Instrumental Competence in Translation: Investigating Workplace Processes and Practices". *International Conference: Translation and Ergonomics*, Université Stendhal Grenoble 3.
- Massey, G. y Ehrensberger-Dow, M. (2011). "Investigating Information Literacy: A Growing Priority in Translation Studies". *Across Languages and Cultures*, 12(2),

- 193-211. Disponible en <http://www.akademai.com/content/h524268024272801/> [Consulta: 8 de mayo de 2012].
- Mayoral, R. (1994). "La documentación en la traducción". En L. Jacoby y E. Benítez (eds.), *Traducción, interpretación, lenguaje*. Madrid: Fundación Actilibre, 107-118.
- McCreary, D. R. y Dolezal, R. (1999). "A Study of Dictionary Use by ESL Students in an American University". *International Journal of Lexicography*, 12(2), 107-146.
- McEnery, T. y Wilson, A. (1993). *Corpora and Translation: Uses and Future Prospects*. Lancaster: UCREL.
- McEnery, T. y Wilson, A. (1996). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Meijssen, G. (2009). "The Philosophy behind OmegaWiki and the Visions for the Future". En H. Bergenholtz, S. Nielsen y S. Tarp (eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 91-98.
- Mel'čuk, I. A., Clas, A. y Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Brussels: Duculot.
- Meyer, I. (1988). "The General Bilingual Dictionary as a Working Tool in *Thème*". *Meta*, 33(3), 368-376. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Meyer, I. (2001). "Extracting knowledge-rich contexts for terminography. A conceptual and methodological framework". En D. Bourigault, C. Jacquemin y M.-C. L'Homme (eds.), *Recent Advances in Computational Terminology*. Amsterdam/Philadelphia: John Benjamins, 279-302.
- Meyer, I. y Mackintosh, K. (1996). "The Corpus from a Terminographer's Viewpoint". *International Journal of Corpus Linguistics*, 1(2), 257-268.
- Miles, M. B. y Huberman, A. M. (1994). *Qualitative Data Analysis*. California: Sage Publications.
- Miller, G. A. (1998a). "Foreword by George A. Miller". En C. Fellbaum (ed.), *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press, xv-xxii.
- Miller, G. A. (1998b). "Nouns in WordNet". En C. Fellbaum (ed.), *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press, 23-46.
- Miller, G. A. y Gildea, P. M. (1985). "How to Misread a Dictionary". *AILA Bulletin*, 13-26.
- Möhrs, C. y Töpel, A. (2011). "The "Online Bibliography of Electronic Lexicography" (OBELEX)". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 199-202. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Mondahl, M. y Jensen, K. A. (1996). "Lexical Search Strategies in Translation". *Meta*, 41(1), 97-113. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Montero Martínez, S. y Faber, P. (2008). *Terminología para traductores e intérpretes*. Granada: Tragacanto.
- Montero Martínez, S. y Faber, P. (2009). "Terminological Competence in Translation". *Terminology*, 15(1), 88-104.
- Moraes, N. S. d. (2006). "How to Use Linguistic Corpora to Improve your Translations". *Proceedings of the 2006 Annual ATA Conference*, New Orleans, 1-4 de noviembre de 2006. Disponible en [http://www.justrightcommunications.com/files/Moraes\\_corpus\\_improve\\_translation\\_2006.pdf](http://www.justrightcommunications.com/files/Moraes_corpus_improve_translation_2006.pdf) [Consulta: 8 de mayo de 2012].

- Muhr, T. (1991). "ATLAS/ti — A Prototype for the Support of Text Interpretation". *Qualitative Sociology*, 14(4), 349-371.
- Müller-Spitzer, C., Koplenig, A. y Töpel, A. (2011). "What Makes a Good Online Dictionary? - Empirical Insights from an Interdisciplinary Research Project". En I. Kosem y K. Kosem (eds.), *Electronic Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 203-208. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Muñoz Sánchez, P. (2007). "La Web 2.0 en el aula de traducción". *Panacea*, 8(26), 168-172. Disponible en [http://medtrad.org/panacea/IndiceGeneral/n26\\_tribuna-Sanchez.pdf](http://medtrad.org/panacea/IndiceGeneral/n26_tribuna-Sanchez.pdf) [Consulta: 8 de mayo de 2012].
- Muráth, J. (2000). *Zweisprachige Fachlexikographie - dargestellt am Beispiel der ungarisch-deutschen, deutsch-ungarischen Wirtschaftskommunikation*. (Tesis doctoral), ELTE BTK Germanistisches Institut, Budapest.
- Nazar, R., Vivaldi, J. y Cabré, M. T. (2008). "A Suite to Compile and Analyze an LSP Corpus". En N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odjik, S. Piperidis y D. Tapias (eds.), *6th International Conference on Language Resources and Evaluation (LREC 2008)*, Marrakech, 1164-1169. Disponible en <http://www.lrec-conf.org/proceedings/lrec2008/> [Consulta: 8 de mayo de 2012].
- Nesi, H. (1998). "Dictionaries on Computer: How Different Markets Have Created Different Products". En A. Wilson y J. Schmied (eds.), *Proceedings of the Chemnitz Symposium on Language Learning and Computers*, 20-21 de febrero de 1998. Disponible en <http://www.tu-chemnitz.de/phil/english/chairs/linguist/real/independent/llc/Conference1998/Paper/s/Nesi.htm> [Consulta: 8 de mayo de 2012].
- Nesi, H. (1999). "A User's Guide to Electronic Dictionaries for Language Learners". *International Journal of Lexicography*, 12(1), 55-66.
- Nesi, H. (2000a). "Electronic Dictionaries in Second Language Vocabulary Comprehension and Acquisition: the State of the Art". En U. Heid, S. Evert, E. Lehmann y C. Rohrer (eds.), *9th EURALEX International Congress*, Stuttgart, 8-12 de agosto de 2000, 839-847. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2000/](http://www.euralex.org/elx_proceedings/Euralex2000/) [Consulta: 8 de mayo de 2012].
- Nesi, H. (2000b). *The Use and Abuse of EFL Dictionaries. How Learners of English as a Foreign Language Read and Interpret Dictionary Entries*. Tübingen: Niemeyer.
- Nesi, H. y Hail, R. (2002). "A Study of Dictionary Use by International Students at a British University". *International Journal of Lexicography*, 15(4), 277-305.
- Neubert, A. (2000). "Competence in Language, in Languages, and in Translation". En C. Schäffner y B. Adab (eds.), *Developing Translation Competence*. Amsterdam/Philadelphia: John Benjamins, 3-18.
- Newmark, P. (1993). *Paragraphs on Translation*. Clevedon: Multilingual Matters.
- Newmark, P. (1998). *A Textbook of Translation*. London: Prentice Hall.
- Nielsen, S. (2009). "Reviewing Printed and Electronic Dictionaries. A Theoretical and Practical Framework". En S. Nielsen y S. Tarp (eds.), *Lexicography in the 21st Century. In honour of Henning Bergenholtz*. Amsterdam/Philadelphia: John Benjamins, 23-42.
- Nielsen, S. (2010). "Specialised Translation Dictionaries for Learners". En P. A. Fuertes-Olivera (ed.), *Specialized Dictionaries for Learners*. Berlin: De Gruyter, 69-82.

- Nielsen, S. y Almind, R. (2011). "From Data to Dictionary". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 141-167.
- Nielsen, S. y Mourier, L. (2005). "Internet Accounting Dictionaries: Present Solutions and Future Opportunities". *Hermes, Journal of Linguistics*, 34, 83-116. Disponible en <https://pure.au.dk/portal/files/9884/Hermes-34-5-Nielsen-Mourier.pdf> [Consulta: 8 de mayo de 2012].
- Nielsen, S. y Mourier, L. (2007). "Design of a function-based internet accounting dictionary". En H. Gottlieb y J. E. Mogensen (eds.), *Dictionary Visions, Research and Practice*. Amsterdam/Philadelphia: John Benjamins, 119-138.
- Nord, B. (1999). "Das zweisprachige Wörterbuch als Hilfsmittel bei der Übersetzung". En A. Gil, J. Haller, E. Steiner y H. Gerzymisch-Arbogast (eds.), *Modelle der Translation*. Frankfurt am Main: Peter Lang, 375-390.
- Nord, C. (1991). *Text Analysis in Translation. Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*. Amsterdam/Atlanta: Rodopi.
- Olaz, Á. (2008). *La entrevista en profundidad*. Oviedo: Septem Ediciones.
- Oppenocht, L. y Schutz, R. (2003). "Developments in electronic dictionary design". En P. Van Sterkenburg (ed.), *A Practical Guide to Lexicography*. Amsterdam/Philadelphia: John Benjamins, 215-227.
- Orozco, M. (2002). "Revisión de investigaciones empíricas en traducción escrita". *TRANS. Revista de Traductología*, 6, 63-85. Disponible en [http://www.trans.uma.es/Trans\\_6/t6\\_63-85\\_MOrozco.pdf](http://www.trans.uma.es/Trans_6/t6_63-85_MOrozco.pdf) [Consulta: 8 de mayo de 2012].
- Orozco, M. y Hurtado Albir, A. (2002). "Measuring Translation Competence Acquisition". *Meta*, 47(3), 375-402. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- PACTE. (2001). "La competencia traductora y su adquisición". *Quaderns. Revista de Traducció*, 6, 39-45. Disponible en [http://grupsderecerca.uab.cat/pacte/sites/grupsderecerca.uab.cat.pacte/files/2001b\\_PACTE\\_QUADERNS.pdf](http://grupsderecerca.uab.cat/pacte/sites/grupsderecerca.uab.cat.pacte/files/2001b_PACTE_QUADERNS.pdf) [Consulta: 8 de mayo de 2012].
- PACTE. (2002). "Una investigación empírico-experimental sobre la adquisición de la competencia traductora". En A. Alcina y S. Gamero (eds.), *La traducción científico-técnica y la terminología en la sociedad de la información*. Castellón de la Plana: Publicaciones de la Universitat Jaume I, 125-138.
- PACTE. (2003). "Building a Translation Competence Model". En F. Alves (ed.), *Triangulating Translation*. Amsterdam/Philadelphia: John Benjamins, 43-66.
- PACTE. (2005a). "Investigating Translation Competence: Conceptual and Methodological Issues". *Meta*, 50(2), 609-619. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- PACTE. (2005b). "Primeros resultados de un experimento sobre la Competencia Traductora". *Actas del II Congreso Internacional de la AIETI (Asociación Ibérica de Estudios de Traducción e Interpretación)*, Madrid, 573-587. Disponible en [http://grupsderecerca.uab.cat/pacte/sites/grupsderecerca.uab.cat.pacte/files/2005a\\_PACTE\\_Actas%20AIETI%20CD-ROM\\_0.pdf](http://grupsderecerca.uab.cat/pacte/sites/grupsderecerca.uab.cat.pacte/files/2005a_PACTE_Actas%20AIETI%20CD-ROM_0.pdf) [Consulta: 8 de mayo de 2012].
- PACTE. (2009). "Results of the Validation of the PACTE Translation Competence Model: Acceptability and Decision Making". *Across Languages and Cultures*, 10(2), 207-230. Disponible en <http://www.akademaii.com/content/rp183p44376w4188/fulltext.pdf> [Consulta: 8 de mayo de 2012].



- PACTE. (2011). "Results of the Validation of the PACTE Translation Competence Model: Translation Project and Dynamic Translation Index". En S. O'Brien (ed.), *IATIS Yearbook 2010*. London: Continuum.
- Palomares Perraut, R. (1999). "Evaluación de recursos de comunicación para el traductor en Internet". En M. Pinto y J. A. Córdón (eds.), *Técnicas documentales aplicadas a la traducción*. Madrid: Síntesis, 179-195.
- Palomares Perraut, R. y Pinto, M. (2000). "Aproximación a las necesidades, hábitos y usos documentales del traductor". *Terminologie et Traduction*, 3, 98-124.
- Pastor, V. (2008a). "Estrategias de búsqueda onomasiológica en el proceso de traducción. La opinión de los traductores". *Fòrum de Recerca*, 14. Disponible en <http://www.uji.es/bin/publ/edicions/jfi14/trad/5.pdf> [Consulta: 8 de mayo de 2012].
- Pastor, V. (2008b). *Estrategias de búsqueda onomasiológica en el proceso de traducción: la opinión de los traductores*. (Trabajo de investigación, Máster en Tecnologías de la Traducción y Localización), Universitat Jaume I, Castellón de la Plana.
- Pastor, V. (2010). "Techniques de recherche onomasiologique dans le processus de traduction". *Seminario Centre de Recherche en Terminologie et Traduction*, Université Lumière Lyon 2. Disponible en <http://recherche.univ-lyon2.fr/crtt/spip.php?article5>
- Pastor, V. y Alcina, A. (2009). "Search techniques in corpora for the training of translators". En I. Ilisei, V. Pekar y S. Bernardini (eds.), *International Workshop on Natural Language Processing Methods and Corpora in Translation, Lexicography and Language Learning*, Bulgaria, 13-20.
- Pastor, V. y Alcina, A. (2010). "Search Techniques in Electronic Dictionaries: A Classification for Translators". *International Journal of Lexicography*, 23(3), 307-354.
- Pastor, V. y Alcina, A. (2011). "Acceso a la información terminológica en Internet: técnicas para traductores". En S. Maruenda-Bataller y B. Clavel-Arroitia (eds.), *Multiple Voices in Academic and Professional Discourse. Current Issues in Specialised Language Research, Teaching and New Technologies*. Newcastle: Cambridge Scholar Publishing, 243-256.
- Pastor, V. y Alcina, A. (en prensa). "Researching the use of electronic dictionaries". En H. Jackson (ed.), *Continuum Companion to Lexicography*. London: Continuum.
- Pavlović, N. (2009). "More ways to explore the translating mind: collaborative translation protocols". En S. Göpferich, A. L. Jakobsen y I. M. Mees (eds.), *Behind the mind. Methods, models and results in translation process research*. Copenhagen: Samfundslitteratur, 81-105.
- Pearson, J. (1996a). "Electronic Texts and Concordances in the Translation Classroom". *TEANGA: The Irish Yearbook of Applied Linguistics*, 16, 85-95. Disponible en <http://www.eric.ed.gov/PDFS/ED414752.pdf> [Consulta: 8 de mayo de 2012].
- Pearson, J. (1996b). "The Expression of Definitions in Specialised Texts: A Corpus-based Analysis". En M. Gellerstam, J. Järborg, S.-G. Malmgren, K. Norén, L. Rogström y C. Røjder Pappmehl (eds.), *7th EURALEX International Congress (Proceedings Part II)*, Göteborg University, 13-18 de agosto de 1996, 817-824. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex1996\\_2/](http://www.euralex.org/elx_proceedings/Euralex1996_2/) [Consulta: 8 de mayo de 2012].
- Pearson, J. (1996c). "Teaching Terminology Using Electronic Resources". En S. Botley, J. Glass, T. McEnery y A. Wilson (eds.), *Proceedings of Teaching and Language Corpora (TALC 1996)* (Vol. Technical Papers 9), University Centre for Computer Corpus Research on Language (UCREL), 92-105.
- Pearson, J. (1998). *Terms in Context*. Amsterdam/Philadelphia: John Benjamins.

- Pearson, J. (2003). "Using Parallel Texts in the Translator Training Environment". En F. Zanettin, S. Bernardini y D. Stewart (eds.), *Corpora in Translator Education*. Manchester: St. Jerome, 15-24.
- Pérez Hernández, M. C. (2002). "Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento". *Estudios de Lingüística del Español (ELiEs)*, 18. Disponible en <http://elies.rediris.es/elies18/index.html> [Consulta: 8 de mayo de 2012].
- Peters, C. y Picchi, E. (1998). "Bilingual Reference Corpora for Translators and Translation Studies". En L. Bowker, M. Cronin, D. Kenny y J. Pearson (eds.), *Unity in Diversity? Current Trends in Translation Studies*. Manchester: St. Jerome, 91-100.
- Peters, C., Picchi, E. y Biagini, L. (1996). "Parallel and Comparable Bilingual Corpora in Language Teaching and Learning". En S. Botley, J. Glass, T. McEnery y A. Wilson (eds.), *Proceedings of Teaching and Language Corpora (TALC 1996)* (Vol. Technical Papers 9), University Centre for Computer Corpus Research on Language (UCREL), 73-85.
- Philip, G. (2009). "Arriving at equivalence. Making a case for comparable general reference corpora in translation studies". En A. Beeby, P. Rodríguez Inés y P. Sánchez-Gijón (eds.), *Corpus Use and Translating*. Amsterdam/Philadelphia: John Benjamins, 59-73.
- Pinto, M. (2000). "Documentación para la traducción en la sociedad de la información". *XV Coloquio de la Association Internationale de Bibliologie*, Salamanca. Disponible en [http://www.mariapinto.es/web/doc/Documentacion\\_para\\_la\\_traducccion.pdf](http://www.mariapinto.es/web/doc/Documentacion_para_la_traducccion.pdf) [Consulta: 8 de mayo de 2012].
- Pinto, M. (2005). "Alfabetización en información para traductores: propuesta del modelo ALFINTRA". En D. Sales (ed.), *La Biblioteca de Babel: Documentarse para traducir*. Granada: Comares, 12-32.
- Pinto, M. y Sales, D. (2007). "A research case study for user-centred information literacy instruction: information behaviour of translation trainees". *Journal of Information Science*, 33(5), 531-550. Disponible en <http://jis.sagepub.com/content/33/5/531> [Consulta: 8 de mayo de 2012].
- Planas, E. (2005). "SIMILIS: Second-generation translation memory software". *Translating and the Computer*, 27. Disponible en <http://www.mt-archive.info/Aslib-2005-Planas.pdf> [Consulta: 8 de mayo de 2012].
- Poirier, C. (1989). "Les différents supports du dictionnaire: livre, microfiche, dictionnaire électronique". En F. J. Hausmann, O. Reichmann, E. Wiegand y L. Zgusta (eds.), *Wörterbücher/Dictionaries/Dictionnaires, Ein internationales Handbuch zur Lexikographie/An International Encyclopedia of Lexicography/Encyclopédie internationale de lexicographie* (Vol. 1). Berlin: de Gruyter, 322-327.
- Polguère, A. (2000). "Towards a theoretically-motivated general public dictionary of semantic derivations and collocations for French". En U. Heid, S. Evert, E. Lehmann y C. Rohrer (eds.), *9th EURALEX International Congress*, Stuttgart, 8-12 de agosto de 2000, 517-527. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2000/](http://www.euralex.org/elx_proceedings/Euralex2000/) [Consulta: 8 de mayo de 2012].
- Prieto Velasco, J. A. y López Rodríguez, C. I. (2009). "Managing graphic information in terminological knowledge bases". *Terminology*, 15(2), 179-213.
- Prinsloo, D. J., Heid, U., Bothma, T. J. y Faaß, G. (2011). "Interactive, dynamic electronic dictionaries for text production". En I. Kosem y K. Kosem (eds.), *Electronic*

- Lexicography in the 21st Century New Applications for New Users (eLex 2011)*, Bled (Eslovenia), 10-12 de noviembre de 2011, 215-220. Disponible en [http://www.trojina.si/elex2011/elex2011\\_proceedings.pdf](http://www.trojina.si/elex2011/elex2011_proceedings.pdf) [Consulta: 8 de mayo de 2012].
- Quirk, R. (1973). "The Social Impact of Dictionaries in the U. K." En R. I. McDavid y A. R. Ducker (eds.), *Lexicography in English*. New York: New York Academy of Sciences, 77-88.
- Quiroz, G., Tebé, C. y Cabré, M. T. (2002). "Metodología de elaboración de terminologías bilingües en programas de memoria de traducción". *Actas del VIII Simposio de RITerm*. Disponible en [http://www.riterm.net/actes/8simposio/Quiroz\\_Tebe\\_Cabre.htm](http://www.riterm.net/actes/8simposio/Quiroz_Tebe_Cabre.htm)
- Recoder, M. J. y Cid, P. (2004). "La documentación en la traducción especializada". En C. Gonzalo García y V. García Yebra (eds.), *Manual de documentación y terminología para la traducción especializada*. Madrid: Arcos/Libros, 73-88.
- Renouf, A., Kehoe, A. y Banerjee, J. (2007). "WebCorp: an integrated system for web text search". En M. Hundt, N. Nesselhauf y C. Biewer (eds.), *Corpus Linguistics and the Web*. Amsterdam: Rodopi, 47-67.
- Resnik, P. (1998). "Parallel Strands: A Preliminary Investigation into Mining the Web for Bilingual Text". En D. Farwell, L. Gerber y E. H. Hovy (eds.), *Machine Translation and the Information Soup: Third Conference of the Association for Machine Translation in the Americas (AMTA 1998)*, Langhorne, 72-82.
- Resnik, P. (1999). "Mining the Web for Bilingual Text". En R. Dale y K. Church (eds.), *37th Annual Meeting of the Association for Computational Linguistics (ACL 1999)*, Maryland, 527-534.
- Resnik, P. y Smith, N. A. (2003). "The Web as a Parallel Corpus". *Computational Linguistics*, 29(3), 349-380.
- Riediger, H. (2009a, octubre de 2009). *Internet per mediatori linguistici*. Disponible en <http://www.term-minator.it/corso/index.html> [Consulta: 8 de mayo de 2012].
- Riediger, H. (2009b). "Term-minator – Il WWW come vocabolario vivo". *mediAzioni*, 7(Monografía/Special Issue: Atti del convegno "Terminologia e mediazione linguistica: approcci e metodi a confronto"). Disponible en <http://www.mediazioni.sitlec.unibo.it/index.php/no7-anno2009.html> [Consulta: 8 de mayo de 2012].
- Riediger, H., Massey, G., Lenz, S. y Wibaut, H.-D. (2006). *RISK: Recherchieren Informationsmanagement Sprache und Kommunikation*. Disponible en [http://www.term-minator.eu/risk/en/en\\_index.html#](http://www.term-minator.eu/risk/en/en_index.html#) [Consulta: 8 de mayo de 2012].
- Rizo Rodríguez, A. J. y Valera Hernández, S. (2000). "Lexicografía bilingüe: el español y la lengua inglesa". En I. Ahumada (ed.), *Cinco siglos de lexicografía del español. IV Seminario de Lexicografía Hispánica*, Jaén, 17-19 de noviembre de 1999, 341-380.
- Robb, T. (2003). "Google as a Corpus Tool?" *ETJ Journal*, 4(1). Disponible en <http://itprofessionals.com/en/resources/etjjournal.html> [Consulta: 8 de mayo de 2012].
- Roberts, R. P. (1990). "Translation and the Bilingual Dictionary". *Meta*, 35(1), 74-81. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Roberts, R. P. (1996). "Dictionaries, dictionary-making and translation". En P. Fernández Nistal y J. M. Bravo Gozalo (eds.), *A Spectrum of Translation Studies*. Valladolid: Universidad de Valladolid, 9-35.

- Roberts, R. P. (1997). "Using dictionaries efficiently". *38th Annual Conference of the American Translators Association*, San Francisco (California). Disponible en <http://www.dico.uottawa.ca/articles-en.htm> [Consulta: 8 de mayo de 2012].
- Roberts, R. P. y Bossé-Andrieu, J. (2006). "Corpora and Translation". En L. Bowker (ed.), *Lexicography, Terminology and Translation*. Ottawa: University of Ottawa Press, 201-214.
- Roberts, R. P. y Langlois, L. (2001). "L'apport de l'informatique à la recherche lexicographique". *Meta*, 46(4), 711-720. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Robertson, T. (1859) *Dictionnaire idéologique*. Paris: Derache.
- Robinson, D. (2003). *An Introduction to the Theory and Practice of Translation*. London: Routledge.
- Rodríguez Inés, P. (2009). "Evaluating the process and not just the product when using corpora in translator education". En A. Beeby, P. Rodríguez Inés y P. Sánchez-Gijón (eds.), *Corpus Use and Translating*. Amsterdam/Philadelphia: John Benjamins, 129-148.
- Rogers, M. y Ahmad, K. (1998). "The Translator and the Dictionary: Beyond Words?" En B. T. S. Atkins (ed.), *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Niemeyer, 193-204.
- Roget, P. M. (1852). In B. Kirkpatrick (Ed.), *Roget's Thesaurus of English Words and Phrases*. London: Penguin Books. (Reprinted from: 2000).
- Ronowicz, E., Hehir, J., Kaimi, T., Kojima, K. y Lee, D.-S. (2005). "Translator's Frequent Lexis Store and Dictionary Use as Factors in SLT Comprehension and Translation Speed - A Comparative Study for Professional, Paraprofessional and Novice Translators". *Meta*, 50(2), 580-596. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Rundell, M. (2000). "The biggest corpus of all". *Humanising Language Teaching*, Year 2(3). Disponible en <http://www.hltmag.co.uk/may00/idea.htm> [Consulta: 8 de mayo de 2012].
- Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Christopher, R. J. y Scheffczyk, J. (2010). "FrameNet II: Extended Theory and Practice". Disponible en <http://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf> [Consulta: 8 de mayo de 2012].
- Rychlý, P. (2007). "Manatee/Bonito – A Modular Corpus Manager". En P. Sojka y A. Horák (eds.), *Proceedings of Recent Advances in Slavonic Natural Language Processing (RASLAN 2007)*, Masaryk University (Brno), 65-70.
- Sager, J. C. (1992). "The Translator as a Terminologist". En C. Dollerup y A. Lindegaard (eds.), *Teaching Translation and Interpreting*. Amsterdam/Philadelphia: John Benjamins, 107-123.
- Sager, J. C. (1993). *Curso práctico sobre el procesamiento de la terminología*. Madrid: Pirámide Salamanca, Fundación Germán Sánchez Ruipérez.
- Sánchez-Gijón, P. (2003a). *Els documents digitals especialitzats: utilització de la lingüística de corpus com a front de recursos per a la traducció*. (Tesis doctoral, Programa de Doctorat de Teoria de la Traducció), Universidad Autónoma de Barcelona, Barcelona. Disponible en <http://www.tesisenred.net/TDX-0123104-173209> [Consulta: 8 de mayo de 2012].
- Sánchez-Gijón, P. (2003b). "És la web pública la nova biblioteca del traductor?" *Tradumàtica*, 2, 1-7.



- Sánchez-Gijón, P. (2004a). *L'ús de corpus en la traducció especialitzada: compilació de corpus ad hoc i extracció de recursos terminològics*. Barcelona: Institut Universitari de Lingüística Aplicada. Universitat Pompeu Fabra.
- Sánchez-Gijón, P. (2004b). "La extracción de conocimiento y terminología a partir de corpus *ad hoc*: el uso de documentos digitales de la web pública". *Lingüística Antverpiensa*, 3, 179-202.
- Sánchez-Gijón, P. (2009). "Developing documentation skills to build do-it-yourself corpora in the specialised translation course". En A. Beeby, P. Rodríguez Inés y P. Sánchez-Gijón (eds.), *Corpus Use and Translating*. Amsterdam/Philadelphia: John Benjamins, 109-127.
- Sánchez-Gijón, P., Aguilar-Amat, A., Mesa-Lao, B. y Solé, M. P. (2009). "Applying terminology knowledge to translation: Problem-based learning for a degree in translation and interpreting". *Terminology*, 15(1), 105-118.
- Sánchez Ramos, M. d. M. (2004a). *El uso de los diccionarios electrónicos y otros recursos de Internet como herramienta para la formación del traductor inglés-español*. (Tesis doctoral), Universitat Jaume I, Castellón de la Plana.
- Sánchez Ramos, M. d. M. (2004b). "Estudio experimental sobre el uso del diccionario como herramienta para el traductor: hacia una descripción de necesidades". En M. A. García Peinado y E. Ortega Arjonilla (eds.), *Panorama actual de la investigación en Traducción e Interpretación*. Granada: Atrio, 477-486.
- Santana, O., Hernández, Z., Pérez, J., Rodríguez, G. y Carreras, F. (1996). "Diccionarios en soportes informáticos. Los diccionarios en el horizonte del siglo XXI". *Cuadernos Cervantes*, 11, 68-77. Disponible en [http://www.gedlc.ulpgc.es/art\\_gied.htm](http://www.gedlc.ulpgc.es/art_gied.htm) [Consulta: 8 de mayo de 2012].
- Santana, O., Rodríguez, G. y Hernández, Z. (1997). "Herramienta para el manejo de diccionarios ideológicos". *Lingüística Española Actual*, XIX(1), 127-135. Disponible en [http://www.gedlc.ulpgc.es/art\\_gied.htm](http://www.gedlc.ulpgc.es/art_gied.htm) [Consulta: 8 de mayo de 2012].
- Sarmento, L., Maia, B. y Santos, D. (2004). "The Corpógrafo - a Web-based environment for corpora research". En M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa y R. Silva (eds.), *4th International Conference on Language Resources and Evaluation (LREC 2004)*, Portugal, 449-452. Disponible en <http://www.lrec-conf.org/proceedings/lrec2004/> [Consulta: 8 de mayo de 2012].
- Schäffner, C. (1998). "Parallel Texts in Translation". En L. Bowker, M. Cronin, D. Kenny y J. Pearson (eds.), *Unity in Diversity? Current Trends in Translation Studies*. Manchester: St. Jerome, 83-90.
- Séguinot, C. (1996). "Some Thoughts About Think-Aloud Protocols". *Target*, 8(1), 75-95.
- Seleskovitch, D. y Lederer, M. (1989). *Pédagogie raisonnée de l'interprétation*. Paris: Didier Erudition.
- Serrahima, L. (2009a). "El programa de reconocimiento de voz Dragon Naturally Speaking". *Panacea*, X(29), 72-75.
- Serrahima, L. (2009b). "Reconocimiento de voz de Windows Vista: ¿mejor, igual o peor que Dragon Naturally Speaking?" *Panacea*, X(29), 76-79.
- Sevilla Muñoz, M. y Sevilla Muñoz, J. (2005a). "Búsqueda terminológica inversa". *El Trujamán*. Disponible en [http://www.cvc.cervantes.es/trujaman/anteriores/marzo\\_05/30032005.htm](http://www.cvc.cervantes.es/trujaman/anteriores/marzo_05/30032005.htm) [Consulta: 8 de mayo de 2012].
- Sevilla Muñoz, M. y Sevilla Muñoz, J. (2005b). "Todo lo que busque el traductor está en Internet". *El Trujamán*. Disponible en

- [http://www.cvc.cervantes.es/trujaman/anteriores/septiembre\\_05/28092005.htm](http://www.cvc.cervantes.es/trujaman/anteriores/septiembre_05/28092005.htm)  
[Consulta: 8 de mayo de 2012].
- Sharoff, S. (2006). "Creating General-Purpose Corpora Using Automated Search Engine Queries". En M. Baroni y S. Bernardini (eds.), *WaCky! Working Papers on the Web as Corpus*. Bologna: Gedit, 63-98.
- Sharpe, P. (1995). "Electronic Dictionaries with Particular Reference to the Design of an Electronic Bilingual Dictionary for English-speaking Learners of Japanese". *International Journal of Lexicography*, 8(1), 39-54.
- Shreve, G. M. (1997). "Terminological Aspects of Text Production". En S. E. Wright y G. Budin (eds.), *Handbook of Terminology Management. Application-Oriented Terminology Management* (Vol. 2). Amsterdam/Philadelphia: John Benjamins, 772-787.
- Siepmann, D. (2006). "Collocation, Colligation and Encoding Dictionaries. Part II: Lexicographical Aspects". *International Journal of Lexicography*, 19(1), 1-39.
- Sierra, G. (1996). "Avances en el desarrollo del Diccionario Electrónico de Búsqueda Onomasiológica". *Actas del V Simposio de RITerm*. Disponible en <http://www.riterm.net/actes/5simposio/sierra2.htm> [Consulta: 8 de mayo de 2012].
- Sierra, G. (1999). *Design of a concept-oriented tool for terminology*. (Tesis doctoral, Doctor of Philosophy), University of Manchester, Manchester.
- Sierra, G. (2000). "The onomasiological dictionary: a gap in lexicography". En U. Heid, S. Evert, E. Lehmann y C. Rohrer (eds.), *9th EURALEX International Congress*, Stuttgart, 8-12 de agosto de 2000, 223-235. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex2000/](http://www.euralex.org/elx_proceedings/Euralex2000/) [Consulta: 8 de mayo de 2012].
- Sierra, G., Alarcón, R., Aguilar, C. y Bach, C. (2010). "Definitional verbal patterns for semantic relation extraction". En A. Auger y C. Barriere (eds.), *Probing Semantic Relations, Exploration and identification in specialized texts*. Amsterdam/Philadelphia: John Benjamins, 73-96.
- Sierra, G., Alarcón, R., Molina, A. y Aldana, E. (2009). "Web Exploitation for Definition Extraction". En E. Chávez, E. Furtado y A. L. Morán (eds.), *Latin American Web Congress*, México, 217-223.
- Sierra, G. y McNaught, J. (2000). "Design of an onomasiological search system: A concept-oriented tool for terminology". *Terminology*, 6(1), 1-34.
- Sinclair, J. (1996). "EAGLES Preliminary Recommendations on Corpus Typology, EAG-TCWG-CTYP/P". Disponible en <http://www.ilc.cnr.it/EAGLES/corpus/corpus.html> [Consulta: 8 de mayo de 2012].
- Sinclair, J. M. (2004). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Snell-Hornby, M. (1987). "Towards a learner's bilingual dictionary". En A. P. Cowie (ed.), *The Dictionary and the Language Learner. Papers of the EURALEX Seminar at the University of Leeds, 1-3 April 1985*. Tübingen: Niemeyer, 159-170.
- Snell-Hornby, M. (1988). *Translation Studies. An Integrated Approach*. Amsterdam/Philadelphia: John Benjamins.
- Sobkowiak, W. (1999). *Pronunciation in EFL Machine-Readable Dictionaries*. Poznań: Motivex.
- Soler, C. (1999). "Obras de consulta. Diccionario Ideológico de Julio Casares". *Métodos de Información*, 6(29-30). Disponible en <http://eprints.rclis.org/archive/00001545/01/1999-29-28.pdf> [Consulta: 8 de mayo de 2012].

- Soler, V. y Alcina, A. (2008). "Patrones léxicos para la extracción de conceptos vinculados por la relación parte-todo en español". *Terminology*, 14(1), 99-123.
- Somers, H. (2003). "The Translator's Workstation". En H. Somers (ed.), *Computers and Translation. A Translator's Guide*. Amsterdam/Philadelphia: John Benjamins, 13-30.
- Spink, A. y Jansen, B. J. (2005). *Web Search: Public Searching of the Web*. Dordrecht: Kluwer Academic Publishers.
- Spohr, D. (2011). "A Multi-Layer Architecture for 'Pluri-Monofunctional' Dictionaries". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 103-120.
- Stewart, D. (2000). "Conventionality, Creativity and Translated Text: The Implications of Electronic Corpora in Translation". En M. Olohan (ed.), *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*. Manchester: St. Jerome, 73-91.
- Tabares Plasencia, E., Ivanova, V. y Krüger, E. (eds.). (2008). *Análisis lingüístico contrastivo de textos especializados en español y alemán*. Berlin: Frank & Timme.
- Tarp, S. (2002). "Translation Dictionaries and Bilingual Dictionaries. Two Different Concepts". *Journal of Translation Studies*, 7, 59-84.
- Tarp, S. (2005). "The pedagogical dimension of the well-conceived specialised dictionary". *Ibérica*, 10, 7-21. Disponible en <http://www.aelfe.org/?s=revista&veure=10> [Consulta: 8 de mayo de 2012].
- Tarp, S. (2007). "Lexicography in the Information Age". *Lexikos*, 17, 170-179.
- Tarp, S. (2008a). *Lexicography in the Borderland between Knowledge and Non-Knowledge*. Tübingen: Niemeyer.
- Tarp, S. (2008b). "Revival of a Dusty Old Profession". *Hermes, Journal of Linguistics*, 41, 175-188. Disponible en <http://download2.hermes.asb.dk/archive/download/Hermes-41-9-tarp.pdf> [Consulta: 8 de mayo de 2012].
- Tarp, S. (2008c). "The Third Leg of Two-legged Lexicography". *Hermes, Journal of Linguistics*, 40, 117-131. Disponible en <http://download2.hermes.asb.dk/archive/download/Hermes-40-8-net-ST.pdf> [Consulta: 8 de mayo de 2012].
- Tarp, S. (2009a). "Beyond Lexicography: New Visions and Challenges in the Information Age". En H. Bergenholtz, S. Nielsen y S. Tarp (eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 17-32.
- Tarp, S. (2009b). "Reflections on Data Access in Lexicographic Works". En S. Nielsen y S. Tarp (eds.), *Lexicography in the 21st Century. In honour of Henning Bergenholtz*. Amsterdam/Philadelphia: John Benjamins, 43-62.
- Tarp, S. (2009c). "Reflections on Lexicographical Research". *Lexikos*, 19, 275-296.
- Tarp, S. (2010). "Functions of Specialised Learners' Dictionaries". En P. A. Fuertes-Olivera (ed.), *Specialized Dictionaries for Learners*. Berlin: De Gruyter, 39-53.
- Tarp, S. (2011a). "Do we need a (new) theory of lexicography?" *16th Annual International AFRILEX Conference UNAM*, Windhoek (Namibia). Disponible en <http://afrilex2011unam.webs.com/paperspresented.htm> [Consulta: 8 de mayo de 2012].
- Tarp, S. (2011b). "Lexicographical and Other e-Tools for Consultation Purposes: Towards the Individualization of Needs Satisfaction". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 54-70.

- Tercedor, M., Alarcón Navío, E., Prieto Velasco, J. A. y López Rodríguez, C. I. (2009). "Images as part of technical translation courses: implications and applications". *JoSTrans. The Journal of Specialised Translation*, 11, 143-168. Disponible en [http://www.jostrans.org/issue11/art\\_tercedoretal.pdf](http://www.jostrans.org/issue11/art_tercedoretal.pdf) [Consulta: 8 de mayo de 2012].
- Tercedor, M. y López Rodríguez, C. I. (2008). "Integrating corpus data in dynamic knowledge bases". *Terminology*, 14(2), 159-182.
- Teubert, W. (1996). "Comparable or Parallel Corpora?" *International Journal of Lexicography*, 9(3), 238-264.
- Teubert, W. y Čermáková, A. (2004). "Directions in Corpus Linguistics". En M. A. K. Halliday, W. Teubert, Y. Colin y A. Čermáková (eds.), *Lexicology and Corpus Linguistics: An Introduction*. London: Continuum, 113-166.
- Tirkkonen-Condit, S. y Jääskeläinen, R. (2000). *Tapping and Mapping the Process of Translation and Interpreting: Outlooks on Empirical Research*. Amsterdam/Philadelphia: John Benjamins.
- Tomaszczyk, J. (1979). "Dictionaries: Users and Uses". *Glottodidactica*, 12, 103-119.
- Tomaszczyk, J. (1989). "L1-L2 Technical Translation and Dictionaries". En M. Snell-Hornby y E. Pöhl (eds.), *Translation and Lexicography. Papers read at the Euralex Colloquium held at Innsbruck 2-5 July 1987*. Amsterdam/Philadelphia: John Benjamins, 177-186.
- Tono, Y. (1989). "Can a Dictionary Help One Read Better?" En G. James (ed.), *Lexicographers and their Works. (Exeter Linguistic Studies 14)*. Exeter: University of Exeter Press, 192-200.
- Tono, Y. (2009). "Pocket Electronic Dictionaries in Japan: User Perspectives". En H. Bergenholtz, S. Nielsen y S. Tarp (eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 33-67.
- Tono, Y. (2010). "A Critical Review of the Theory of Lexicographical Functions". *Lexicon*, 40, 1-16.
- Tutin, A., Grossmann, F., Falaise, A. y Kraif, O. (2009). "Autour du projet Scientext: étude des marques linguistiques du positionnement de l'auteur dans les écrits scientifiques". *6èmes Journées Internationales de Linguistique de Corpus*, Lorient. Disponible en <http://scientext.msh-alpes.fr/scientext-site/spip.php?article7> [Consulta: 8 de mayo de 2012].
- Valero, E. y Alcina, A. (2010). "Exploración de características conceptuales en contextos ricos en conocimiento mediante un programa de análisis cualitativo". *Revista de Lingüística y Lenguas Aplicadas*, 5, 241-254.
- Valles, M. S. (2001). "Ventajas y desafíos del uso de programas informáticos (e.g. ATLAS.ti y MAXqda) en el análisis cualitativo. Una reflexión metodológica desde la grounded theory y el contexto de la investigación social española." *Seminario sobre Investigación Avanzada Cualitativa Asistida por Ordenador*, Granada. Disponible en [http://centrodeestudiosandaluces.es/biblio/imagendoc/00000001\\_00000500/00000125/00000125\\_090h0101.PDF](http://centrodeestudiosandaluces.es/biblio/imagendoc/00000001_00000500/00000125/00000125_090h0101.PDF) [Consulta: 8 de mayo de 2012].
- Valles, M. S. (2002). *Entrevistas cualitativas* (Vol. 32). Madrid: Centro de Investigaciones Sociológicas.
- Valles, M. S. (2005). "El reto de la calidad en la investigación social cualitativa: de la retórica a los planteamientos de fondo y las propuestas técnicas". *Revista Española de Investigaciones Sociológicas*, 110, 91-114. Disponible en



- [http://www.reis.cis.es/REIS/PDF/REIS\\_110\\_051168261781313.pdf](http://www.reis.cis.es/REIS/PDF/REIS_110_051168261781313.pdf) [Consulta: 8 de mayo de 2012].
- Varantola, K. (1994). "The Dictionary User as Decision Maker". En W. Martin, W. Meijs, M. Moerland, E. Ten Pas, P. Van Sterkenburg y P. Vossen (eds.), *6th EURALEX International Congress*, Amsterdam, 30 de agosto-3 de septiembre de 1994, 606-611. Disponible en [http://www.euralex.org/elx\\_proceedings/Euralex1994/](http://www.euralex.org/elx_proceedings/Euralex1994/) [Consulta: 8 de mayo de 2012].
- Varantola, K. (1998). "Translators and their Use of Dictionaries: User Needs and User Habits". En B. T. S. Atkins (ed.), *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*. Tübingen: Niemeyer, 179-192.
- Varantola, K. (2000). "Translators, dictionaries and text corpora". En S. Bernardini y F. Zanettin (eds.), *I corpora nella didattica della traduzione - Corpus use and learning to translate*. Bologna: CLUEB, 117-.
- Varantola, K. (2002a). "Disposable Corpora as Intelligent Tools in Translation". En S. E. O. Tagnin (ed.), *Cadernos de Tradução: Corpora e Tradução* (Vol. 1). Florianópolis: NUT, 171-189.
- Varantola, K. (2002b). "Use and Usability of Dictionaries: Common Sense and Context Sensibility?" En M.-H. Corréard (ed.), *Lexicography and Natural Language Processing: A Festschrift in Honour of B. T. S. Atkins*. Grenoble: EURALEX, 30-44.
- Varantola, K. (2003a). "Translators and Disposable Corpora". En F. Zanettin, S. Bernardini y D. Stewart (eds.), *Corpora in Translator Education*. Manchester: St. Jerome, 55-70.
- Varantola, K. (2003b). "Translators and their Use of Dictionaries: User Needs and User Habits". En R. R. K. Hartmann (ed.), *Lexicography. Dictionaries, Compilers, Critics, and Users* (Vol. I). London: Routledge, 336-354.
- Varantola, K. (2006). "The Contextual Turn in Learning to Translate". En L. Bowker (ed.), *Lexicography, Terminology, and Translation*. Ottawa: University of Ottawa Press, 215-226.
- Vargas Sierra, C. (2002). "Utilización de los programas de concordancias en la traducción especializada". En P. Hernández y L. González (eds.), *El español, lengua de traducción. Actas del I Congreso Internacional*, Almagro, 468-483. Disponible en <http://www.ua.es/personal/chelo.vargas/Documentos/Vargas.pdf> [Consulta: 8 de mayo de 2012].
- Vargas Sierra, C. (2006). "Diseño de un corpus especializado con fines terminográficos: el corpus de la piedra natural". *Revista Debate Terminológico*, 2. Disponible en <http://riterm.net> [Consulta: 8 de mayo de 2012].
- Vázquez Béjar, D. (2008). "La competencia terminológica: su lugar en la formación de traductores". En M. T. Cabré, R. Estopà y C. Tebé (eds.), *La terminología en el siglo XXI: contribución a la cultura de la paz, la diversidad y la sostenibilidad. Actas del IX Simposio Iberoamericano de Terminología RITERM04*. Barcelona: Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, 713-720.
- Velásquez, G. (2002). "La traducción y la terminología en la comunicación bilingüe mediada". *Meta*, 47(3), 444-459. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Verlinde, S. (2011). "Modelling Interactive Reading, Translation and Writing Assistants". En P. A. Fuertes-Olivera y H. Bergenholtz (eds.), *e-Lxicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum, 275-286.

- Verlinde, S. y Binon, J. (2009). "Pedagogical Lexicography Revisited". En H. Bergenholtz, S. Nielsen y S. Tarp (eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 69-89.
- Verlinde, S., Leroyer, P. y Binon, J. (2010). "Search and You Will Find. From Stand-Alone Lexicographic Tools to User Driven Task and Problem-Oriented Multifunctional Leximats". *International Journal of Lexicography*, 23(1), 1-17.
- Vossen, G. y Hagemann, S. (2007). *Unleashing Web 2.0: From Concepts to Creativity*. Burlington: Morgan Kaufmann.
- Walker, D. G. (1993). "Translation Problems as They Occur in Everyday Practice". En K. D. Schmitz (ed.), *3rd International Congress on Terminology and Knowledge Engineering (TKE 1993)*, Colonia, 221-224.
- Walter, S. (2006). "Efficient Web Search Strategies or How to Use Google in an Appropriate Manner in a T&I Context". *ProZ.com Regional Conference*, Edinburgh. Disponible en [http://static.proz.com/event\\_resources/event8\\_presentation103.pdf](http://static.proz.com/event_resources/event8_presentation103.pdf) [Consulta: 8 de mayo de 2012].
- Weitzman, E. A. (1999). "Analyzing qualitative data with computer software". *Health Services Research*, 34(5 Part 2), 1241-1263. Disponible en <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1089062/?page=1> [Consulta: 8 de mayo de 2012].
- Weitzman, E. A. (2000). "Software and Qualitative Research". En N. K. Denzin y Y. S. Lincoln (eds.), *Handbook of Qualitative Research*. London: Sage Publications, 803-820.
- Welker, H. A. (2010). *Dictionary Use: A General Survey of Empirical Studies*. Brasília: Author's Edition.
- Wengraf, T. (2001). *Qualitative Research Interviewing. Biographic Narrative and Semi-Structured Methods*. London: Sage Publications.
- Wheatley, A. (2006). *Corpora & E-learning questionnaire. Results Summary*. MeLLANGE Multilingual eLearning in LANGuage Engineering. Disponible en <http://mellange.eila.jussieu.fr/Mellange-Results-1.pdf> [Consulta: 8 de mayo de 2012].
- Wiegand, H. E. (1998). *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie* (Vol. 1). Berlin/New York: de Gruyter.
- Wiegand, H. E. (2001). "Was eigentlich sind Wörterbuchfunktionen? Kritische Anmerkungen zur neueren und neuesten Wörterbuchforschung". *Lexicographica*, 17, 217-248.
- Wilkinson, M. (2005a). "Discovering Translation Equivalents in a Tourism Corpus by Means of Fuzzy Searching". *Translation Journal*, 9(4). Disponible en <http://translationjournal.net/journal/34corpus.htm> [Consulta: 8 de mayo de 2012].
- Wilkinson, M. (2005b). "Using a Specialized Corpus to Improve Translation Quality". *Translation Journal*, 9(3). Disponible en <http://translationjournal.net/journal/33corpus.htm> [Consulta: 8 de mayo de 2012].
- Wilkinson, M. (2006a). "Compiling Corpora for Use as Translation Resources". *Translation Journal*, 10(1). Disponible en <http://accurapid.com/Journal> [Consulta: 8 de mayo de 2012].
- Wilkinson, M. (2006b). "The corpus analysis tool - an under-exploited translation aid". *Kääntäjä*, 7.

- Wilkinson, M. (2007). "Corpora, Serendipity & Advanced Search Techniques". *JoSTrans. The Journal of Specialised Translation*, 7. Disponible en [http://www.jostrans.org/issue07/art\\_wilkinson.php](http://www.jostrans.org/issue07/art_wilkinson.php) [Consulta: 8 de mayo de 2012].
- Wilkinson, M. (2010). "Quick Corpora Compiling. Using Web as Corpus". *Translation Journal*, 14(3). Disponible en <http://translationjournal.net/journal//53corpus.htm> [Consulta: 8 de mayo de 2012].
- Williams, I. A. (1996). "A Translator's Reference Needs: Dictionaries or Parallel Texts?" *Target*, 8(2), 275-299.
- Winston, M. E., Chaffin, R. y Herrmann, D. (1987). "A Taxonomy of Part-Whole Relations". *Cognitive Science*, 11(4), 417-444.
- Woods, W. A. (2004). "Searching versus Finding: Why Systems Need Knowledge to Find What You Really Want". *ACM Queue Magazine*, 2(2), 27-35. Disponible en <http://queue.acm.org/> [Consulta: 8 de mayo de 2012].
- Wright, S. E. y Budin, G. (1997). *Handbook of Terminology Management. Basic Aspects of Terminology Management* (Vol. 1). Amsterdam/Philadelphia: John Benjamins.
- Wright, S. E. y Wright, L. D. (1997). "Terminology Management for Technical Translation". En S. E. Wright y G. Budin (eds.), *Handbook of Terminology Management. Basic Aspects of Terminology Management* (Vol. 1). Amsterdam/Philadelphia: John Benjamins, 147-159.
- Yuste, E. (2002). "Language resources and the language professional". En E. Yuste (ed.), *1st International Workshop on Language Resources for Translation Work and Research (LREC 2002 Workshop Proceedings)*, Las Palmas de Gran Canaria, 33-37.
- Zanettin, F. (1998). "Bilingual Comparable Corpora and the Training of Translators". *Meta*, 43(4), 616-630. Disponible en <http://www.erudit.org/revue/meta/> [Consulta: 8 de mayo de 2012].
- Zanettin, F. (1999). *Iper testi corpora e traduzioni: la narrativa di Salman Rushdie in italiano*. (Tesis doctoral), Università degli Studi di Bologna, Bologna.
- Zanettin, F. (2000). "Parallel Corpora in Translation Studies: Issues in Corpus Design and Analysis". En M. Olohan (ed.), *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*. Manchester: St. Jerome, 103-118.
- Zanettin, F. (2001a). "IperGrimus. Ipermedia e traduzione". *inTRAlinea*, Hypermedia. Disponible en <http://www.intralineait> [Consulta: 8 de mayo de 2012].
- Zanettin, F. (2001b). "Swimming in Words: Corpora, Translation, and Language Learning". En G. Aston (ed.), *Learning with Corpora*. Bologna: CLUEB, 177-197.
- Zanettin, F. (2002a). "Corpora in Translation Practice". En E. Yuste (ed.), *1st International Workshop on Language Resources for Translation Work and Research (LREC 2002 Workshop Proceedings)*, Las Palmas de Gran Canaria, 10-14. [15 de julio de 2008].
- Zanettin, F. (2002b). "DIY Corpora: The WWW and the Translator". En B. Maia, J. Haller y M. Ulrych (eds.), *Training the Language Services Provider for the New Millennium*. Porto: Universidade do Porto, 239-248.

## Recursos analizados

- About* [en línea]. Disponible en <http://www.about.com/>
- Antconc* (3.2.4w). Anthony, L. (2011). Disponible en <http://www.antlab.sci.waseda.ac.jp/software.html>
- Ask jeeves* [en línea]. Disponible en <http://uk.ask.com/>
- Base lexicale du français – BLF* [en línea]. Verlinde, S., Katholieke Universiteit Leuven. Disponible en <http://ilt.kuleuven.be/blf/>
- Beaucoup* [en línea]. Disponible en <http://www.beaucoup.com/>
- Bing* [en línea]. Disponible en <http://www.bing.com/>
- BNCweb (CQP-edition)* [en línea]. Lancaster University (1996-2008). Disponible en <http://bncweb.lancs.ac.uk>
- British National Corpus (BNC)* [en línea]. University of Oxford. Disponible en <http://www.natcorp.ox.ac.uk/>
- Buscopio* [en línea]. Disponible en <http://www.buscopio.net/>
- Bwananet* [en línea]. Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra. Disponible en <http://brangaene.upf.es/bwananet/indexes.htm>
- BYU-BNC: British National Corpus* (Based on the British National Corpus from Oxford University Press) [en línea]. Davies, M., Brigham Young University (2004-). Disponible en <http://corpus.byu.edu/bnc/>
- Cercaterm* [en línea]. Centre de Terminologia TERMCAT. Disponible en <http://www.termcat.cat/>
- CLUVI* [en línea]. Gómez Guinovart, X., Seminario de Lingüística Informática (SLI) da Universidade de Vigo (2003-). Disponible en <http://sli.uvigo.es/CLUVI/>
- Collins BETA* [en línea]. Collins (2011). Disponible en <http://www.collinsdictionary.com/>
- Collins English Dictionary* (Electronic Edition 1.5) [CD-ROM]. HarperCollins Publishers.
- Compara* [en línea]. Linguateca. Disponible en <http://linguateca.pt/COMPARA/>
- Copernic Agent Personal* (6.20). Copernic Inc. Disponible en <http://www.copernic.com>
- Corpógrafo V4* [en línea]. Linguateca. Disponible en <http://www.linguateca.pt/corpografo/>
- Corpus de referencia del español actual (CREA)* [en línea]. Real Academia Española. Disponible en <http://www.rae.es>
- Corpus del Español: 100 million words, 1200s-1900s* [en línea]. Davies, M., Brigham Young University (2002-). Disponible en <http://www.corpusdelespanol.org>
- Corpus diacrónico del español (CORDE)* [en línea]. Real Academia Española. Disponible en <http://www.rae.es>
- Corpus do Português: 45 million words, 1300s-1900s* [en línea]. Davies, M., Brigham Young University (2006-). Disponible en <http://www.corpusdoportugues.org>
- Déjà Vu X2*. ATRIL (2011). Disponible en <http://www.atril.com/>
- Diatopix 2.1* [en línea]. Drouin, P. Disponible en [http://olst.ling.umontreal.ca/~drouinp/diatopix/index\\_es.html](http://olst.ling.umontreal.ca/~drouinp/diatopix/index_es.html)
- Diccionario de colocaciones del español (DiCE)* [en línea]. Alonso Ramos, M., Grupo DiCE, Universidade da Coruña. Disponible en <http://www.dicesp.com/>
- Diccionario de la lengua española* (22.<sup>a</sup> edición) [en línea]. Real Academia Española. Disponible en <http://www.rae.es>
- Diccionario de la lengua española* (Edición electrónica 21.1.0) [CD-ROM]. Real Academia Española (1995). Madrid: Espasa Calpe.
- Diccionario de la llengua catalana* (Segona edició) [en línea]. Institut d'Estudis Catalans. Disponible en <http://dlc.iec.cat/>
- Diccionario de uso del español (DUE)* [CD-ROM]. Moliner, M. (2002). Madrid: Gredos.



- DiCoEnviro (Le dictionnaire fondamental de l'environnement)* [en línea]. Observatoire de Linguistique Sens-Texte (OLST), Université de Montréal. Disponible en <http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search-enviro.cgi?ui=es>
- DiCoInfo (Le dictionnaire fondamental de l'informatique et de l'internet)* [en línea]. Observatoire de Linguistique Sens-Texte, Université de Montréal. Disponible en <http://olst.ling.umontreal.ca/dicoinfo/>
- DiCoInfo Visuel* [en línea]. Observatoire de Linguistique Sens-Texte (OLST), Université de Montréal. Disponible en <http://olst.ling.umontreal.ca/dicoinfo/visuel.php>
- Dicouèbe (Dictionnaire en ligne de combinatoire du français)* [en línea]. Observatoire de Linguistique Sens-Texte (OLST), Université de Montréal (2005). Disponible en <http://olst.ling.umontreal.ca/dicouebe/>
- Dictionnaire analytique de la distribution* [en línea]. Dancette, J. (2000). Disponible en <http://olst.ling.umontreal.ca/dad/>
- Dictionnaire d'Apprentissage du Français des Affaires (DAFA)* [en línea]. GRELEP (2001). Disponible en <http://www.projetdafa.net/>
- Dirae* [en línea]. Rodríguez Alberich, G. y Real Academia Española. Disponible en <http://dirae.es/>
- Dogpile* [en línea]. Disponible en <http://www.dogpile.com/>
- EcoLexicon* [en línea]. LexiCon Research Group, Universidad de Granada. Disponible en <http://ecolexicon.ugr.es/en/>
- EOHS Term* [en línea]. ISPEL. Disponible en <http://eohsterm.org/>
- EUR-Lex* [en línea]. Oficina de Publicaciones de la Unión Europea. Disponible en <http://eur-lex.europa.eu/es/index.htm>
- EuroWordNet*. University of Amsterdam. Disponible en <http://www.illc.uva.nl/EuroWordNet/>
- Finderseeker: The Search Engine for Search Engines* [en línea]. Disponible en <http://www.finderseeker.com/>
- Foreigndesk* [en línea]. Disponible en <http://www.foreigndesk.net/>
- FrameNet* [en línea]. International Computer Science Institute (2000-). Disponible en <https://framenet.icsi.berkeley.edu/fndrupal/>
- Genoma* [en línea]. IULATERM, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra (2003). Disponible en <http://genoma.iula.upf.edu:8080/genoma/index.jsp>
- Goodrae* [en línea]. Abad, S. Disponible en <http://goodrae.es/>
- Google* [en línea]. Disponible en <http://www.google.com>.
- Grand dictionnaire terminologique* [en línea]. Office québécois de la langue française, Gouvernement du Québec (2002). Disponible en <http://www.oqlf.gouv.qc.ca/ressources/gdt.html>
- Hellenic National Corpus (HNC) (3.0)* [en línea]. Institute for Language and Speech Processing (1999-2009). Disponible en <http://hnc.ilsp.gr/en/>
- IATE (InterActive Terminology for Europe)* [en línea]. European Communities (2007). Disponible en <http://iate.europa.eu>
- Jaguar* [en línea]. Disponible en <http://melot.upf.edu/cgi-bin/jaguar/jaguar.pl>
- Just The Word* [en línea]. Sharp Laboratories of Europe (2005-2011). Disponible en <http://www.just-the-word.com/>
- Keybot* [en línea]. Disponible en <http://www.keybot.com/>
- KWiCFinder*. Fletcher, W. H. (2007). Disponible en <http://www.kwicfinder.com/KWiCFinder.html>
- Le migou* [en línea]. Observatoire de Linguistique Sens-Texte (OLST), Université de Montréal. Disponible en <http://olst.ling.umontreal.ca/migou/>

- Le Trésor de la Langue Française informatisé* [en línea]. Atilf. Disponible en <http://atilf.atilf.fr/tlf.htm>
- LEXA: Corpus Processing Software* [en línea]. Hickey, R., Universität GH Essen. Disponible en <http://icame.uib.no/lexainf.html>
- Lexical Freenet* (2.0) [en línea]. Datamuse (2003). Disponible en <http://www.lexfn.com>
- Linguee* [en línea]. Linguee GmbH. Disponible en <http://www.linguee.es/>
- Lycos* [en línea]. Disponible en <http://www.lycos.com/>
- Macmillan English Dictionary* [en línea]. Macmillan Publishers Limited (2009-). Disponible en <http://www.macmillandictionary.com>
- MemoQ*. Kilgray. Disponible en <http://kilgray.com/products/memoq>
- Merriam-Webster online* [en línea]. Merriam-Webster, Incorporated. Disponible en <http://www.merriam-webster.com>
- Metacrawler: Search the search engines!* [en línea]. Disponible en <http://www.metacrawler.com/>
- Metasearch* [en línea]. Disponible en <http://metasearch.com/>
- MetaTaxis*. Bruns, H. Disponible en <http://www.metataxis.net/espanol/>
- MicroConcord*. Scott, M. y Johns, T. (1993). Oxford: Oxford University Press.
- Monoconc Pro*. Barlow, M., Athel. Disponible en <http://www.athel.com/mono.html>
- Olster* [en línea]. Observatoire de Linguistique Sens-Texte (OLST), Université de Montréal (2010). Disponible en <http://olst.ling.umontreal.ca/olster/>
- OmegaT* [en línea]. Disponible en <http://www.omegat.org/>
- OncoTerm: Sistema Bilingüe de Información y Recursos Oncológicos* [en línea]. Grupo de Investigación OncoTerm, Universidad de Granada. Disponible en <http://www.ugr.es/~oncoterm/>
- OneLook Reverse Dictionary* [en línea]. Beeferman, D. (2003). Disponible en <http://www.onelook.com/reverse-dictionary.shtml>
- Oneseeek* [en línea]. Disponible en <http://www.oneseeek.com>
- Open Language Tools*. Oracle. Disponible en <http://open-language-tools.java.net/>
- OPUS open parallel corpus* [en línea]. Tiedemann, J. Disponible en <http://opus.lingfil.uu.se/index.php>
- Oxford English Dictionary* (Electronic Edition. Second Edition. 1.0) [CD-ROM]. Oxford University Press.
- Paraconc*. Barlow, M., Athel. Disponible en <http://www.athel.com/mono.html>
- Phrases in English* [en línea]. Fletcher, W. H. Disponible en <http://phrasesinenglish.org/>
- Scientext* [en línea]. Agence Nationale de la Recherche (ANR), LIDILEM (Grenoble 3), LiCoRN (Bretagne sud), LLS (Savoie). Disponible en <http://scientext.msh-alpes.fr>
- SDL Trados Studio 2009*. SDL (2009).
- Simple Concordance Program*. Reed, A. (1997-). Disponible en <http://www.textworld.com/scp/>
- Sketch Engine* [en línea]. Lexical Computing Ltd. Disponible en <http://www.sketchengine.co.uk/>
- Taus search* [en línea]. TAUS Data Association. Disponible en <http://www.tausdata.org/index.php/home>
- Teoma* [en línea]. IAC Search & Media. Disponible en <http://www.teoma.com/>
- TerminoWeb 2.0* [en línea]. Barrière, C., Interactive Language Technologies Group, NRC Institute for Information Technologies (2009). Disponible en <http://terminoweb.iit.nrc.ca/TE.html>
- TERMIUM Plus* [en línea]. Government of Canada. Disponible en <http://www.btb.termiumplus.gc.ca>
- Term-minator* [en línea]. Riediger, H. Disponible en <http://www.term-minator.it/>

- Terra* [en línea]. Disponible en <http://www.terra.com/>
- Text Analysis Computing Tools (TACT)*. University of Toronto. Disponible en <http://projects.chass.utoronto.ca/tact/>
- TextSTAT* [en línea]. Hüning, M., Freie Universität Berlin. Disponible en <http://neon.niederlandistik.fu-berlin.de/en/textstat/>
- The Corpus of Contemporary American English: 425 million words, 1990-present* [en línea]. Davies, M., Brigham Young University (2008-). Disponible en <http://corpus.byu.edu/coca/>
- TIME Magazine Corpus: 100 million words, 1920s-2000s* [en línea]. Davies, M., Brigham Young University (2007-). Disponible en <http://corpus.byu.edu/time/>
- Transit*. STAR. Disponible en <http://www.star-group.net/>
- Transolution* [en línea]. Disponible en <http://sourceforge.net/projects/eviltrans/files/Transolution/>
- Ultralingua* [en línea]. Ultralingua, inc. Disponible en <http://ultralingua.com/onlinedictionary/index.html>
- UNTERM* [en línea]. United Nations. Disponible en <http://unterm.un.org/>
- Web Concordancer* [en línea]. Disponible en <http://webascorpus.org/searchwac.html>
- WebCONC* [en línea]. Hüning, M., Freie Universität Berlin. Disponible en <http://www.niederlandistik.fu-berlin.de/cgi-bin/web-conc.cgi>
- WebCorp* [en línea]. Research and Development Unit for English Studies, Birmingham City University (1999-). Disponible en <http://www.webcorp.org.uk>
- Webcrawler* [en línea]. Disponible en <http://www.webcrawler.com/>
- Webitext* [en línea]. Terminotix Inc. Disponible en <http://www.webitext.com/bin/webitext.cgi>
- Wisnut*. Disponible en <http://en.wisnut.com/>
- WordCruncher*. Brigham Young University. Disponible en <http://wordcruncher.byu.edu/wordcruncher/default.htm>
- Wordfast*. Wordfast LLC. Disponible en <http://www.wordfast.net/>
- WordNet 3.1* [en línea]. Princeton University. Disponible en <http://wordnet.princeton.edu/wordnet/>
- WordReference Online Language Dictionaries* [en línea]. Disponible en <http://wordreference.com>.
- WordSmith Tools 4*. Scott, M. (2004). Oxford: Oxford University Press.
- Wordsmyth* [en línea]. Wordsmyth (2011). Disponible en <http://www.wordsmyth.net>
- WWW Search Interfaces for Translators* [en línea]. Harvey Ciampi, T. Disponible en <http://www.multilingual.ch/>
- Yahoo!* [en línea]. Disponible en [www.yahoo.com/](http://www.yahoo.com/)