

Departament de Traducció i d'Interpretació

UAB

Universitat Autònoma de Barcelona

**USO DE CORPUS ELECTRÓNICOS
EN LA FORMACIÓN DE TRADUCTORES
(INGLÉS-ESPAÑOL-INGLÉS)**

Volumen 1

TESIS DOCTORAL

presentada por

PATRICIA RODRÍGUEZ INÉS

Dirigida por

Amparo Hurtado Albir

y

Allison Beeby

Barcelona, 2008

*A mi padre,
que me enseñó a caminar contando los pasos en inglés.*

*A mi madre,
que siempre me animó a seguir caminando.*

*A Paul,
con quien camino cada día.*

AGRADECIMIENTOS

En primer lugar, quisiera dar las gracias a las directoras de esta tesis, Amparo Hurtado y Allison Beeby, por poner a mi disposición su saber y su experiencia para que esta tesis sea hoy una realidad. El apoyo moral, académico y personal que me han brindado hace de ellas unas mentoras con las que siempre estaré en deuda.

Agradezco también a aquellas compañeras que en mayor o menor medida han contribuido a esta tesis, ya sea solucionando alguna duda o simplemente dando ánimos. En este sentido, debo agradecer a la profesora Olivia Fox toda la confianza que siempre ha depositado en mí y a todos los miembros del grupo PACTE por su compañerismo.

Gracias al Departament de Traducció i d'Interpretació de la Universitat Autònoma de Barcelona y al Comissionat per a Universitats i Recerca del Departament de la Presidència de la Generalitat de Catalunya por el apoyo recibido. Gracias también a los organismos oficiales que me concedieron una beca predoctoral de formación y diversas estancias de investigación de suma importancia para la elaboración de esta tesis. Las cinco universidades internacionales punteras en aplicación de corpus a la traducción donde he disfrutado de estancias de investigación y a las que me gustaría agradecer su hospitalidad son: Université Paris 7 Denis-Diderot, Università di Bologna, Universidade do Porto, Tampereen Yliopistoon y la Universidade Federal de Minas Gerais.

Abreviaturas

En el cuerpo de la tesis se utilizan las siguientes abreviaturas:

CT = competencia traductora / TC = *translation competence*

FTI = Facultat de Traducció i d'Interpretació

LM = lengua meta

LO = lengua original

TM = texto meta o texto traducido

TO = texto original

TP = texto paralelo

UAB = Universitat Autònoma de Barcelona

UD = unidad didáctica

ÍNDICE DE CONTENIDOS

<u>INTRODUCCIÓN</u>	1
<u>I. LA LINGÜÍSTICA DE CORPUS Y SUS APLICACIONES</u>	
1. LA LINGÜÍSTICA DE CORPUS	17
1.1. Definición de corpus	17
1.2. Definición de lingüística de corpus	18
1.3. Ámbito y evolución de la lingüística de corpus.....	19
1.3.1. Críticas desde la lingüística generativa	20
1.4. Fases del trabajo con corpus.....	23
1.4.1. Establecimiento del objeto de estudio e hipótesis.....	24
1.4.2. Selección de textos/corpus.....	25
1.4.2.1. Tipos de corpus.....	26
1.4.2.1.1. Según el porcentaje y la distribución de los diferentes tipos de texto.....	27
1.4.2.1.2. Según la especificidad de los documentos.....	27
1.4.2.1.3. Según las cantidad de texto en cada uno de los documentos	28
1.4.2.1.4. Según la codificación y la anotación.....	28
1.4.2.1.5. Según la documentación que acompaña a los textos.....	28
1.4.2.2. Procedimientos de selección.....	29
1.4.2.3. Cuestiones de almacenamiento	30
1.4.2.4. Derechos de autor.....	31
1.4.2.5. Captura de textos	33
1.4.2.6. Tamaño	33
1.4.2.7. Representatividad	34
1.4.3. Preparación del corpus.....	35
1.4.3.1. Anotación gramatical.....	38
1.4.3.2. Lematización	39
1.4.3.3. Anotación sintáctica	40
1.4.3.4. Anotación semántica	40
1.4.4. Análisis	41
1.4.4.1. Operaciones básicas.....	42
1.4.4.2. Herramientas informáticas	42
1.4.5. Explotación y reelaboración de hipótesis	46

2. CORPUS ELECTRÓNICOS EN ESPAÑA Y TRADUCCIÓN	47
2.1. Corpus escritos.....	48
2.1.1. Corpus generales	48
2.1.2. Corpus específicos	49
2.2. Corpus orales	51
2.2.1. Corpus generales	51
2.2.2. Corpus específicos	52
2.3. Tesis doctorales relacionadas con la lingüística de corpus en traducción.....	53
2.4. Utilización de corpus en el mundo profesional de la traducción: agencias de traducción.....	55
3. CAMPOS DE APLICACIÓN	59
3.1. Lexicografía y descripciones de la lengua	60
3.2. Lingüística computacional	62
3.3. Investigación del habla oral	63
3.4. Pragmática y análisis del discurso	63
3.5. Sociolingüística.....	65
3.6. Lingüística histórica	66
3.7. Dialectología y variación lingüística	67
3.8. Estilística y estudios literarios	67
3.9. Psicolingüística y psicología social	68
3.10. Adquisición de la lengua	69
3.11. Didáctica de lenguas.....	69
3.11.1. Antecedentes	70
3.11.2. Aproximaciones al uso de corpus.....	74
3.11.3. Usos de los corpus	79
3.11.3.1. Diseño de materiales didácticos.....	81
3.11.3.2. Diseño de temarios.....	82
3.11.3.3. Complemento de materiales de referencia ya existentes.....	82
3.11.3.4. Aprendizaje autónomo.....	83
3.11.3.5. Fuente de actividades de aprendizaje.....	83
3.11.4. Situación actual y necesidades	85
4. USO DE CORPUS EN LA DIDÁCTICA DE LA TRADUCCIÓN	89
4.1. Ventajas de los corpus electrónicos frente a diccionarios y textos paralelos en papel	90
4.2. Tipos de corpus más relevantes para la didáctica de la traducción escrita	92
4.2.1. Corpus paralelo	93

4.2.2. Corpus comparable bilingüe.....	95
4.2.3. Corpus comparable monolingüe (o corpus especializado en lengua meta) ...	100
4.2.4. Corpus ad hoc, virtual, desechable o de fabricación propia	102
4.2.5. Estrategias de búsqueda de equivalentes de traducción en un corpus especializado monolingüe en lengua meta	106
4.3. Pruebas de eficacia	111
4.4. Situación actual y necesidades.....	113

II. MARCO CONCEPTUAL PARA LA UTILIZACIÓN DE LA METODOLOGÍA DE CORPUS EN LA FORMACIÓN DE TRADUCTORES

5. LA DIDÁCTICA DE LA TRADUCCIÓN ESCRITA.....	119
5.1. La competencia traductora y su adquisición.....	119
5.1.1. La competencia traductora.....	119
5.1.1.1. La investigación del grupo PACTE.....	120
5.1.2. La adquisición de la competencia traductora.....	122
5.2. Traducción pedagógica.....	124
5.3. Enfoques de la didáctica de la traducción	125
5.3.1. La didáctica tradicional de la traducción	125
5.3.2. Los estudios contrastivos	126
5.3.3. Los tratados esencialmente teóricos.....	127
5.3.4. La enseñanza por objetivos de aprendizaje.....	128
5.3.5. Los enfoques profesionales	130
5.3.6. El constructivismo social	131
5.3.7. El enfoque por tareas de traducción	131
5.3.7.1. El enfoque por tareas en la enseñanza de lenguas.....	131
5.3.7.2. El enfoque por tareas de traducción.....	132
5.3.8. La formación por competencias	134
5.3.8.1. La formación por competencias.....	134
5.3.8.2. La formación por competencias en traducción	138
5.3.8.2.1. La importancia del perfil profesional.....	138
5.3.8.2.2. La formación por competencias en traducción.....	139
6. MARCO METODOLÓGICO Y DE DISEÑO CURRICULAR PARA EL USO DE CORPUS EN LA FORMACIÓN DE TRADUCTORES.....	143
6.1. Competencia traductora y uso de corpus para traducir: ubicación y componentes.....	143
6.2. Objetivos de aprendizaje.....	145

6.2.1.	Fuentes de los objetivos.....	145
6.2.2.	Objetivos relacionados con el uso de corpus para traducir	147
6.2.2.1.	Objetivos de la iniciación en el uso de corpus para traducir.....	148
6.2.2.2.	Objetivos de la consolidación en el uso de corpus para traducir.....	149
6.2.3.	Objetivos relacionados con competencias genéricas	150
6.2.4.	Objetivos relacionados con materias de traducción.....	151
6.2.4.1.	Bases profesionales para el diseño de objetivos	151
6.2.4.2.	Objetivos relacionados con materias de traducción	153
6.3.	Contenido disciplinario	155
6.3.1.	Desarrollo del contenido	159
6.4.	Marco metodológico.....	165
6.5.	Materias de aplicación	166
6.5.1.	La traducción directa	167
6.5.1.1.	Iniciación en traducción directa.....	167
6.5.1.2.	Traducción directa especializada.....	167
6.5.2.	La traducción inversa	169
6.5.2.1.	Iniciación en traducción inversa.....	170
6.5.2.2.	Traducción inversa especializada.....	170
6.6.	Unidades didácticas	172
6.6.1.	Criterios para la elaboración de las unidades didácticas.....	172
6.6.1.1.	Criterios generales.....	172
6.6.1.1.1.	Variedad.....	172
6.6.1.1.2.	Progresión.....	175
6.6.1.2.	Criterios específicos.....	177
6.6.1.2.1.	Selección de objetivos de aprendizaje	177
6.6.1.2.2.	Selección de materiales	181
6.6.1.2.2.1.	Selección de textos.....	182
6.6.1.2.2.2.	Selección de recursos lexicográficos.....	185
6.6.1.2.2.3.	Elaboración de fichas	185
6.6.1.2.2.4.	Selección de tipos de corpus	186
6.6.1.2.2.5.	Selección de software.....	188
6.6.1.2.2.6.	Ejemplo de selección de materiales de una unidad.....	189
6.7.	Evaluación	195
6.7.1.	La evaluación como fuente de información.....	196
6.7.2.	Indicadores	198
6.7.3.	Procedimiento de evaluación del uso de corpus cuando se traduce un texto ...	199
6.7.4.	Herramientas de evaluación.....	200
6.7.4.1.	Herramientas de evaluación del aprendizaje	201
6.7.4.1.1.	Diario de aprendizaje	201
6.7.4.1.2.	Ficha de autoevaluación	203
6.7.4.1.3.	Ficha de recogida de datos sobre el uso de corpus al traducir	204
6.7.4.1.4.	Ficha de cruce de datos	205

6.7.4.1.5.	Ejemplo	208
6.7.4.2.	Herramientas de evaluación de la enseñanza	216
6.7.4.2.1.	Ficha de registro de observación	216
6.7.4.2.2.	Ficha de valoración de la unidad didáctica.....	217
6.8.	Evolución de la propuesta	218
6.8.1.	Fase previa al diseño	218
6.8.2.	Diseño y experimentación	219
6.8.3.	Propuesta definitiva.....	220

III. PROPUESTA DE UNIDADES DIDÁCTICAS

INTRODUCCIÓN.....	225
1. Procedimiento de presentación de las unidades didácticas	225
2. Unidades didácticas.....	229
(1) UD “Introducción al uso de corpus electrónicos para traducir”	
❶ Presentación	232
❷ Tareas.....	235
❸ Evaluación	244
❹ Evolución	250
(2) UD “Contrastes de uso en un corpus paralelo”	
❶ Presentación	252
❷ Tareas.....	254
❸ Evaluación	261
❹ Evolución	264
(3) UD “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”	
❶ Presentación	266
❷ Tareas.....	269
❸ Evaluación	279
❹ Evolución	289
(4) UD “Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”	
❶ Presentación	292
❷ Tareas.....	296
❸ Evaluación	301
❹ Evolución	303

(5) UD “El corpus ad hoc como remedio para la traducción médica al inglés”

- ① Presentación 306
- ② Tareas..... 309
- ③ Evaluación 316
- ④ Evolución 318

(6) UD “Corpus y memoria de traducción”

- ① Presentación 320
- ② Tareas..... 323
- ③ Evaluación 328
- ④ Evolución 333

(7) UD “Sólo utilizar corpus; utilizar corpus solo”

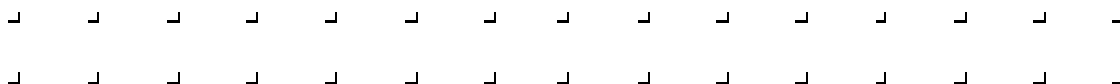
- ① Presentación 336
- ② Tareas..... 339
- ③ Evaluación 343
- ④ Evolución 350

CONCLUSIONES 353

GLOSARIO 375

ÍNDICE DE TABLAS, DOCUMENTOS Y FIGURAS 381

BIBLIOGRAFÍA 385



ÍNDICE DEL VOLUMEN 2

(1) UD “Introducción al uso de corpus electrónicos para traducir”	
📁 Materiales	413
Fichas de trabajo	413
Fichas de evaluación.....	420
Textos originales y paralelos.....	422
Textos de apoyo.....	437
Descripciones.....	439
↔ Soluciones	440
☑ Evaluación	450
(2) UD “Contrastes de uso en un corpus paralelo”	
📁 Materiales	459
Fichas de trabajo	459
Fichas de evaluación.....	462
Texto original y traducción	468
↔ Soluciones	470
☑ Evaluación	482
(3) UD “Utilidad de un corpus comparable bilingüe para traducir textos técnicos”	
📁 Materiales	487
Fichas de trabajo	487
Fichas de evaluación.....	493
Textos originales y paralelos.....	499
Descripciones.....	514
Concordancias	514
Listas	515
↔ Soluciones	516
☑ Evaluación	531

(4) UD “Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”

☰ Materiales	557
Fichas de trabajo.....	557
Fichas de evaluación.....	583
Textos originales y paralelos.....	585
Textos de apoyo.....	590
Descripciones.....	596
Concordancias	597
Colocaciones.....	597
↔ Soluciones	598
☑ Evaluación	613

(5) UD “El corpus ad hoc como remedio para la traducción médica al inglés”

☰ Materiales	617
Fichas de trabajo.....	617
Textos originales y paralelos.....	621
Descripciones.....	626
↔ Soluciones	627
☑ Evaluación	635

(6) UD “Corpus y memoria de traducción”

☰ Materiales	641
Fichas de trabajo.....	641
Fichas de evaluación.....	653
Textos originales y paralelos.....	655
Textos de apoyo.....	671
↔ Soluciones	672
☑ Evaluación	681

(7) UD “Sólo utilizar corpus; utilizar corpus solo”689

☰ Materiales	691
Fichas de trabajo.....	691
Fichas de evaluación.....	696
Texto original.....	701
Textos de apoyo.....	702
↔ Soluciones	704
☑ Evaluación	713

INTRODUCCIÓN

OBJETIVOS

El objetivo general de esta tesis doctoral es integrar el uso de corpus electrónicos en la enseñanza de la traducción. Para llegar a éste nos planteamos los siguientes objetivos específicos:

- 1) Elaborar el marco conceptual de la lingüística de corpus y sus campos de aplicación.
- 2) Elaborar el marco conceptual de la didáctica de la traducción apropiado para integrar la metodología de la lingüística de corpus.
- 3) Diseñar una propuesta de uso de corpus electrónicos en la enseñanza de la traducción.
- 4) Evaluar la propuesta diseñada de uso de corpus electrónicos en la enseñanza de la traducción.

MOTIVACIÓN DE LA INVESTIGACIÓN

Si entendemos que en la vida siempre hay más de un área que atrae nuestra atención y curiosidad, también lo es así en el ámbito académico. Lo ideal es, entonces, que podamos aunar en una investigación varias de esas áreas que despiertan nuestro interés. La lengua como instrumento de comunicación desde el punto de vista del traductor o del traductólogo, la investigación con base empírica en los estudios sobre la traducción en general y sobre la competencia traductora en particular, la traducción hacia la lengua extranjera o la materna y su enseñanza en diversas áreas de especialidad y las nuevas tecnologías aplicadas a la traducción (*Tradumática*), son todas áreas de interés personal. La lingüística de corpus, como metodología que es, permite investigar en varias áreas.

La principal motivación para la realización de este trabajo llegó con el descubrimiento de la metodología de la lingüística de corpus, sus técnicas y herramientas. Una primera incursión en el mundo de los corpus electrónicos tuvo su reflejo en la memoria de máster y el trabajo de investigación de la autora de esta tesis dedicados a la ideología en traducción (Rodríguez Inés 1999, 2000), dirigidos por las doctoras Mona Baker y Allison Beeby, respectivamente. Tras este trabajo de índole descriptiva y analítica, se

optó por continuar con una de las vertientes de la rama aplicada de los estudios sobre la traducción, la didáctica de la traducción, y el uso de corpus. La aplicación de esta metodología en las aulas es, hasta la fecha, bastante limitada y casi inexistente en el campo de la enseñanza de la traducción. En otras palabras, existe un vacío académico en este campo, lo que ha supuesto una de las principales razones para el inicio de esta investigación, que esperamos despierte el interés entre otros investigadores por su novedad o aplicabilidad.

IMPORTANCIA DEL ESTUDIO

Dentro de la aún joven disciplina de los estudios sobre la traducción, la aplicación de corpus no está sino en sus inicios en algunas de sus ramas, prueba de lo cual es que en varios libros sobre lingüística de corpus se mencionen campos diversos de aplicación, pero no nuestra disciplina y, en el caso de que se mencione la enseñanza de la traducción, se haga dentro del apartado de enseñanza de lenguas, como muy bien observa Olohan en su libro (2004: 13). De hecho, el uso de los corpus en la didáctica de la traducción es una de las áreas menos desarrolladas y que sólo cuenta con un número relativamente escaso de estudios puntuales, para lo cual pueden buscarse razones como la falta de conocimiento de la metodología, de manejo de la informática a nivel de usuario avanzado o de medios para ponerlo en práctica.

Por otro lado, la enseñanza de la traducción aún adolece en nuestro país de una incorporación contextualizada de las nuevas tecnologías en el aula. Se ha observado una necesidad creciente de aplicar recursos y herramientas informáticas dentro de las clases de traducción y no dejar esta carga únicamente para las asignaturas de informática aplicada. Con el fin de que se obtenga un aprendizaje significativo, esta tesis propone el uso de corpus electrónicos en las clases de traducción en combinación con otras herramientas informáticas y otros recursos electrónicos, sobre todo documentales, como los que se encuentran en Internet (diccionarios, glosarios, buscadores, etc.).

La propuesta de trabajo que aquí se expone apunta al establecimiento de una metodología de aplicación de corpus a la didáctica de la traducción, trabajo que sería pionero en lo que respecta a su envergadura y profundidad. Además, la parte de esta propuesta que se refiere a la evaluación es también novedosa en tanto que no se queda solamente en la evaluación de la traducción, es decir, del producto final, sino que se interesa por todo el proceso, desde la documentación al análisis y la fase previa a la

resolución final de un problema de traducción. Finalmente, esta tesis puede también considerarse un paso hacia las premisas del Espacio Europeo de Enseñanza Superior, centrado en el estudiante y en el aprendizaje y que pone un énfasis especial en el uso de las nuevas tecnologías y la autonomía del alumno.

ANTECEDENTES

Como ya se ha apuntado, la aplicación de corpus a la disciplina de los estudios sobre la traducción es aún reciente y limitada. Sin embargo, algunos de los campos que más se han beneficiado de la aplicación de corpus han sido la terminología, la lexicografía y la traducción automática. Otros campos afines que también han incorporado corpus a sus estudios han sido la lingüística en diversas ramas (desde la lingüística histórica al análisis del discurso), la enseñanza de lenguas, etc.

Desde el punto de vista de la aplicación de corpus, la didáctica de la traducción en general ha recibido hasta el momento una atención muy escasa, y en especial por parte de los docentes de traducción. Algunos profesores de lengua son responsables de algunos intentos de aplicación de corpus a la traducción que dejan ver un claro trasfondo filológico o de utilización de la traducción como medio para enseñar una lengua (traducción pedagógica), si bien es cierto que alguno de estos ejercicios puede ser de interés en las clases de traducción inversa.

Desde mediados de los años 90 se ha visto un creciente interés en el tema, dando como resultado la celebración de tres conferencias bajo el nombre de CULT (*Corpus Use and Learning to Translate*) en 1997 y 2000 en Bertinoro y en 2004 en Barcelona, donde se presentaron casos prácticos de aplicaciones y otras cuestiones relacionadas. La impresión es, sin embargo, que aquellos docentes que aplican corpus en sus clases de traducción lo hacen a título individual y sin seguir una metodología preestablecida que les asista. En este sentido, se indican a continuación algunas universidades en las que una persona, o a lo sumo un grupo muy reducido de personas, están implicadas o han mostrado su interés en la aplicación de corpus a la enseñanza de la traducción en sus respectivos idiomas: Università di Bologna, Universidad de Málaga, Universidade do Porto, Universitat Jaume I, Tampereen Yliopistoon (Universidad de Tampere), Université Paris 7 Denis-Diderot, Univerza v Ljubljani (Universidad de Liubiana), entre otras.

Respecto a los fundamentos pedagógicos sobre los que nos apoyamos, creemos que las palabras que mejor resumen el espíritu de este trabajo son las de Hurtado (1999).

“El diseño de cualquier enfoque didáctico supone plantear los objetivos que se persiguen, el/los método/s utilizado/s (la manera de organizar las actividades de aprendizaje), los medios empleados (materiales y objetos de todo tipo), la evaluación (criterios y tipos de pruebas), la progresión, la planificación del curso y de cada unidad didáctica...” (Hurtado 1999: 44)

En este sentido, y como ya se defenderá en el cuerpo de la tesis, el enfoque por tareas de traducción y el diseño de objetivos expresados en términos de competencias nos ha parecido ser especialmente idóneo para la enseñanza del uso de corpus electrónicos para traducir ya que aporta toda la sistematicidad de la que carecían los primeros ejercicios de uso de corpus diseñados. Tales primeros esfuerzos, llevados a cabo con herramientas limitadas en sus funciones, corpus limitados en su tamaño y manejabilidad, aulas limitadas por su disponibilidad y capacidad, y profesores limitados por el desamparo en la investigación en el campo de la didáctica de la traducción con corpus, merecen sin duda toda la atención, si bien es cierto que adolecen, por las razones expuestas, de una metodología bien establecida o de una sistematicidad en el diseño, aplicación y evaluación.

METODOLOGÍA

Tipo de investigación

El presente trabajo está enmarcado dentro de la rama aplicada de los Estudios sobre la Traducción, si bien también se nutre de la rama descriptiva. En el contexto de la investigación educativa, nos situamos en una de las perspectivas que distinguen Latorre, Del Rincón y Arnal (1996), “la perspectiva orientada a la práctica educativa, con una aplicación directa en la política o prácticas educativas y centrada en aportar información que guíe la toma de decisiones y los procesos de cambio” (en Hurtado 2001: 174). En nuestro caso estaríamos hablando de un tipo de investigación-acción, que sigue procesos de intervención en la práctica docente y cuyo objetivo es transformarla. El estudio que aquí se presenta se enmarca también dentro de la investigación empírica, usando métodos cualitativos ya que pretende describir la aplicación de corpus en la didáctica de la traducción y descubrir sus ventajas, a la vez que elaborar una propuesta didáctica y evaluar el funcionamiento de la misma. Por otro lado, esta aplicación tiene como objetivo general contribuir al desarrollo de la

competencia traductora en lo que se refiere a la parte instrumental y, en concreto, aportar mejoras a la formación en traducción para que esté menos basada en la experiencia del profesor y más centrada en el alumno, su proceso de aprendizaje, sobre todo a través de la reflexión y el descubrimiento, y el desarrollo de su autonomía.

Para llevar a cabo este trabajo ha sido necesario primero establecer el estado de la cuestión en dos áreas, la lingüística de corpus y su aplicación en la didáctica de lenguas y traducción, y la didáctica de la traducción por sí misma. Una vez sentadas las bases conceptuales, se han combinado la metodología, técnicas y herramientas de la lingüística de corpus con varios enfoques de la didáctica de la traducción para proporcionar una base multidisciplinaria a nuestra propuesta.

Como decimos, existe una metodología propia de la lingüística de corpus que también debe tenerse en cuenta a la hora de integrarla en el estudio. Hablando de la investigación con corpus, Partington resume el proceso con las siguientes palabras que se pueden hacer extensibles al trabajo con corpus en general:

“A researcher has an intuition about language, checks this against the data the corpus provides, and this checking process frequently suggests other avenues of research to be taken, often entirely unsuspected at the start of the process...” (Partington 1998: 1)

De forma más detallada, se puede decir que la metodología de corpus se basa en varios pilares o fases que siguen a continuación:

1. Planteamiento del objeto de estudio o del problema (de traducción) que se quiere resolver.
2. Selección del corpus o de los textos que lo van a componer.
3. Preparación de los textos, si cabe.
4. Búsqueda de datos, observación y análisis.
5. Generalización de resultados y conclusiones.
6. Replanteamiento del objeto de estudio, del problema o de las búsquedas realizadas.

Dentro de cada fase se cuenta con una serie de recursos y herramientas informáticas que ayudan a optimizar el trabajo y aplicar las diversas técnicas de corpus existentes (extracción de concordancias, estadísticas, listas de palabras, etc.), aspecto que se desarrolla en detalle en el cuerpo de esta tesis.

Fases

Las fases que se han seguido en la realización del presente trabajo son las siguientes:

1. FASE CONCEPTUAL

En el plano conceptual, se realizó un vaciado bibliográfico y lectura de las publicaciones más destacables sobre lingüística de corpus, con especial exhaustividad en la parte de aplicación de corpus a la didáctica de la traducción, y sobre didáctica de la traducción.

2. DISEÑO DE LA PROPUESTA

Para poder realizar una propuesta de aplicación de corpus en la formación de traductores fue necesario primero seleccionar un enfoque pedagógico de la propia didáctica de la traducción. En este sentido, el adoptado fue el enfoque por tareas de traducción (Hurtado 1992, 1993, etc.), así como su integración con la formación por competencias.

La propuesta que efectuamos consiste en la elaboración de un marco metodológico y de diseño curricular para el uso de corpus en la formación de traductores. Esta propuesta se concreta en una competencia específica, el uso de corpus electrónicos para traducir, que se ubica dentro de la subcompetencia instrumental de la competencia traductora. Esta competencia específica se define a través de cuatro componentes que a su vez se recogen en cuatro objetivos generales. Estos objetivos y, ulteriormente, la competencia en el uso de corpus para traducir, se alcanzan a través de la apropiación del contenido disciplinario asociado a dicha competencia, y que nosotros hemos desarrollado en este trabajo. También se han seleccionado las materias de traducción donde pueden encajar estos contenidos, los cuales hemos agrupado y organizado para producir varias unidades didácticas. Finalmente, la propuesta incluye también un diseño de evaluación de la propia propuesta.

3. APLICACIÓN Y VALORACIÓN DE LA PROPUESTA

Dado que realizamos una investigación-acción y utilizamos métodos cualitativos, tras el diseño de la propuesta se pusieron en práctica todas las unidades didácticas en el aula y se recogieron los datos de su evaluación. Las herramientas de recogida de información son variadas y tuvieron que ser diseñadas expresamente debido a la novedad del estudio. El primer paso fue establecer los indicadores para la evaluación de la competencia objeto de estudio. Se crearon

fichas diversas para recoger, por un lado, la evaluación del aprendizaje y, por otro, la evaluación de la enseñanza.

Para la evaluación del aprendizaje se diseñaron las siguientes herramientas:

- (1) Diario de aprendizaje, para que los alumnos recojan todos aquellos conocimientos y habilidades que van adquiriendo y el docente tenga acceso a sus reflexiones.
- (2) Ficha de autoevaluación, para que, al final de cada unidad didáctica, los alumnos puntúen su propia adquisición de contenidos según su percepción.
- (3) Ficha de recogida de datos sobre el uso de corpus al traducir, para que los alumnos reflejen las operaciones realizadas con un corpus durante una tarea de traducción.
- (4) Ficha de cruce de datos, para que el docente refleje la puntuación obtenida por el alumno respecto a su uso de corpus a la hora de traducir, a partir de los indicadores para la evaluación de la competencia, las operaciones realizadas por el alumno y la aceptabilidad de sus soluciones de traducción.

Para la evaluación de la enseñanza se diseñaron las siguientes herramientas:

- (1) Ficha de registro de observación, para que el docente recoja sus impresiones sobre la puesta en práctica de la propuesta.
- (2) Ficha de valoración de la unidad, para que los alumnos reflejen también de forma argumentada sus impresiones sobre la propuesta.

En el *Volumen 2* se ofrecen las pruebas o evidencias de la evaluación, es decir, las fichas de cruces de datos rellenas para cada alumno en su tarea final, si es que la unidad didáctica la contempla, y las respuestas redactadas (no numéricas) de todos los estudiantes a fichas como la de autoevaluación o valoración de la unidad.

Las siete unidades didácticas diseñadas se probaron con varios grupos de estudiantes diferentes. Al principio de cada unidad se especifica su ubicación, es decir, el grupo y la materia o asignatura en la que de hecho se puso en práctica, y en las secciones dedicadas a la evaluación se podrá comprobar que el número e iniciales de los alumnos evaluados cambian de una unidad a otra. La mayoría de los grupos pertenecen a la licenciatura en Traducción e Interpretación, pero también se trabajó con otros grupos. Creemos que es importante recalcar el hecho de que la autora de esta tesis ha podido poner a prueba en el aula todas las

unidades didácticas porque ha tenido acceso a estas materias, bien por medio del plan docente establecido, bien como horas extraordinarias en materias que pertenecían a otros colegas, bien en cursos de especialización como máster en los que la autora participa como docente, o bien gracias a una estancia de investigación que incluía horas de docencia.

Respecto a los grupos de prueba, hay que puntualizar que las 7 unidades se probaron total o parcialmente en los siguientes grupos y asignaturas:

- (1) Tres grupos de 3er curso de Licenciatura en Traducción e Interpretación, FTI, UAB; asignatura Traducción Especializada (inglés B-español): UD de traducción técnica directa.
- (2) Tres grupos de 4º curso de Licenciatura en Traducción e Interpretación, FTI, UAB; asignatura Seminario de Traducción Especializada (español-inglés B): UDs de traducción especializada inversa.
- (3) Un grupo de 4º curso de Licenciatura en Traducción e Interpretación, FTI, UAB; asignatura Seminario de Traducción (inglés C-español): UDs de iniciación a la traducción directa.
- (4) Dos grupos de 2º curso del Máster Oficial en Traducción y Mediación Intercultural, FTI, UAB; asignatura Traducción Especializada (español-inglés B): UD de concienciación sobre la traducción inversa.
- (5) Un grupo de último curso de Licenciatura en Letras, Faculdade de Letras, Universidade Federal de Minas Gerais; asignatura Traducción al inglés B: UD de concienciación sobre la traducción inversa.

Sin embargo, aunque las unidades se probaron en estos 10 grupos mencionados, los resultados de evaluación que se muestran en esta tesis pertenecen a 6 grupos y a 4 asignaturas:

- (1) Un grupo de 3er curso de Licenciatura en Traducción e Interpretación, FTI, UAB; asignatura Traducción Especializada (inglés B-español).
- (2) Tres grupos de 4º curso de Licenciatura en Traducción e Interpretación, FTI, UAB; asignatura Seminario de Traducción Especializada (español-inglés B).
- (3) Un grupo de 4º curso de Licenciatura en Traducción e Interpretación, FTI, UAB; asignatura Seminario de Traducción (inglés C-español).
- (4) Un grupo de 2º curso del Máster Oficial en Traducción y Mediación Intercultural, FTI, UAB; asignatura Traducción Especializada (español-inglés B).

4. ELABORACIÓN DE LA PROPUESTA DEFINITIVA

Tras la puesta en práctica de la propuesta didáctica y la valoración de la misma tanto por parte de los estudiantes como de la docente a través de las herramientas citadas, se introdujeron algunas modificaciones para llegar a elaborar una propuesta definitiva. Las modificaciones, que siempre se presentan en el apartado “Evolución” de cada unidad didáctica, han estado en la línea de reformular los enunciados de algunas fichas para facilitar su comprensión, la adición de tareas o postareas tras la comprobación de que algún objetivo no había sido plenamente alcanzado, o la inclusión de más recursos gratuitos, tanto documentales como de software, con el fin de facilitar su accesibilidad.

ESTRUCTURA DE LA TESIS

Esta tesis consta de dos volúmenes:

- El *Volumen 1*, que incluye el marco conceptual de la lingüística de corpus y la didáctica de la traducción escrita y el marco metodológico para la utilización de corpus en la formación de traductores.
- El *Volumen 2*, que incluye los materiales pertenecientes a las unidades didácticas diseñadas.

El contenido de la tesis se estructura en tres partes bien diferenciadas, la primera dedicada a la lingüística de corpus y sus aplicaciones, la segunda a la didáctica de la traducción escrita y la tercera a la propuesta didáctica concreta que efectuamos. En otras palabras, la lingüística de corpus y la didáctica de la traducción son las dos grandes áreas en las que se inscribe este trabajo.

Parte I

El *Volumen 1* empieza con una parte dedicada a la lingüística de corpus en general. Esta parte comienza en el *Capítulo 1* con una introducción al tema y las implicaciones del trabajo con corpus electrónicos como son sus fases, las cuestiones a tener en cuenta en cada fase y algunas de las herramientas informáticas que existen para automatizar procesos y facilitar el análisis.

Yendo de lo general a lo concreto, el *Capítulo 2* recoge lo más destacado del uso de corpus en España y en lengua española (variante peninsular). Ya en 2003 una empresa

de traducción con sede en Barcelona, Star, en su publicación digital *Starlette* llamaba la atención sobre la utilidad de los corpus electrónicos en la traducción en tanto que permiten “explorar una gran cantidad de texto e identificar patrones comunes que sería imposible capturar mediante análisis manual” (Star Servicios Lingüísticos S.L. 2003). Este capítulo también incluye un sondeo a pequeña escala sobre la práctica real de la recopilación de textos por agencias de traducción españolas, gracias al cual se comprueba que éstas trabajan con textos originales, traducidos y paralelos tanto para almacenarlos como para su posterior explotación. Por otro lado, los datos de Sánchez-Gijón (2003) obtenidos de una encuesta dirigida a traductores de empresa, traductores del Servicio de traducción de organismos europeos y traductores autónomos, muestra que éstos utilizan textos ampliamente para diversos propósitos documentales y que se apoyan sobre todo en Internet y en traducciones realizadas anteriormente. En conclusión, todos estos datos no pueden sino conducirnos a reflexionar y concluir que es indispensable entrenar a los alumnos de traducción en el manejo de corpus o de textos (originales con sus traducciones y paralelos), desde su evaluación y recogida hasta su aprovechamiento como fuente de documentación.

El *Capítulo 3* se centra en las aplicaciones de los corpus en diversas áreas, destacando aquellos casos donde existan ejemplos relacionados con la traducción. En este campo los corpus han interesado principalmente desde dos puntos de vista, uno descriptivo, sobre todo para establecer las características de los textos traducidos en comparación con sus originales o también por sí mismos para estudiar los universales de la traducción; y desde otro punto de vista, como ayuda en los procesos de la traducción humana y automática (Aston 1999: 290). Respecto a las aplicaciones de corpus hasta la fecha, mientras la lista de estudios que se han llevado a cabo con corpus es bastante extensa, se debe apuntar que la gran mayoría tratan la lengua inglesa y así, por ejemplo, existen numerosos estudios descriptivos del inglés que se centran en la gramática. Otras áreas dentro de la lingüística también han utilizado corpus como campo de pruebas para sus teorías o como recurso para extraer datos o materiales. Éste es el caso de la lingüística histórica y la psicolingüística, así como la sociolingüística, la pragmática y el análisis del discurso; la dialectología y los estudios en variación lingüística también han adoptado un enfoque basado en corpus. De la misma manera, la literatura y la estilística han utilizado corpus para producir, sobre todo, estudios comparativos sobre diferentes autores o análisis del idiolecto de un autor o traductor en concreto. La lingüística computacional, y las áreas de traducción automática y procesamiento del lenguaje natural en particular, también han aprovechado las ventajas de la lingüística de corpus y la disponibilidad de corpus electrónicos que, a

cambio, se han beneficiado de los avances en dicha disciplina. Finalmente, los corpus también han encontrado aplicación en la enseñanza de lenguas.

En el *Capítulo 4* se muestra cómo, desafortunadamente, la aplicación de la metodología de corpus a la enseñanza de la traducción no está tan extendida como cabría esperar y su influencia hasta ahora ha sido muy limitada. Entre los distintos tipos de corpus que existen, los más utilizados en esta área son los comparables y paralelos y éstos pueden dar pie a una gran variedad de actividades de aprendizaje, yendo desde la propia compilación de corpus hasta la extracción de información de diverso tipo. Para esto último, en el capítulo se ofrece una recopilación de estrategias de búsqueda en corpus especializados en lengua meta.

Parte II

La *Parte II* de este trabajo es la dedicada a tratar el marco conceptual apropiado para la metodología de la lingüística de corpus en el que se sitúa esta tesis respecto a la didáctica de la traducción escrita. Esta parte comienza con el *Capítulo 5*, donde se presenta un repaso a la evolución de los enfoques en la enseñanza en esta área para más adelante justificar la adopción del diseño de objetivos de aprendizaje, y el marco metodológico del enfoque por tareas de traducción de Hurtado (1992, 1993, etc.). También se expone de forma resumida el trabajo del grupo de investigación PACTE sobre la competencia traductora y su adquisición, poniendo especial énfasis en el componente instrumental dada la naturaleza del uso de corpus. Dado que en nuestra concepción de la competencia traductora priman los conocimientos operativos sobre los declarativos, la práctica totalidad de los objetivos que se proponen en el siguiente capítulo se centran en que el estudiante “aprenda a hacer” y “aprenda haciendo”.

En el *Capítulo 6* se presenta el marco de la propuesta que se realiza en esta tesis. Se ubica la competencia en el uso de corpus para traducir dentro de la competencia traductora; se desglosan los objetivos de aprendizaje generales y específicos tanto de la fase de iniciación en el uso de corpus para traducir como de la fase de consolidación, así como se explicitan los objetivos relacionados con competencias genéricas y los relacionados con materias de traducción; se presentan los bloques de contenidos incluidos en la propuesta y el marco metodológico en que ésta se inscribe; se comentan las materias de traducción en las que se aplica el uso de corpus; se justifican la selección y ubicación de las unidades didácticas diseñadas; se explican los criterios para la elaboración de dichas unidades y se pone de relieve la variedad de las mismas.

También se trata la cuestión de la evaluación dentro de la propuesta, se exponen las herramientas utilizadas para este fin y se ofrecen ejemplos reales de su puesta en práctica. Creemos que, si se incluye el manejo de corpus para traducir como uno de los objetivos de un curso de traducción, también se debe entonces evaluar su adquisición (y no quedarse solamente en una evaluación de la traducción, es decir, del producto final de todo el proceso de documentación, análisis, etc.). Para terminar, se repasa la evolución que ha experimentado la propuesta realizada.

Parte III

La *Parte III* muestra la aplicación concreta de la propuesta de unidades didácticas, es decir, todas las unidades diseñadas, completas y probadas en el aula. Se incluyen aquí tres unidades didácticas de traducción directa y cuatro de inversa, unas para nivel inicial y otras de consolidación. Estas unidades son:

- 1) “Introducción al uso de corpus electrónicos para traducir” (traducción directa, nivel de iniciación)
- 2) “Contrastes de uso en un corpus paralelo” (traducción directa, nivel de iniciación)
- 3) “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos” (traducción directa, nivel de consolidación)
- 4) “Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés” (traducción inversa, nivel de iniciación)
- 5) “El corpus ad hoc como remedio para la traducción médica al inglés” (traducción inversa, nivel de consolidación)
- 6) “Corpus y memoria de traducción” (traducción inversa, nivel de consolidación)
- 7) “Sólo utilizar corpus; utilizar corpus solo” (traducción inversa, nivel de consolidación)

Dado que el objetivo global de la enseñanza de la traducción es desarrollar la competencia traductora en el estudiante y ésta comprende entre sus subcompetencias la instrumental o relativa a la documentación y uso de recursos informáticos, se ha puesto un gran énfasis en trabajar con una amplia variedad de recursos, como tipos de corpus diferentes, herramientas, fases, direcciones, campos de especialidad y tipos de problemas de traducción, etc.

VARIEDAD EN LAS UNIDADES DIDÁCTICAS			
FASE PEDAGÓGICA EN EL USO DE CORPUS	DIRECCIÓN DE LA TRADUCCIÓN	CAMPO DE ESPECIALIDAD	TIPO DE CORPUS UTILIZADOS
Iniciación	Directa Inversa	Economía Política	Corpus generales Corpus especializados (ya construidos y de fabricación propia)
Consolidación		Técnica Científica Médica	

Esta variedad se ha procurado, en parte, porque nos adscribimos a una forma de pensar que no apuesta por la ultra-especialización del alumnado de traducción en ningún campo en particular, si bien es cierto que sí creemos en el trabajo con textos de temática, género y registro variados:

“... esbozamos una concepción de la práctica docente de la traducción científica y técnica que prefiere no recrearse excesivamente en el producto, en definitiva en el “saber” (lo que está bien y lo que está mal), sino fomentar el “saber dónde y cómo encontrar soluciones” y que pretende orientarse justamente hacia conocimientos sin fecha de caducidad en el ejercicio profesional del futuro traductor”. (Vega 1999: 242)



Tras el apartado de *Conclusiones*, donde se expone la consecución de los objetivos de esta tesis y se destacan los resultados principales y posibles perspectivas de la investigación, se incluye un *Glosario* en español e inglés de términos relacionados con los corpus electrónicos elaborado a partir de definiciones ya existentes de varios autores. Finalmente, el *Volumen 1* termina con el *Índice de tablas, documentos y figuras* y el apartado de *Bibliografía* en el que las obras citadas y consultadas se han clasificado en tres grandes bloques, la lingüística de corpus y temas afines, traducción, estudios sobre la traducción, didáctica de lenguas y traducción y competencia traductora, y conferencias no escritas.

En el *Volumen 2* se incluyen todos los materiales utilizados en las siete unidades didácticas elaboradas y puestas en práctica en el marco de esta tesis. Se muestran las fichas de trabajo y de evaluación diseñadas para cada unidad, así como los textos utilizados u otros materiales facilitados a los estudiantes. Se ofrecen las soluciones o ejemplos de soluciones de las fichas de trabajo y se muestran las evidencias más relevantes de la evaluación como, en especial, las respuestas redactadas de los estudiantes a la ficha de autoevaluación y de valoración de cada unidad.

PARTE I.

**LA LINGÜÍSTICA DE CORPUS
Y SUS APLICACIONES**

“... as Knowles (1990) points out, the distinction between teaching and study and research becomes very blurred in the case of corpus-based work, since even at the level of the most obvious kind of corpus use, one is entering into the realms of discovery, possibly novel discovery”
(Fligelstone 1993: 101)

1. LINGÜÍSTICA DE CORPUS

1.1. Definición de “corpus”

La construcción y uso de corpus no es algo exclusivo de nuestra sociedad de la información y las nuevas tecnologías. Desde hace siglos los investigadores en lenguas han recogido textos con el fin de utilizarlos para un propósito concreto. Por ejemplo, Francis (1992) nos recuerda que el uso de corpus en lexicografía se retrotrae al menos hasta principios del siglo XVII, aunque esto no significa que no existan ejemplos anteriores. En un sentido tradicional, un corpus puede ser una colección o cuerpo de textos como las obras completas de un autor en particular o una colección de citas. Hoy en día, sin embargo, el término “corpus” tiende a relacionarse con ciertas características que se reflejan en las definiciones ofrecidas por diversos académicos relevantes al tema. Por ejemplo, el ya fallecido John Sinclair, experimentado lingüista y experto en lingüística de corpus aplicada a la lexicografía, define un corpus como “a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language” (Sinclair 1995: 19). Mona Baker, académica de la traducción y experta en estudios sobre la traducción basados en corpus, lleva más allá la definición de Sinclair al afirmar que un corpus es “any collection of running texts (as opposed to examples/sentences), held in electronic form and analysable automatically or semi-automatically (rather than manually)” (Baker 1995: 226). Según Tony McEnery y Andrew Wilson, profesores ambos en el Departamento de Lingüística y Lengua Inglesa de la Lancaster University, un corpus es “a finite collection of machine-readable text, sampled to be maximally representative of a language or variety” (McEnery y Wilson 1996: 177). Finalmente, para Graeme Kennedy, profesor de lingüística aplicada en la Victoria University of Wellington, un corpus es “a body of written text or transcribed speech which can serve as a basis for linguistic analysis and description” (Kennedy 1998: 1).

Como se puede observar partiendo de las definiciones apuntadas, el término “corpus”, tal y como se entiende hoy en día, reúne las siguientes características:

- (a) Es (o tiende a ser) finito.
- (b) Está en formato electrónico.
- (c) Tiene unos criterios de diseño definidos.
- (d) Es representativo de una lengua o variedad de lengua.

La mayoría de los corpus existentes son finitos en el sentido de que tienen un punto de partida y un punto final tanto en términos de tamaño como de cobertura. Dado el esfuerzo de mantenimiento que requieren, existen muy pocos corpus abiertos que se encuentren en permanente crecimiento. Uno de ellos es el *Corpus de Referencia del Español Actual (CREA)* de la Real Academia Española de la Lengua, y otro se conoce como el *Bank of English*, un corpus vastísimo (en 2008 tiene unos 450 millones de palabras) que pertenece a la editorial HarperCollins y la University of Birmingham. El formato electrónico es una característica de los corpus modernos que ha ayudado de manera significativa a que la lingüística de corpus se desarrollara y extendiera como metodología. En cuanto a criterios de diseño se refiere, un corpus diseñado para el análisis lingüístico se caracteriza por la compilación sistemática, planificada y estructurada de sus textos. Este aspecto permite distinguir a los corpus de otros recursos lingüísticos como los archivos, que normalmente son repositorios de textos, o las bases de datos, que son colecciones de información diseñadas para facilitar la introducción y recuperación de datos. En resumen, los corpus tienen el rasgo distintivo de estar diseñados y construidos con el objetivo de ser representativos del uso “real” dentro de una determinada lengua o muestra de lengua en un momento determinado.

1.2. Definición de “lingüística de corpus”

Si “corpus” se define tal y como hemos visto, la “lingüística de corpus” será entonces el estudio de la lengua a través de corpus y, aunque el término ya se hubiera utilizado antes de forma esporádica, no fue hasta 1984 cuando apareció “corpus linguistics” titulado un libro (Aarts y Meijs 1984). En otro orden de cosas, no está tan claro que la lingüística de corpus sea una rama de la lingüística ya que no puede considerarse como tal en el mismo plano que la morfología, la sintaxis o la semántica, por ejemplo. Tratamos más bien con una metodología dentro de la lingüística que utiliza determinadas técnicas y herramientas. En este sentido, la lingüística de corpus estaría al mismo nivel que otros métodos dentro de la lingüística, como la observación, la introspección, la elicitación o la experimentación. La siguiente cita de Laviosa aclara la situación totalmente:

“What characterises corpus linguistics as an independent discipline within general linguistics is not just its specific methodology nor the particular nature of its object of study, but its unique approach to the study of language which is firmly based on the integration of four interdependent, equally important elements: data, description, theory and methodology” (Laviosa 2002: 8).

La misma autora (ídem: 8-9) adapta lo aportado por Stubbs (1993: 2; 1996: 23) para ofrecernos los principios más importantes sobre los que se sustenta la aproximación al estudio de la lengua desarrollado por la lingüística de corpus, que aquí ofrecemos de forma sintética:

- rejection of the Saussurian *langue-parole*, the Chomskian competence-performance and internalised-externalised language dualisms
- language is viewed as a social phenomenon
- language in use involves both routine and creative processes, individuality and generality
- language in use is systematically heterogeneous
- language is studied empirically in large collections of authentic [...] spoken and written texts
- the meaning of words is enshrined in language use and studied in actual corpus citations
- the role of linguistics is to discover, understand and describe the “genuine” rules of language
- linguistics is essentially a social and an applied science

Principios más importantes en los que se basa la aproximación al estudio de la lengua desarrollado por la lingüística de corpus (Laviosa 2002: 8-9).

1.3. Ámbito y evolución de la lingüística de corpus

El principal interés de la lingüística de corpus es el de proporcionar pruebas empíricas para procurar unas mejores descripciones de la lengua a partir de datos y así explicar, probar o ejemplificar teorías ya existentes o asunciones respecto a la lengua. Aún más, la lingüística de corpus es un medio de descubrir patrones y revelar hechos sobre la lengua que quizás antes ni siquiera hubiéramos pensado buscar. En este punto, parece apropiado recordar un comentario hecho por el académico de la traducción Ian Mason¹. Hablando sobre investigación en los estudios sobre la traducción, Mason quiso remarcar la diferencia entre la investigación que nos permite afirmar que “un fenómeno ocurrió”, “un fenómeno ocurre” o “un fenómeno ocurre frecuentemente, típicamente o siempre”. Gracias a la lingüística de corpus parece seguro afirmar que ahora estamos más cerca que nunca de poder sostener que un hecho ocurre frecuentemente, típicamente o siempre, respecto de una lengua sin miedo a confiar exclusivamente en nuestra intuición, siempre que se encuentre la manera de formular la pregunta. Sin embargo, también hay que reconocer que en los estudios llevados a cabo con esta metodología se da la tendencia a centrarse demasiado en lo léxico, lo cual

¹ Comentario recogido en Inglaterra, en la conferencia de UMIST del 2000 titulada “Research Models in Translation Studies”.

se debe a que el software disponible en la actualidad para el manejo de corpus está diseñado principalmente para analizar elementos léxicos dentro de contextos limitados.

Sin embargo, la lingüística de corpus ya existía mucho antes que los ordenadores. Como mencionamos al principio de este apartado, hay ejemplos del uso de corpus ya en el siglo XVII y en el XVIII, como es el caso del diccionario de Samuel Johnson en el que cada entrada estaba acompañada por una cita de un corpus de obras literarias, o el *Old English Dictionary* del siglo XIX. Durante el siglo XIX y la primera mitad del siglo XX varias áreas de estudio hicieron uso de esta metodología. Un ejemplo tal es la investigación en adquisición del lenguaje que, en aquel tiempo, estaba basada en diarios cuidadosamente escritos por padres que recogían las locuciones de sus hijos. Käding (1897) estudió las convenciones ortográficas utilizando un corpus de alemán de 11 millones de palabras y Estoup (1916) analizó un corpus para determinar la frecuencia de grafemas para la taquigrafía. Acercándonos a la mitad del siglo XX, los estudios basados en corpus en la pedagogía de la lengua estaban centrados sobre todo en la extracción de listas del vocabulario más frecuente para su uso en la enseñanza de lenguas extranjeras. Otra disciplina, la lingüística contrastiva, también se aprovechó de los estudios basados en corpus y así Eaton (1940) llevó a cabo un sofisticado estudio comparativo de la frecuencia de significados en diferentes lenguas. Y, finalmente, investigadores como Fries (1952) y Gougenheim *et al.* (1956) llevaron a cabo estudios basados en corpus con el fin de describir el uso de la gramática en inglés y francés respectivamente.

A pesar de estos datos, no puede obviarse el hecho de que los avances tanto en software como en hardware han tenido un gran impacto en la lingüística de corpus en términos de velocidad de procesamiento y capacidad de almacenamiento de datos. Hoy en día los ordenadores pueden manejar enormes cantidades de datos en cuestión de segundos con exactitud y fiabilidad estadística, algo que ha aportado a esta metodología unas bases sólidas sobre las que fundarse.

1.3.1. Críticas desde la lingüística generativa

Como podemos comprobar, los corpus ya se construían y utilizaban para diversos propósitos lingüísticos incluso, como es el caso de alguno de los estudios mencionados, en un tiempo en el que ni siquiera el término “lingüística de corpus” se había acuñado aún. Sin embargo, este periodo de prosperidad y optimismo que disfrutó la investigación basada en corpus en la primera mitad del siglo XX no duraría mucho. Un

periodo productivo de investigación con corpus fue seguido por otro de casi total abandono y falta de popularidad, periodo que empezó en los años 50 con la llegada de Noam Chomsky y el estructuralismo. En una serie de publicaciones de gran influencia, Chomsky (1957, 1962) cambió la dirección de la lingüística alejándola del empirismo (basado en la observación de datos que ocurren naturalmente, como los contenidos en un corpus) y llevándola hacia el racionalismo (basado en la introspección). Su interés principal era la definición de un modelo de la competencia lingüística, entendida como el conocimiento interiorizado de una lengua. Chomsky consideraba la actuación, la prueba externa de la competencia lingüística, como un “reflejo pobre” de la misma al estar influenciada por factores ajenos a la propia competencia. La investigación basada en corpus, centrada en el estudio del uso “real” de la lengua lo cual podría implicar la inclusión de muchos errores gramaticales en un corpus, simplemente carecía de interés para Chomsky. Sumado a esta falta de intereses comunes, algunas convicciones erróneas mantenidas por los primeros lingüistas de corpus también alentaron las críticas de Chomsky hacia la metodología. Se asumía, en aquel tiempo, que la lengua era finita y que todos los ejemplos (frases) podían recogerse y enumerarse, de lo que se desprendía que la observación y descripción “completa” de una lengua debía de ser posible a través de corpus. En contra de este parecer, Chomsky sostenía la opinión de que la lengua era infinita e innumerable, por lo que los corpus eran incompletos, parciales y sesgados, llevándole a concluir que nunca podrían ser el único medio de estudiar y explicar la lengua. Además, Chomsky también criticaba las “pseudo-técnicas” utilizadas para el análisis de corpus en aquel tiempo y que básicamente descansaban en el procesamiento manual y el manejo de los datos por parte del investigador quien, como ser humano, era propenso a cometer errores. El resultado final de este argumento fue un rechazo general de las metodologías empíricas basadas en corpus en la lingüística.

Sin embargo, a pesar de las críticas, es menester decir que no todo el trabajo con corpus se detuvo en este punto. Había campos de estudio como la fonética y la adquisición de lenguas que continuaron utilizando corpus como fuente de pruebas por dos razones principales: su estudio requería datos que ocurrieran naturalmente, y la introspección era insuficiente o ni siquiera era una opción, como en el caso de niños que aprendían una lengua. En otras áreas, otros investigadores también continuaron construyendo y trabajando con corpus cuando se hizo evidente que las aproximaciones racionalistas al estudio de la lingüística también tenían sus limitaciones mientras que los corpus aportaban diversas ventajas, algunas de las cuales listamos a continuación:

- Las observaciones basadas en corpus son intrínsecamente más verificables que los juicios basados en la introspección.
- Los datos basados en frecuencias no pueden recuperarse con tanta exactitud a través de la introspección.
- Un corpus proporciona la base para una aproximación mucho más sistemática al análisis de la lengua en términos de verificación objetiva de resultados.
- Las oraciones no gramaticales no son tan numerosas como se suponía en un principio y, en cualquier caso, pueden ser incluso interesantes para determinados tipos de estudio.

Algunas ventajas del uso de corpus en el estudio de la lengua (basado en McEnery y Wilson 1996).

Al final, algunas de las primeras críticas de la metodología de corpus resultaron ciertas. Los académicos chomskianos, por ejemplo, estaban acertados cuando afirmaban que la lengua era infinita y que la actuación se veía afectada por factores ajenos a la competencia lingüística. Otras críticas resultaron ser incorrectas o parcialmente incorrectas. Las críticas a los datos cuantitativos como datos sin importancia o utilidad, o la afirmación radical de que sólo la introspección o sólo la investigación basada en corpus podrían ser la única explicación de la lingüística perdieron su fuerza gradualmente. Finalmente, el argumento chomskiano contra las técnicas manuales (pseudo-técnicas) utilizadas por los primeros lingüistas de corpus se tornó inválido con los avances de la tecnología y los ordenadores.

La reconsideración de los argumentos mencionados llevó a una re-adopción del estudio de la lengua basado en corpus como una metodología en lingüística. Este cambio empezó en 1980 y se debió a varias razones tanto prácticas como teóricas:

- La disponibilidad de material para corpus en formato electrónico.
- La aceptación de que un corpus solo nunca podría explicar toda la lingüística.
- La creciente concienciación de la utilidad de los datos cuantitativos.
- El entendimiento de que el corpus y la intuición del lingüista eran complementarios y no excluyentes.

En conclusión, estos avances teóricos y prácticos han ayudado a la lingüística de corpus y los corpus mismos a desarrollarse a través de los años. En su algo fluctuante historia, los corpus han sido utilizados para propósitos y estudios lingüísticos variados. Desde el momento en el que el término “lingüística de corpus” ni siquiera existía y el análisis de la lengua se llevaba a cabo manualmente, hasta el día de hoy, cuando los investigadores trabajan con corpus ingentes y software sofisticado, los corpus se han utilizado en muchos y diversos campos de investigación que veremos en detalle más adelante (cfr. infra 3.). Insistiremos aquí, sin embargo, que todos ellos tienen algo en común: el uso

de corpus como recursos que contienen lengua extraída de la “vida real”, con el objetivo de buscar pruebas empíricas que apoyen o refuten teorías ya existentes o asunciones sobre una lengua. Además, también merece la pena destacar que, como se ha sugerido previamente, en ocasiones los corpus han facilitado al investigador datos que éste no esperaba o ni siquiera buscaba. Como afirma Fillmore (1992: 35), “(...) every corpus I have had the chance to examine, however small, has taught me facts I couldn’t imagine finding out any other way.”

1.4. Fases del trabajo con corpus

Ahora que ya hemos establecido lo que es un corpus y comentado algunos aspectos de la lingüística de corpus como metodología, los temas de debate que ha suscitado y algunas de sus aplicaciones, podemos pasar a describir en detalle las diferentes fases que implica un estudio realizado con corpus, siempre con la mirada puesta en la aplicación ulterior que le hemos dado a los corpus dentro de este trabajo, la formación de traductores. La importancia del seguimiento de determinados pasos o fases en el trabajo con corpus viene reforzada por citas como la de Corpus Pastor y Seguiri Domínguez.

“[...] if the collection of texts is to qualify as a “corpus” and be considered as representative of a particular field, it is essential that it conforms to clear design parameters that are set out from the beginning followed by a specific compilation protocol. This protocol is divided into four distinct phases: (a) locating and accessing resources; (b) downloading data; (c) text formatting; and (d) data storage (Corpus Pastor y Seguiri Domínguez 2008).

Así, en un estudio con corpus, y tal y como nosotros lo hemos estructurado, una vez decidido qué se desea investigar, explicitadas las hipótesis y valorado el hecho de si es éste el método más adecuado para ello, nos encontramos con la fase de selección de los textos que se incluirán en el corpus, o la selección del corpus en sí y los criterios implicados en estas decisiones. Tras esto está la preparación de los textos, fase a la que sigue el análisis del corpus con software apropiado y, por último, la extracción de resultados y conclusiones.

1.4.1. Establecimiento del objeto de estudio e hipótesis

El primer paso en cualquier estudio de corpus debería ser siempre el establecimiento del objetivo de dicho estudio y el consecuente planteamiento de las hipótesis. El hecho de que en estos días los estudios de corpus estén en un momento de gran popularidad no debe hacer perder al investigador la perspectiva de lo que quiere estudiar y la imparcialidad a la hora de elegir aplicar una u otra metodología de investigación. En cuanto a las hipótesis, éstas no dejan de ser preguntas que el investigador desea responder y, en el caso del trabajo con corpus, la formulación de las preguntas es clave ya que va estrechamente ligada a la manera como se interrogará el corpus. Para poner un ejemplo, alguien podría estar interesado en saber si hombres y mujeres escriben de forma diferente, pero para responder a esta pregunta abierta debería ponerse en clave de “caracteres” que se pudieran introducir en un programa de análisis de corpus. Esto nos llevaría a pensar en cómo esas diferencias pueden manifestarse en la superficie del texto, es decir, a plantear nuestras hipótesis respecto a este estudio. Una de las hipótesis posibles que ayudaría a buscar marcadores en la superficie del texto sería pensar que hombres y mujeres utilizan de forma diferente las palabras malsonantes, por ejemplo.

En el tema que nos ocupa en este trabajo, la formación de traductores, también hay objeto de estudio e hipótesis. En el anterior trabajo de investigación con corpus de la autora (Rodríguez Inés 1999, 2000) había un solo gran objeto de estudio, la ideología en traducción, y unas hipótesis, que hay marcas ideológicas en la lengua que se pueden observar estudiando sólo textos traducidos y que esta observación se puede hacer con un corpus comparable bilingüe documentado. En esta tesis, sin embargo, se presenta más adelante (cfr. *infra Parte III.*) una amplia variedad de objetos de estudio e hipótesis a pequeña escala que darán lugar a pequeños estudios de corpus llevados a cabo bien por la docente bien por los estudiantes en su proceso de formación como traductores. Para poner un ejemplo, en una tarea se puede tener como objeto de estudio la intuición de un nativo sobre su lengua. Una posible manifestación de este hecho en la superficie de los textos serían aquellos elementos que un nativo “sabe” por experiencia o intuición que son correctos o incorrectos, frecuentes o infrecuentes; los estudiantes con el español como lengua materna podrían tener la intuición, por ejemplo, de que utilizar la conjunción “pero” a principio de frase no es correcto. Una comprobación en un corpus general extenso de la lengua española les ratificaría que este uso está perfectamente extendido y aceptado.

1.4.2. Selección de textos/corpus

Una vez decidido qué se quiere investigar y forjada una idea de cómo se puede llevar a cabo tal investigación, el siguiente paso debería ser la selección cuidada de los textos que formarán el corpus con el que vamos a trabajar. Tanto si decidimos utilizar un corpus ya construido como si preferimos construir el nuestro propio, hay varios aspectos que deben tenerse en cuenta ya que tendrán una influencia directa sobre el estudio que queremos efectuar. Estos aspectos se describen en los apartados que siguen, pero si el caso en el que nos encontramos es el de la construcción de un corpus, Torruella y Llisterra (1999: 59 y ss) listan y comentan a un nivel bastante elaborado los aspectos principales que han de tenerse en cuenta en el diseño de un corpus, sea éste del tipo que sea.

Aspectos principales	Comentarios/Recomendaciones
Finalidad	Cuestión que condicionará todo lo demás.
Límites del corpus	Fechas de los documentos, lenguas incluidas y variantes geográficas.
Tipo de corpus	Es necesario definir: <ul style="list-style-type: none"> a) porcentaje y distribución de los diferentes tipos de textos, b) especificidad de los textos, c) cantidad de texto que se toma de cada documento, d) codificación y anotaciones que se le añaden, e) documentación que le acompaña.
Proporciones de los diferentes grupos temáticos del corpus	Depende de otros aspectos.
Población y muestra	La relación entre estos aspectos es muy importante porque se pretende que de la observación de las muestras se puedan extraer generalizaciones sobre la lengua.
Número y longitud de los textos de la muestra	La selección se puede hacer de tres formas: <ul style="list-style-type: none"> a) al azar, b) dividiendo los textos en tres partes de extensión parecida y extrayendo de cada una de ellas las muestras en número y proporciones aproximadamente iguales, c) determinando la estructura externa de los textos y decidiendo qué niveles estructurales se usarán para el

	muestreo.
Captura de los textos y etiquetado	Utilización de escáneres, OCR y programas de etiquetado.
Procesamiento del corpus	Tan importante es el corpus como las herramientas.
Crecimiento del corpus y <i>feedback</i>	Es necesario colaborar conjuntamente tanto con expertos estadísticos como con los usuarios del corpus.
Hardware y software	Es necesario hacer una estimación de la infraestructura informática necesaria para desarrollar y explotar el corpus.
Aspectos legales	Cuestiones de propiedad intelectual, para las que en ocasiones la legislación no ofrece soluciones claras.
Presupuesto y etapas	Es necesario considerar cuestiones como el personal humano, software, hardware y adquisición de derechos de autor, si cabe.

Tabla 1. Aspectos principales del diseño de corpus (adaptado de Torruella y Llisterri 1999: 59 y ss).

Es menester añadir que, dependiendo del objetivo para el que se construye un corpus de un determinado tipo, las fases de selección y recogida de textos pueden ser más o menos flexibles. En otras palabras, si, por ejemplo, lo que se plantea es la construcción de un corpus reducido en tamaño y alcance para un propósito muy concreto e incluso limitado a muy pocos usuarios, como es el caso de los corpus ad hoc o desechables que se verán más adelante (cfr. infra 4.2.4.), entonces la selección de textos puede ser recursiva en el sentido de que el usuario puede seguir alimentando su corpus con textos a medida que le surgen necesidades de documentación. Casi con toda probabilidad estos corpus construidos generalmente con escasos recursos y tiempo no serán equilibrados; quizá no serán representativos y la selección de textos se basará seguramente en la disponibilidad de los mismos. Sin embargo, la experiencia propia y ajena nos dice que la utilidad de estos corpus construidos de forma ad hoc, incluyendo el proceso de selección de los propios textos, proporciona grandes beneficios que se irán desgranando en estas páginas.

1.4.2.1. Tipos de corpus

Dada la relevancia del corpus en un trabajo basado en este recurso, es importante que aquél que utilice o construya uno sepa qué tipos de corpus existen, razón por la cual hemos incluido una tipología. La que presentamos aquí es la tomada de Corpas (2001: 157 y ss), que a su vez adopta la tipología de Torruella y Llisterri (1999) y le añade modificaciones basadas en otras tipologías. En definitiva, Corpas presenta diversas

tipologías, dependiendo de los criterios clasificatorios, y ofrece definiciones, que también están recogidas en el *Glosario* de esta tesis (cfr. infra *Volumen 2*):

1.4.2.1.1. Según el porcentaje y la distribución de los diferentes tipos de texto

- Corpus grande o extenso:
Corpus cuya extensión no está delimitada voluntariamente y que recoge, típicamente, un número de palabras muy elevado, sin tener en cuenta aspectos como el equilibrio o la representatividad.
- Corpus equilibrado:
Corpus que contiene diversas variedades de lengua en porcentajes similares.
- Corpus piramidal:
Corpus que contiene textos distribuidos por niveles, los cuales se caracterizan por aumentar progresivamente la complejidad de las variedades temáticas, en detrimento del número de textos incluidos en cada variedad.
- Corpus monitor:
Corpus que mantiene un volumen textual constante, pero actualizado, al ir añadiendo elementos nuevos a la vez que elimina cantidades equivalentes de material antiguo.
- Corpus paralelo:
Corpus formado por una serie de textos en la lengua de origen junto con sus traducciones en una (o varias) lengua(s) meta. Se habla de *corpus paralelo bilingüe* cuando hay sólo dos lenguas implicadas, y de *corpus paralelo multilingüe* cuando se trata de más de dos.
- Corpus comparable:
Corpus que, en relación a otro u otros corpus de lenguas distintas, incluyen tipos similares de textos originales. (...) Se habla de *corpus comparable bilingüe* cuando se incluyen dos lenguas, y *corpus comparable multilingüe*, cuando hay más de dos.

1.4.2.1.2. Según la especificidad de los documentos

- Corpus general:
Corpus representativo de la lengua común, que incluye una gran variedad de textos producidos en situaciones comunicativas cotidianas.
- Corpus especializado:
Corpus representativo de un tipo particular de lengua, que incluye textos pertenecientes exclusivamente a tal variedad.
- Corpus genérico:
Corpus representativo de un determinado género, que incluye textos pertenecientes a dicho género.

- Corpus canónico:
Corpus que contiene las obras completas de un determinado autor.
- Corpus periódico o cronológico:
Corpus que incluye textos producidos durante un determinado periodo de tiempo.
- Corpus diacrónico:
Corpus que incluye textos pertenecientes a etapas sucesivas en el tiempo con objeto de estudiar la evolución lingüística.

1.4.2.1.3. Según la cantidad de texto en cada uno de los documentos

- Corpus textual:
Corpus que incluye textos completos con objeto de ofrecer una muestra representativa de la lengua común, así como de sus variedades más importantes, de manera que pueda ser utilizado como base para la elaboración de gramáticas, diccionarios y otras obras de referencia.
- Corpus de referencia:
Corpus que no incorpora documentos enteros, sino fragmentados, ya que el interés no reside en el texto en sí, sino en el estadio de lengua representado.
- Corpus léxico:
Corpus que recoge documentos formados por fragmentos muy pequeños de texto de igual longitud.

1.4.2.1.4. Según la codificación y la anotación

- Corpus no anotado:
Corpus en formato de sólo texto o ASCII, que no ha sido etiquetado de forma alguna y, por tanto, presenta un alto grado de simplicidad.
- Corpus anotado:
Corpus cuyos documentos han sido etiquetados lingüística o metatextualmente, ya sea de forma manual o de forma automática.

1.4.2.1.5. Según la documentación que acompaña a los textos

- Corpus documentado:
Corpus en el que cada documento textual lleva asociado un archivo DTD (Document Type Definition) o una cabecera (header) descriptiva de su procedencia.
- Corpus no documentado:
Corpus cuyos documentos integrantes carecen de tales archivos asociados o cabeceras.

(Corpas 2001: 157 y ss)

En lo que respecta a la formación de traductores, y como se explicará más adelante en detalle (cfr. infra 4.2.), los tipos de corpus más relevantes son el paralelo bilingüe y el comparable tanto monolingüe en lengua meta como bilingüe, aunque si atendemos a otro criterio de clasificación entrarían también los corpus generales, especializados y genéricos, todos ellos utilizados en el presente trabajo.

1.4.2.2. *Procedimientos de selección*

Los criterios y procedimientos de selección de textos son de vital importancia para aquellos que compilen corpus ya que tendrán una influencia directa sobre la utilidad del propio corpus. Sin embargo, a pesar de que un corpus ha de estar diseñado de acuerdo con unos principios bien definidos para ser lingüísticamente representativo, se debería reconocer también que, en muchos casos, los textos se seleccionan al azar o en base a su disponibilidad, hecho que restringirá el tipo de investigaciones que puedan llevarse a cabo.

No obstante, también existen importantes ejemplos de corpus que se han construido siguiendo criterios de selección bien establecidos con el fin de conferirle la mayor representatividad y exhaustividad posibles. Un ejemplo de ello es el *British National Corpus (BNC)*, un corpus general representativo del uso del inglés actual. La selección de textos para el *BNC* está basada en datos de producción y recepción con el fin de obtener una descripción bastante cercana a la lengua que los hablantes actuales de inglés producen y perciben. Para tal fin, se siguieron diversos procedimientos a la hora de seleccionar los textos. Para el componente escrito del corpus se seleccionaron algunos títulos al azar de entre candidatos disponibles identificados en el *Whitaker's Books in Print (1992)*, mientras que otros se tomaron de las listas de los libros más vendidos o listas de premios literarios. También se tomaron en cuenta los préstamos de las bibliotecas, así como las publicaciones periódicas, revistas y diarios, y otros textos impresos como folletos publicitarios, ensayos académicos o material escrito de programas de noticias para la televisión. Los textos también fueron escogidos según algunos rasgos de selección específicos, tales como la factualidad (imaginativo e informativo, en términos generales), un periodo de tiempo (aproximadamente de 1960 en adelante) y el formato (libro, publicación seriada, etc.). En cuanto al componente oral del corpus, se tomaron muestras y se seleccionaron textos en base a cuestiones demográficas, tomando también en cuenta otros aspectos como la región, el nivel de alfabetización, el género de los hablantes y el tema.

El *BNC* ha establecido un estándar en la compilación de corpus en cuanto a la definición rigurosa de procedimientos de selección y criterios observables para procurar la cobertura más amplia posible del panorama actual de la lengua inglesa. Obviamente, la exhaustividad en términos de alcance es, en la mayoría de los casos, imposible de conseguir en la compilación de un corpus, excepto en el caso de que se trate con una población de textos limitada y cerrada, como las obras del periodo del inglés antiguo, por ejemplo, o toda la obra de un autor en particular. A pesar de la falta de exhaustividad total, si los criterios de selección seguidos para la compilación del corpus se explicitan, esto debería proveer a la investigación de una base firme y la justificación requerida ante las críticas.

En la enseñanza de la traducción la selección de corpus y, en especial, la selección de textos para su inclusión en un corpus es de una importancia tal que incluso supone una actividad de aprendizaje en sí misma. Por otro lado, los condicionamientos prácticos imponen en ocasiones que en la selección de textos se priorice la disponibilidad de los mismos, sin desatender por ello la calidad.

1.4.2.3. *Cuestiones de almacenamiento*

Una vez que se han seleccionado los textos que se incluirán en el corpus deberían almacenarse y catalogarse siguiendo un método de referenciación sistemático para facilitar su acceso más adelante. Teniendo esto presente, Kennedy (1998) apunta que también se debería realizar un registro lo más detallado posible de los rasgos extratextuales tales como la autoría, la fuente de los textos o cualquier otro dato relevante sobre el texto en cuestión.

Llevando el tema más allá, Sinclair llama la atención sobre la necesidad de mantener la integridad del texto tan intacta como sea posible. En este sentido, ha sugerido que sería conveniente preservar el texto original y recomienda hacer tanto una copia digitalizada como una copia de trabajo de cada texto, ésta última sin ningún tipo de codificación. También subraya la importancia de mantener la legibilidad y sugiere asignar a cada texto un número de identificación. El objetivo de Sinclair con estas recomendaciones no es otro que el de evitar la corrupción de los textos en el proceso de preparación para su inclusión en un corpus (Sinclair 1999).

El almacenamiento y referenciación de los textos son temas también relevantes al uso de corpus para la enseñanza de la traducción, especialmente cuando se construye un

corpus en el aula. Si el trabajo es además colaborativo, también aumenta la importancia de contar con un método para nombrar cada archivo y con unidades de disco compartidas de bastante capacidad.

1.4.2.4. *Derechos de autor*

Uno de los primeros problemas que se deben resolver en aquellos casos en los que se quiera compilar un corpus para objetivos propios es el del *copyright* o derechos de autor. Como cualquier otro usuario de textos, los compiladores de corpus deben observar ciertos requisitos legales antes de copiar un texto en un corpus para su explotación. Este hecho no sólo implica la obtención del permiso sobre los derechos de autor, sino también el consentimiento de los individuos cuyo derecho a la privacidad debe reconocerse y respetarse. En la mayoría de los casos, la selección de textos se planifica cuidadosamente antes de empezar con los trámites relativos a los derechos de autor. Sin embargo, expertos en el campo como Tim Johns² (2000) sugieren que el permiso debería obtenerse antes de que el diseño y la estructura del corpus estén finalizados o, al menos, ambas tareas deberían llevarse a cabo simultáneamente. Según Johns, se podría ahorrar mucho tiempo en el diseño y compilación de un corpus si el compilador supiera por adelantado qué textos tendrá a su alcance.

Aquél que compile un corpus también debería ser consciente de las dificultades que representa obtener este tipo de permiso ya que no siempre está claro a quién pertenecen los derechos de autor de una obra en concreto. En ocasiones puede darse que la editorial haya sido absorbida por otra compañía, puede ser difícil dar con el paradero de los autores de los artículos, puede que una obra que ya no está sujeta a los derechos de autor en una edición sí lo esté en otra, etc. (Kennedy 1998). En el caso de las traducciones, la cuestión podría ser incluso más complicada ya que también se debería obtener el permiso del traductor, cuyo nombre no siempre se encuentra explicitado. Se ha constatado que las editoriales de mayor envergadura son a veces las menos propensas a cooperar aunque, por otro lado, también hay casos en los que se ha concedido el permiso para fines investigadores no comerciales o, rara vez, para la realización de una tesis.

En el caso del uso de material para fines docentes no es necesario solicitar el permiso de los derechos de autor, especialmente porque los corpus no se “dan” a los alumnos o

² El experto en corpus, Tim Johns, de la University of Birmingham, hizo esta sugerencia durante la IV conferencia TALC, celebrada en Graz, Austria, en julio de 2000.

incluso se “eliminan” después de haber trabajado con ellos. En cualquier caso, la mayoría de los textos y recursos que se utilizan en la enseñanza de la traducción son aquellos recogidos o diseñados para este mismo propósito, se encuentran disponibles en Internet o tienen el permiso de uso para fines académicos. La disponibilidad y la facilidad de acceso en este caso también han de primar si se desea facilitar tanto la labor del docente como la del discente a la hora de trabajar con determinados materiales.

Como resumen a las distintas posiciones que se pueden encontrar, Wilkinson (2006b) ofrece una visión completa de la situación referente a los derechos de autor en cuanto al trabajo con corpus. Según las observaciones del autor, se dan en este tema varias actitudes por parte del compilador de corpus, que van desde la tranquilidad de que si la finalidad es educativa, investigadora o no comercial, se está a salvo, hasta la precaución absoluta de que es mejor solicitar el permiso de uso en cualquier caso. Las leyes aplicables al uso de corpus son, en todo caso, las del país desde donde se da acceso al corpus y no el origen de los textos que lo forman. Así, en España, en lo que respecta al uso de fragmentos de obras ajenas, la SGAE habla del “derecho de cita”, al que podemos acogernos en lo que respecta a ofrecer concordancias extraídas de un corpus, como ocurre en esta tesis.

“[El derecho de cita] constituye una limitación establecida legalmente a los derechos que corresponden a los creadores intelectuales. Permite la inclusión en una obra propia de fragmentos de otras ajenas de naturaleza escrita [...], siempre que se trate de obras ya divulgadas y su inclusión se realice a título de cita o para su análisis, comentario o juicio crítico. La utilización del derecho de cita tiene como límites: ser para fines docentes o de investigación; estar justificado por el fin de la incorporación a una obra, y la necesidad de indicar la fuente y el nombre del autor de la misma” (Sociedad General de Autores y Editores 2002).

Por otro lado, en cuanto a la descarga por parte del docente y los alumnos de documentos no protegidos desde Internet a una unidad local para fines educativos y uso dentro o fuera del aula, pero no comercial, nos asiste el hecho de que no existe diferencia entre esta actividad y descargar cualquier otro archivo no protegido desde la Internet pública siempre y cuando, insistimos, sea para uso personal, didáctico y no comercial.

1.4.2.5. *Captura de textos*

Una vez obtenido el permiso sobre los derechos de autor, si es el caso, los textos están listos para incluirse, en formato electrónico, en un corpus. En cuanto a los textos escritos en papel, existen básicamente dos formas de conversión a formato electrónico: el teclado y el escaneado; ambas requieren el empleo de bastante tiempo ya que es necesario comprobar y editar los textos. Por último, y más problemática incluso, es la compilación de textos de origen oral ya que se deben transcribir las grabaciones para poder así incluirlas en un corpus.

En la formación de traductores, y sin ignorar la valía de las obras de consulta y textos que se puedan encontrar en papel, es frecuente que se prime la disponibilidad de textos en formato electrónico a la hora de crear o trabajar con un corpus ya que la conversión de los textos de papel a formato digital demoraría el proceso, si bien es cierto que consideramos como un objetivo de aprendizaje realizar dicho proceso.

1.4.2.6. *Tamaño*

En la fase de diseño se deben tomar decisiones sobre cuántas palabras, categorías, muestras, palabras en cada muestra, etc. van a conformar el corpus. Además, tales decisiones deben basarse en principios explícitos y bien establecidos.

El tamaño medio de los primeros corpus electrónicos que aparecieron en los años 60 era de un millón de palabras, una cifra enorme si pensamos en la tecnología que estaba disponible en aquel tiempo. Sin embargo, el hecho de que cada vez haya mayor disponibilidad de textos electrónicos, sumado al aumento de memoria y capacidad de disco de los ordenadores, ha tenido un gran impacto en el tamaño actual de los corpus. Hasta tal punto es cierto este hecho que, hoy en día, existen corpus de varios millones e incluso cientos de millones de palabras. Esto ha sido especialmente beneficioso para algunas observaciones y estudios que necesitan de grandes muestras, como por ejemplo la descripción de fenómenos de baja frecuencia. Por otro lado, la disponibilidad de grandes corpus no implica que otros corpus pequeños, bien equilibrados y de alta calidad no puedan servir a objetivos concretos. Patton parece encontrar el equilibrio en este tema: “The sample should be large enough to be credible given the purpose of the evaluation, but small enough to permit adequate depth and detail for each case or unit in the sample” (Patton 1987: 58-59).

Los corpus utilizados en la enseñanza de la traducción pueden ser de todo tipo, tanto grandes como pequeños, si bien es cierto que el alumno debe aprender que más grande no siempre significa mejor, aunque sí necesitará corpus de gran tamaño para resolver algunas cuestiones de traducción. Así, sin poder fijar una cantidad de palabras como una cifra óptima, los corpus que se utilicen para enseñar a traducir serán tan grandes como útiles sean para las necesidades documentales y los objetivos de aprendizaje.

1.4.2.7. *Representatividad*

Según Sinclair, uno de los aspectos más importantes y distintivos de los corpus es que están contruidos de acuerdo con criterios de diseño precisos para un objetivo específico. Esto es así para asegurarse de que el corpus es representativo de una lengua o variedad que pretende explicar. Sin embargo, la representatividad en los corpus es siempre una cuestión controvertida. Los criterios de diseño implican cuestiones teóricas como la selección del campo temático que se va a investigar, la tipología textual, el modo escrito u oral o la decisión de si incluir muestras o textos completos. También hay cuestiones pragmáticas que se deben tener en cuenta, tales como la financiación o la disponibilidad de los textos, siendo conscientes de las posibles dificultades a la hora de tener dichos textos en formato electrónico (p.ej. lenguaje oral) o de obtener el *copyright*, o que podría ser problemático encontrar ejemplos de todas las categorías textuales deseadas. En suma, todos estos aspectos imponen restricciones en el diseño de corpus y es posible que afecten a la representatividad de un corpus³.

Deberíamos subrayar que en la mayoría de los casos estaremos tratando con una muestra de una población más amplia. De aquí se deduce que se debería proporcionar una clara definición de los límites de la población que se pretende estudiar, así como una definición de nuestro marco de muestreo, es decir, la población completa de textos de la cual tomaremos algunos. El tamaño de la muestra en términos de la extensión de cada texto y el número de textos que se incluirán también tiene que determinarse previamente. Por otro lado, aunque la representatividad está, hasta cierto punto, relacionada con el tamaño del corpus, esto no necesariamente se aplica en todos los casos si las proporciones son exactas y los criterios para la inclusión o exclusión de sujetos, textos o partes de textos están bien definidos y justificados de forma razonable.

³ Para ampliar la información sobre el tema de la representatividad en los corpus, léase Halverson, H. (1998) "Translation Studies and representative corpora: establishing links between translation corpora, theoretical/descriptive categories and a conception of the object of study", *Meta*, 43(4), 494-514.

Después de todo, la representatividad no puede juzgarse sólo en términos de cantidad sino también de calidad.

En este sentido, es de esperar que cualquier corpus utilizado para traducir también se rija por los mismos principios de representatividad, si bien es cierto que el corpus más bien se habrá de valorar en tanto en cuanto sea útil para alcanzar los objetivos de aprendizaje fijados. Por otro lado, siendo la representatividad un concepto relativo y flexible, también se da la posibilidad de alterar la composición de un corpus durante el proceso de análisis si éste no sirve a los propósitos documentales o de aprendizaje previstos.

Corpas Pastor y Seguiri Domínguez (2008) han estudiado la cuestión de la representatividad en los corpus para llegar a una solución que se ha materializado en el diseño de un programa, el *ReCor*, que permite evaluar de forma exacta y a posteriori la representatividad de un corpus. En sus razonamientos previos al diseño del programa se aprecia la lógica en la que éste se basa.

“A possible solution could be to analyse the lexical density of a corpus in relation to the increase in documentary material included (Corpas Pastor and Seghiri Domínguez, 2006, and Seghiri Domínguez, 2006). In other words, if the ratio between the actual number of different words in a text and the total number of words (types/tokens) is an indicator of lexical density or richness, it may be possible to create a formula that can represent increases in the corpus on a document by document basis: the number of types does not increase in proportion to the number of words the corpus contains, once a certain number of texts has been achieved (Corpas Pastor y Seguiri Domínguez 2008).

1.4.3. Preparación del corpus

Hasta el momento hemos descrito varios temas directamente relacionados con el diseño de corpus en general y la selección y manejo de textos en particular. Avanzamos ahora hasta la siguiente fase, la de la preparación de los textos para su consiguiente análisis con el software apropiado. Esta fase de preparación implica el marcado o la anotación de algunos rasgos relevantes al tipo de estudio que se quiera llevar a cabo.

Según la definición de Sinclair, el término “marcado” o *mark-up* significa “added codes which encode only surface features of texts that would otherwise be lost in transfer to a machine” (Sinclair 1995: 25), mientras que la “anotación” o *annotation* se define como “added codes which encode analyses and interpretations” (ídem.: 20).

Esta fase de preparación es, hasta cierto punto, opcional, ya que puede que el investigador, compilador o docente prefieran trabajar bien con texto plano (textos que no han sufrido ningún tipo de marcado o adición de información) o bien con textos mínimamente marcados (textos en los que sólo se han marcado algunos rasgos esenciales). Para poner un ejemplo, en un texto se pueden haber marcado divisiones como frases, párrafos, capítulos y así sucesivamente. Los segmentos como los encabezados, títulos, notas al pie y similares, o estilos como la negrita, la cursiva, etc. también podrían marcarse de tal manera que el usuario del corpus tuviera acceso a esa información. Hay diversas maneras en que un corpus puede estar anotado a través de etiquetas. Hay ejemplos de anotación de tipo gramatical, sintáctica, semántica, discursiva, fonética, prosódica e incluso anotación de problemas, entre otros. El tipo de anotación seleccionada dependerá del tipo de estudio que se quiera llevar a cabo. Por ejemplo, si nuestra investigación se centra en la desambiguación de significados, entonces estaríamos interesados en trabajar con un corpus anotado semánticamente.

También hay distintos formatos de anotación, aunque se debería apuntar que se recomienda ampliamente el uso de los estándares creados para tal fin. Las iniciativas más relevantes vienen del grupo oficialmente establecido y conocido como EAGLES (Expert Advisory Group for Language Engineering Standards). Es un comité de la UE que, entre otras cosas, promueve el uso de un formato de marcado de documentos en concreto, el SGML (Standard General Marked Language), y el uso de convenciones como la TEI (Text Encoding Initiative). Esta iniciativa tiene como objetivo proporcionar implementaciones estandarizadas para el intercambio de textos en formato electrónico. Finalmente, el esquema de anotación utilizado para un corpus debe estar descrito en el “Document Type Description”, que es la representación formal que informa al usuario o programa informático de qué elementos contiene el texto y cómo se combinan.

En cuanto a la tarea en sí de anotar o asignar etiquetas, es posible hacerlo manualmente o automáticamente. Las ventajas del etiquetado o la anotación son que permite la recuperación automática de todos los ejemplos de una palabra o expresión que porten la misma etiqueta, que es un código que se adjunta a cada palabra o segmento de un texto. La decisión de trabajar con un corpus anotado o de texto plano dependerá del tipo de estudio que se quiere llevar a cabo. Como fase opcional que es, no entraremos a dar una descripción muy detallada de todas las cuestiones teóricas y prácticas implicadas en la anotación, pero ofreceremos una idea global de sus principios y el tipo de herramientas informáticas relacionadas.

Como ya hemos apuntado, existen básicamente dos procesos dentro de la fase de preparación del texto: el marcado y la anotación. Parafraseando la definición de Sinclair de los términos en cuestión, el marcado consiste en reflejar o marcar rasgos que ya están presentes en el texto con el fin de mantenerlos, mientras que la anotación implica añadir rasgos que no estaban antes. Ambos procesos pueden llevarse a cabo adjuntando etiquetas o *tags*. Aquí nos centraremos en la anotación lingüística, que es el proceso más complejo, si bien se debería subrayar la importancia del marcado, especialmente si el análisis que se quiere efectuar implica la alineación de textos.

En cuanto a la anotación lingüística, existen varios tipos y todos ellos pueden llevarse a cabo manualmente. Por otro lado, se han desarrollado herramientas informáticas para aliviar, aunque sólo hasta cierto punto, la tarea del anotador cuando trabaje con determinados tipos de anotación.

La anotación “discursiva” y la “textual” son dos tipos de anotación que necesitan realizarse a mano. La primera implica añadir etiquetas del discurso a categorías lingüísticas como las disculpas, los rodeos (p.ej. en inglés “sort of”, “kind of”), saludos y similares, para poder llevar a cabo algunos tipos de análisis discursivo. Un ejemplo de los elementos que se marcan en una anotación lingüística textual es la anáfora para poder así estudiar las relaciones anafóricas establecidas dentro de un texto.

La anotación de la “transcripción fonética” y la “prosodia” también han de realizarse manualmente. En el primer caso, la anotación puede resultar especialmente problemática ya que los límites de los sonidos no son claros y la transcripción se hace normalmente utilizando los caracteres disponibles en vez de los símbolos fonéticos. La anotación de rasgos prosódicos no sólo tiene que efectuarse a mano, sino también bajo la cuidada atención de expertos fonetistas que quizás tampoco se pondrían de acuerdo entre ellos. Mantener la consistencia se presenta, por tanto, como una tarea difícil y la recuperación del material puede verse obstaculizada por los símbolos introducidos entre las sílabas o los segmentos de un elemento.

La anotación de “problemas” tiene dos características fundamentales: (1) no es exhaustiva, y (2) utiliza un esquema de anotación seleccionado por la relevancia de las distinciones que hace en términos de las preguntas específicas que cada analista desea preguntar. Por estas características tan específicas, este tipo de anotación se tiene que hacer inevitablemente a mano ya que a un investigador le interesará, por ejemplo, anotar problemas de un tipo y a otro investigador de otro tipo.

1.4.3.1. Anotación gramatical

El tipo más común de anotación es la gramatical o morfosintáctica, que implica la asignación de una etiqueta a cada unidad léxica del texto para indicar su categoría gramatical. Este proceso, que requiere del empleo de mucho tiempo si se hace manualmente, puede agilizarse con etiquetadores automáticos (programas informáticos que, a priori, pueden realizar la misma tarea que un anotador humano). El primer programa de este tipo fue el *TAGGIT* (Greene y Rubin 1971), desarrollado específicamente para la anotación gramatical del corpus Brown. Funcionaba siguiendo plantillas basadas en reglas, es decir, conjuntos de reglas gramaticales creadas por expertos lingüistas que se basaban en su conocimiento e intuición. Como resultado de un proyecto pensado para la anotación del corpus *LOB*, varios años después se desarrolló un nuevo sistema de etiquetación automática. En esta ocasión, el *CLAWS* (*Constituent Likelihood Automatic Word-tagging System*) (Garside 1987) no estaba basado en reglas lógicas o lingüísticas, sino en principios probabilísticos. Más adelante se desarrollarían varias versiones para distintos propósitos tomando la versión inicial del *CLAWS* como base. Un caso así es el de la etiquetación gramatical del *BNC*, para la que se utilizó la cuarta versión (*CLAWS4*).

¿Qué son y qué aspecto tienen las etiquetas?

Según la definición de McEnery y Wilson, una etiqueta o *tag* es “a code attached to words in a text representing some feature or set of features relating to those words” (McEnery y Wilson 1996: 179).

Para ejemplificarlo, el conjunto de etiquetas morfológicas del *BNC* consiste en 139 etiquetas como las que se ofrecen en la lista que sigue. Se debería apuntar que aún no se ha llegado a consensuar un conjunto de etiquetas estándar, lo que ha propiciado una gran diversidad en la representación de las mismas. Seguidamente se muestran algunos ejemplos de etiquetas extraídas del conjunto utilizado para la anotación del *BNC*:

NN – nombre común, número neutro (p.ej. *sheep*)

NN1 – nombre común, singular (p.ej. *difficulty*)

NN2 – nombre común, plural (p.ej. *corridors*)

AJO – adjetivo (p.ej. *white*)

AVO – adverbio (p.ej. *usually*)

VVO – forma básica de un verbo léxico (p.ej. *give, work*)

VVD – tiempo pasado de un verbo léxico (p.ej. *gave, worked*)

VVG – participio en -ing de un verbo léxico (p.ej. *giving, working*)

VVZ – forma en -s de un verbo léxico (p.ej. *gives, works*)

Respecto al uso de los etiquetadores, normalmente hay cuatro pasos implicados en el proceso automático de etiquetación (Kennedy 1998):

- (1) El sistema busca una palabra determinada en un lexicón en formato electrónico.
- (2) Si no encuentra esa palabra, el sistema la envía a un analizador morfológico que la descompone en lexema y afijos y la devuelve al lexicón, y así sucesivamente, hasta que la palabra se reconoce.
- (3) Puede darse el caso de que el texto contenga una secuencia imposible de descomponer, como una expresión idiomática sintáctica (dos o más palabras que van juntas y tienen una sola función sintáctica, como por ejemplo, los verbos frasales ingleses). Para estos casos el sistema cuenta normalmente con un lexicón especializado que contiene secuencias de este tipo para que puedan ser reconocidas y se les pueda añadir una etiqueta.
- (4) Si aún existe un grado de ambigüedad el sistema puede acudir a un set de reglas lingüísticas o cálculos probabilísticos con el fin de conseguir la etiqueta correcta.

Se debería apuntar que ningún etiquetador morfológico ha conseguido nunca llegar a un 100% de precisión en la asignación de etiquetas. Esto significa que siempre serán necesarias una revisión y una post-edición cuidadas hechas por un humano.

1.4.3.2. Lematización

La lematización es un tipo de anotación lingüística diferente que consiste en la reducción de las palabras en un corpus a sus respectivos lexemas. Su éxito depende de una etiquetación morfológica fiable lo cual, como ya hemos mencionado (cfr. supra 1.4.3.1.), está aún por perfeccionar. Además, se hace necesario aproximarse a un corpus lematizado con precaución ya que no puede tratarse de la misma forma que uno que contenga material “en estado puro”. Por poner un ejemplo, la lematización puede afectar a los datos extraídos de un corpus en términos de cómputo de palabras y análisis estadístico. Un ejemplo de lematizador automático es *PALIC*, que también es un etiquetador y se ha desarrollado en la Universitat Pompeu Fabra, en Barcelona, para utilizarlo con los textos en español que pertenecen al corpus terminológico del IULA (*Institut Universitari de Lingüística Aplicada*).

1.4.3.3. *Anotación sintáctica*

Pasando a un nivel superior de complejidad, la anotación sintáctica o *parsing* no sólo necesita de una etiquetación morfológica fiable, sino también de un análisis sintáctico fiable. El *parsing* consiste en la asignación de categorías sintácticas a las oraciones contenidas en un corpus. Como también ocurre con la anotación morfológica, existen dos aproximaciones al desarrollo de *parsers* automáticos: (1) basados en reglas lingüísticas, y (2) basados en probabilidades de aparición. Sin embargo, se dan demasiadas dificultades en un texto para que un sistema automático de análisis sintáctico funcione con un 100% de acierto. Algunas de estas dificultades dignas de mencionar son los sintagmas preposicionales, las elipsis o las entradas en falso, que son elementos que probablemente dificulten la tarea a un analizador automático, de lo que se desprende que, una vez más, la post-edición humana será necesaria.

1.4.3.4. *Anotación semántica*

Por último, la anotación semántica es incluso más difícil de realizar para un ordenador con cierto grado de precisión aún cuando el sistema sea probabilístico o basado en reglas. Existen dos tipos de anotación semántica: (1) el marcado de las relaciones semánticas entre elementos del mismo texto, y (2) el marcado de significados. Un etiquetador semántico automático encontrará en ambos casos dificultades considerables relacionadas con la desambiguación de sentidos, la polisemia y las metáforas.

En conclusión, la imagen descorazonadora que ofrece este apartado en cuanto a las herramientas informáticas que se han desarrollado para la anotación de textos no se debe a defectos en los programas mismos. La razón por la que el software de anotación disponible en la actualidad no ha alcanzado una precisión total tiene menos que ver con los ordenadores que con los problemas que aún tenemos los humanos para explicar la lengua de una forma procesable por un ordenador. Como veremos en el apartado siguiente, se dan muchos otros casos en los que los ordenadores pueden ayudar a los humanos a obtener un extremadamente alto nivel de precisión y fiabilidad en el trabajo con corpus.

En la enseñanza de la traducción también se da el trabajo con corpus anotados, sobre todo con corpus ya construidos como puede ser el caso del *BNC* o el corpus del IULA,

etiquetados morfológicamente para permitir la búsqueda de, por ejemplo, cadenas de categorías gramaticales, algo a lo que se le pueden dar usos didácticos.

1.4.4. Análisis

Una vez que se han completado la fase 1 (selección del corpus o los textos y temas relacionados) y la fase 2 (preparación del texto, si es el caso), el corpus está preparado para ser analizado, siempre teniendo en cuenta el tipo de estudio o trabajo para el que está destinado.

En cuanto a la parte técnica del análisis, en las últimas décadas se han desarrollado muchos programas para asistir a los investigadores en el procesamiento de textos en formato electrónico. Sin embargo, la disponibilidad de estas herramientas no significa que se pueda ignorar el aspecto metodológico del análisis. De hecho, algunos de los asuntos más importantes relacionados con la metodología están basados en el tipo de estudio que se quiere llevar a cabo (discursivo, sociolingüístico, contrastivo, etc.) y el tipo de preguntas que pueden dirigirse a un corpus (abiertas, cerradas, relacionadas con información estadística, etc.). Un corpus es una fuente de datos potencialmente rica de donde el investigador extraerá resultados o en base a la cual llegará a conclusiones. Para hacer esto, tendrá que decidir primero cómo enfocar la tarea de extraer del corpus datos relevantes para su estudio. Por ejemplo, un corpus no podría proporcionar una respuesta a una pregunta abierta como “¿Cuáles son las principales diferencias entre los textos producidos por hombres y mujeres en lo que respecta al discurso?”. Y lo mismo se aplicaría a una pregunta cerrada como “¿Está marcado ideológicamente el idiolecto de un autor determinado?”. Para encontrar respuestas a estas preguntas el investigador debería pensar en las formas como se explicita el discurso o la marca ideológica, es decir, en cómo se reflejan estos aspectos en la superficie del texto para después introducirlos en forma de búsquedas en un programa de análisis de corpus. Para responder a la primera cuestión se podrían buscar rasgos discursivos tales como evasivas, adjetivos o rasgos suprasegmentales como la entonación. La segunda cuestión podría investigarse buscando palabras que tienen una carga ideológica potencial.

Volviendo al software disponible para el trabajo y la investigación basados en corpus, ya hemos comentado algunas herramientas apropiadas para la fase de preparación del corpus. Nos centraremos ahora en el análisis de corpus y algunas de las herramientas

desarrolladas para automatizar y, por ende, facilitar la tarea del investigador y usuario de corpus.

1.4.4.1. Operaciones básicas

Hay una amplia gama de paquetes de software que incluyen operaciones estándar disponibles para el público. Una de las operaciones más conocidas y utilizadas es la generación de concordancias. Una concordancia, según la definición de McEnery y Wilson, es “a formatted version of display of all the occurrences or tokens of a particular type in a corpus” (McEnery y Wilson 1996: 247). Otras funciones estándar incluyen la generación de listas de palabras y estadísticas.

1.4.4.2. Herramientas informáticas

Un programa de extracción de concordancias permite al usuario ver todos los casos de una palabra o frase en particular dentro de un corpus, junto con el segmento de texto en el que aparecen. Este segmento se conoce como la línea de concordancia (Pearson 1996). Según Zanettin (1998), el procedimiento básico de investigación con corpus consiste en producir múltiples líneas de concordancias, también conocido como concordancia KWIC (Key Word In Context). La *key word* o palabra clave se refiere a la palabra buscada y que también aparece en la bibliografía en inglés como *target item*, *node word* o *search item*. Aparte de listar todos los casos de una palabra (algunos programas también facilitan concordancias de oración o párrafo, un formato que se conoce como KWAL - Key Word and Line), y revelar los contextos en los que aparece, las concordancias KWIC también pueden ordenarse de distintas formas (alfabéticamente, a la derecha o la izquierda de la palabra clave). También pueden expandirse las líneas para que se muestre más contexto, y algunos programas incluso ofrecen un perfil colocacional de la palabra clave en un conjunto de concordancias. Otras utilidades ofrecidas por algunos paquetes de software son, por ejemplo, el listado de todas las palabras de un corpus o de una parte del mismo, ordenado por frecuencia o alfabéticamente. Los perfiles de frecuencias de palabras ofrecen información estadística útil para investigar diferentes medidas en un corpus, como el ratio formas/palabras o *type-token ratio* y la densidad léxica. El *type-token ratio* es la frecuencia de aparición de formas (*types*) comparado con el número total de palabras (*tokens*) en el corpus, mientras que la densidad léxica es “the percentage of lexical as opposed to grammatical items in a given text or corpus of texts” (Baker 1995: 237). Ésta última ofrece información sobre la variedad del vocabulario utilizado en un corpus. Por último, las

herramientas informáticas también facilitan información extralingüística si se han diseñado para tal efecto. Esto se puede conseguir a través de cabeceras que pueden mostrarse ofreciendo distintos tipos de datos, como la fuente original del texto, el nombre del autor o del traductor, si es el caso, etc.

Entre las herramientas diseñadas para la búsqueda y recuperación de datos en un corpus en general, el *Oxford Concordance Program (OCP)* estuvo entre los más utilizados en el pasado. Creado en 1981, puede que el *OCP* no sea el programa más rápido del mercado, pero es fiable y flexible y produce concordancias KWIC y listas de palabras, además de ofrecer estadísticas textuales básicas.

También creado en los 80, *WordCruncher* también fue uno de los programas más utilizados para extraer concordancias KWIC y KWAL, pudiendo también generar concordancias de colocaciones seleccionadas por el usuario, así como facilitar información estadística sobre la frecuencia de un elemento. Necesita que los textos estén indexados en un formato especial, lo que significa que los textos deben estar organizados en archivos anotados especialmente. Por esta razón, el *WordCruncher* consiste de hecho en dos programas: *WC Index*, un indexador, y *WC View*, que presenta todas las funciones mencionadas. Gracias al proceso de indexado previo este programa es capaz de funcionar más rápido que otros y también es interactivo.

Existe un programa llamado *TACT (Text Analysis Computer Tools)* que es muy parecido al *WordCruncher*. Un rasgo diferente del *TACT* es que permite al usuario producir gráficos de la distribución de las palabras a lo largo de un texto o un corpus. También ofrece información estadística de mayor complejidad y permite agrupar palabras para realizar búsquedas de acuerdo con categorías definidas por el usuario, tales como campos semánticos o una clase de palabra específica. *TACT* también trabaja con texto que antes ha de haber sido indexado por el programa *MAKBAS*. En un plano un tanto más sencillo, el *SCP 4.0.7 (Simple Concordance Program)* es también un generador de concordancias y de listas de palabras, con la característica de que es gratis. También gratis pero mucho más potente como herramienta, el *AntConc* destaca por la variedad de funciones que ofrece, muy similares éstas a las del conocido programa *WordSmith Tools*. Comercializado a través de Oxford University Press y desarrollado por Mike Scott (1996) en base a otro programa llamado *MicroConcord*, el *WordSmith Tools* pertenece a una categoría de software de nueva generación diseñado para ser capaz de manejar enormes cantidades de datos con velocidad, precisión y flexibilidad. Aparte de ofrecer las operaciones estándar como la creación de

concordancias, listas de palabras y similares, *WordSmith Tools* también realiza otras operaciones que le hacen aún más atractivo. Un ejemplo de ello es la opción de realizar búsquedas complejas introduciendo etiquetas, asteriscos, operadores booleanos y secuencias discontinuas, además de ofrecer información colocacional, todo ello de una forma muy agradable para el usuario. Las nuevas versiones del programa (4.0 y 5.0) integran también un localizador y descargador de páginas web con el objetivo de facilitar la creación de corpus ad hoc.

Otros programas también de la categoría de software de nueva generación son el *LEXA* (Hickey 1993) y el programa que utiliza el corpus ICE (*ICECUP*) (Quinn 1993). El *LEXA* es una herramienta potente y versátil para el manejo y análisis de corpus que consta de una amplia gama de programas útiles que ofrecen un gran número de operaciones diversas. El *ICECUP* destaca por sus interesantes funciones adicionales como la búsqueda en un corpus o una parte del mismo de cadenas de elementos, tanto si son colocaciones como estructuras gramaticales. Digna de mención también es la aplicación *SGML-Aware Retrieval Application (SARA)*, que es el motor de búsqueda por defecto diseñado para el *British National Corpus (BNC)* y, por último, el *CQP* (Christ 1994), que es un software de búsqueda y recuperación avanzado capaz de manejar corpus muy vastos y hacer búsquedas complejas de forma rápida.

También hay paquetes de software que han sido diseñados para propósitos más específicos. El *Child Language Analysis (CLAN)* es un ejemplo de programa diseñado para su uso con la base de datos CHILDES (MacWhinney y Snow 1990) para la investigación de la adquisición de la lengua. *COALA* es un sistema de software más complejo y particularmente adecuado para la investigación de sistemas lingüísticos de transición, es decir, interlenguas (Pienemann 1992). *VocabProfile*, por otro lado, es un programa de *MS-DOS* para comparar el solapamiento de vocabulario entre diferentes textos y el grado de repetición.

También hay ejemplos de software que ha sido desarrollado especialmente para objetivos educacionales. Éste es el caso del *Longman Mini Concordancer (LMC)* (Chandler y Tribble 1989), *MicroConcord* (Murison-Bowie 1993), *MonoConc 1.5* y su versión mejorada, *MonoConc Pro 2.0* (Barlow 2000). Básicamente, todos estos paquetes de software son programas de concordancias diseñados para su uso en el aula. Ofrecen diversas opciones de búsqueda, reordenamiento, navegación, edición y formateado de los datos, con un especial énfasis en proporcionar un diseño agradable.

Otro tipo de programa utilizado en la fase de análisis es el alineador, que coloca normalmente dos textos en paralelo para mostrar frente a frente las frases o párrafos de un texto y sus correspondencias en otro (normalmente, la traducción del primer texto). Para que esta operación sea posible, los textos deben estar previamente marcados al menos a nivel de frase o párrafo. *MultiConcord*, un programa creado por David Woolls dentro del marco del proyecto europeo LINGUA, es un ejemplo de software diseñado para el manejo de corpus que contiene un alineador. Consta de dos aplicaciones, *MinMark*, que marca los textos, y *MultiConcord*, que permite al usuario llevar a cabo el análisis de los textos. Además de esto, *MultiConcord* también cuenta con funciones que permiten la generación de ejercicios de blancos, al igual que otros programas en lenguaje PERL (Practical Extraction and Reports Language). El *WordSmith Tools* ofrece la posibilidad de alinear y mostrar textos en paralelo sin necesidad de que los textos estén marcados ya que, por defecto, el programa se guía por los signos de puntuación del texto a la hora de segmentarlo. De la misma forma, otros programas concebidos para la traducción asistida por ordenador (TAO), tales como el *WinAlign* de Trados, por ejemplo, pueden alinear textos y almacenarlos alineados para más adelante utilizarlos como una memoria de traducción o un corpus paralelo.

Las herramientas que se han descrito en este apartado son sólo algunos ejemplos del software más relevante y disponible en la actualidad. Sin embargo, sería interesante apuntar que hay otros tipos de software que, aunque no están diseñados específicamente para el trabajo con corpus, sí pueden ser de gran utilidad. Por ejemplo, sin entrar en demasiados detalles, existen programas que facilitan la ubicación de páginas web, como el *Copernic Agent 6.1*, y otros capaces de descargar páginas web completas, como el *WebRecorder 2.0*, el *WebCopier 2.1a* o el *Offline Explorer*, que son muy convenientes a la hora de compilar corpus. Si ya se tienen archivos almacenados en una unidad local del ordenador el programa *InfoRapid Search & Replace*, originalmente pensado para hacer búsquedas y realizar cambios en varios documentos a la vez, permite extraer concordancias simples. Finalmente, los corpus que se pueden acceder en la web ofrecen sus propios mecanismos de búsqueda. En estos casos, el usuario normalmente tiene que introducir una palabra o cadena de palabras de búsqueda y seleccionar el corpus o la parte del corpus en la que está interesado. Otros programas independientes permiten tratar toda la Internet pública como si fuera un corpus; dos ejemplos son el *WebCorp* y el *KwicFinder*. En el apéndice de esta tesis ofrecemos un listado de corpus gratis en Internet que incluyen entre sus lenguas el inglés, el español o ambas y el software citado más relevante.

Hasta cierto punto todas las herramientas presentadas en este apartado han contribuido al desarrollo de la lingüística de corpus como metodología. En el lado cuantitativo del análisis, el software diseñado específicamente para la investigación con corpus ha proporcionado a la lingüística de corpus resultados fiables y un medio de justificar todos los datos. Sin embargo, no se debe ignorar la parte cualitativa. Por defecto, el software disponible se centra normalmente en proporcionar datos cuantitativos en términos de listas de frecuencia, estadísticas y similares, lo que implica el riesgo de que el investigador desatienda otras variables. En este sentido, parece que McEnery y Wilson han encontrado un punto de equilibrio cuando afirman que “qualitative analysis can provide greater richness and precision whereas quantitative analysis can provide statistically reliable and generalisable results” (McEnery y Wilson 1996: 63).

En cualquier caso, es de esperar que la disponibilidad de corpus grandes y software potente mejore las descripciones de las lenguas e influya de manera importante en otras aplicaciones. Mirando hacia el futuro, aún parece faltar mucho antes de que el software sea capaz de manejar corpus ingentes en cuestión de segundos, tratar con diferentes esquemas de codificación de corpus y diferentes alfabetos, o combinar imagen y texto, por ejemplo.

Como se ha apuntado al inicio de este apartado, la fase del análisis representa el momento de realizar las preguntas al corpus, valorar su utilidad y, en el caso que nos ocupa, aprender a traducir descubriendo hechos gracias al corpus.

1.4.5. *Explotación y reelaboración de hipótesis*

El presente apartado se refiere a lo que hemos llamado la fase final en un estudio de corpus, que es la explotación real del corpus, es decir, la extracción final de resultados y conclusiones y, si cabe, la reformulación de las hipótesis iniciales. Los resultados, en cualquier caso, siempre se habrán de interpretar a la luz de la cantidad y calidad de los textos que conforman el corpus, de ahí la importancia de la fase de selección de textos.

2. CORPUS ELECTRÓNICOS EN ESPAÑA Y TRADUCCIÓN

A continuación se presenta una visión general de la situación del trabajo basado en corpus en España o con la lengua española. A pesar de que no es España uno de los países más punteros en lo que a investigación basada en corpus se refiere, se debería subrayar que esta situación está cambiando gradualmente gracias al interés que algunas instituciones públicas y privadas y empresas han puesto en el uso de corpus para diversos fines. Entre las instituciones públicas que han fomentado la construcción y estudio de corpus están la Real Academia Española, la Universitat Autònoma de Barcelona, la Universitat de Barcelona, la Universitat Pompeu Fabra, la Universitat Jaume I, la Universidad Autónoma y Universidad Politécnica de Madrid, la Universidade de Vigo y el Instituto Cervantes, entre otros. En cuanto a entes privados, compañías como Telefónica e IBM, y editoriales como Anaya y SM también han contribuido notablemente al desarrollo del trabajo con corpus.

Al igual que en los países anglosajones, en España se ha trabajado con corpus en una gran variedad de campos. También aquí estos recursos se han utilizado en la creación y mejora de diccionarios, como en los casos del *Diccionario de la Real Academia Española de la Lengua (DRAE)* y algunos Vox, la extracción de terminología, el estudio de la adquisición de la lengua y el estudio de la lengua española en general y sus variantes geográficas. Otras áreas que también se han beneficiado de enfoques basados en corpus incluyen los estudios literarios, la psicolingüística y el procesamiento del lenguaje natural. En este último caso destacan los esfuerzos centrados en el desarrollo de sistemas de reconocimiento de voz o los trabajos con analizadores y desambiguadores morfológicos (Màrquez y Padró 1997; Civit, Martí y Padró 2003). Todas estas aplicaciones han tenido un efecto de mayor o menor grado tanto en el mundo de la empresa como el académico. Sin embargo, también es cierto que el conocimiento acerca de la investigación y aplicaciones basados en corpus no está aún muy extendido dentro de nuestras fronteras, lo cual podría deberse a que el acceso a muchos de los corpus existentes en España está restringido así como que muchos proyectos de corpus están aún en marcha y no están disponibles.

Para compensar esta falta de conocimiento sobre lo que se ha hecho y se está haciendo en España en cuanto a trabajo con corpus se refiere, ofrecemos aquí información sobre algunos de los corpus que existen en la actualidad para el estudio de la lengua española,

sin dejar por ello de mencionar importantes corpus en otras lenguas oficiales del país como son el *Corpus de referencia do galego actual (CORGA)* o el *Corpus Técnico do Galego (CTG)*, el *Corpus estadístico de eusquera del siglo XX* y el *Corpus Textual Informatizat de la Llengua Catalana* del Institut d'Estudis Catalans, que han proporcionado a la comunidad científica tanto estudios como productos concretos. De entre los corpus descritos aquí, algunos representan proyectos completados y otros aún en marcha, contienen textos escritos, orales o ambos, y han sido construidos para fines generales o específicos. Se debería apuntar, no obstante, que esta relación no es exhaustiva y sólo pretende dar una idea de las aplicaciones que han atraído más atención y de las instituciones que han jugado un papel más importante en este campo de investigación hasta la fecha. También se debería subrayar que aunque la mayoría de los corpus descritos son exclusivamente de español, también se incluyen otros multilingües que tienen un componente en esta lengua.

2.1. Corpus escritos

2.1.1. Corpus generales

Como ya se estableció en un apartado anterior (cfr. supra 1.4.2.1.2.), los corpus pueden dividirse entre generales o especializados, dependiendo de si sus textos pertenecen a un dominio en particular o no. Así, un corpus general puede servir como fuente de información textual para muchos y diversos usos.

La Real Academia Española (RAE) es la responsable de la creación de dos de los corpus de español general más extensos que tienen como fin representar la lengua española. Éstos son el *Corpus de Referencia del Español Actual (CREA)* y el *Corpus Diacrónico del Español (CORDE)*. El *CREA*, diseñado para ser un corpus de referencia del español en continua actualización, está compuesto por textos de ámbitos diversos tanto escritos como orales, de distintas variedades del español y que pertenecen a los últimos 25 años de la historia reciente. Su aplicación de búsqueda gratuita en Internet facilita la selección de subcorpus en términos de campo temático y variante geográfica. Por otro lado, el *CORDE* es un corpus diacrónico que cubre la lengua española desde sus inicios hasta donde comienza el *CREA* y, juntos, suman unos 410 millones de palabras. Pero la RAE no sólo se ocupa de la creación y mantenimiento de los citados corpus, sino que además dedica esfuerzos al diseño de herramientas de procesamiento del lenguaje natural y recursos para anotar los corpus morfosintácticamente. De manera similar,

aunque a menor escala, el ya citado profesor Mark Davies ha compilado un corpus de 100 millones de palabras que también pretende ser representativo de la lengua española en general (desde el siglo XIII hasta el XX), también se puede acceder gratuitamente en Internet y además ofrece la posibilidad de realizar búsquedas por lema, categoría gramatical, sinónimos y antónimos, etc., lo que le convierte en un recurso especialmente valioso.

El *Archivo de textos hispánicos de la Universidad de Santiago (ARTHUS)* es también un corpus general que incluye textos en español de diferentes periodos y géneros e incluye transcripciones de material oral.

CUMBRE es un corpus de 20 millones de palabras que contiene material compilado por la editorial española SGEL S.A. para usos descriptivos y lexicográficos y que contiene una gran variedad de textos de origen literario, periodístico y oral.

2.1.2. Corpus específicos

El número de corpus construidos en España para fines específicos supera al número de corpus generales, debido muy probablemente al esfuerzo de compilación que suponen los últimos. Entre los primeros destacan varios corpus monolingües y bilingües que contienen textos de campos específicos, como el *CRATER (Corpus Resources and Terminology Extraction)*, que es un corpus multilingüe, alineado y anotado, compuesto de textos del campo de las telecomunicaciones en español, inglés y francés. El *CRATER* es un ejemplo de cooperación entre socios académicos y comerciales para investigar en lingüística computacional y traducción automática. Dentro del marco de este proyecto también se creó un etiquetador morfosintáctico para el español y un alineador automático, lo que demuestra que España no sólo es activa en términos de compilación de corpus, sino también en el desarrollo de software relacionado. De manera similar, el Institut Universitari de Lingüística Aplicada (IULA) de la Universitat Pompeu Fabra (UPF) está implicado en la compilación de un corpus terminológico multilingüe en cinco lenguas (español, catalán, inglés, francés y alemán) que cubre varios campos temáticos (derecho, economía, medioambiente, medicina e informática). El objetivo con el que se construye este *Corpus Tècnic de l'IULA* es el de realizar investigaciones en áreas como el reconocimiento y extracción de terminología, la variación lingüística, el alineamiento de textos, la creación de diccionarios y tesauros electrónicos, la enseñanza de lenguas, etc. Además, para la preparación y análisis del corpus en el IULA también se han desarrollado varias herramientas informáticas originales como un alineador

automático, un lematizador (*PALIC*) y un desambiguador para español y catalán (*AMBILIC*). Otro proyecto del IULA es el *BancTrad*, una plataforma que permite acceso web a corpus anotados mono y multilingües; uno de ellos contiene traducciones de estudiantes de la UPF corregidas, de profesores, editoriales e Internet y los textos suman 2 millones de palabras anotadas a diversos niveles.

En la Universidad de Deusto también se ha construido otro corpus especializado, el *LEGE BIDUN*, que es un corpus paralelo de textos jurídico-administrativos en euskera y español diseñado para contribuir a la estandarización terminológica y la traducción automática. Otro corpus paralelo, en este caso de documentos de la Organización Internacional del trabajo, es el *ECI/MCI (European Corpus Initiative Multilingual Corpus)*, que también incluye textos de otras fuentes. El proyecto *LEJES*, desarrollado conjuntamente entre las universidades de Granada y Bonn, es un corpus de textos legales que se ha utilizado para la investigación del lenguaje legal en español y, para llevar a cabo investigaciones de épocas pasadas, está el *AITJUM (Archivo Informatizado de Textos Jurídicos Medievales)* de la Universitat Autònoma de Barcelona (UAB). Sirviendo a otros propósitos, diversos textos de la Unión Europea y las Naciones Unidas también se han recogido en corpus como el *JOC-CES*, que contiene documentos procedentes del *Journal of the European Community* y del *Corpus Encoding Standard*, o el *UN Parallel Text Corpus*, que contiene textos de la Oficina del Servicio de Conferencias de la ONU, respectivamente. El corpus *VISL* permite hacer búsquedas en línea a textos del Parlamento Europeo y también periodísticos. En relación con éste último, el *Corpus Textual del Español Periodístico* se encuentra en la UAB y, a pesar de que se limita a textos de un medio específico, aún puede servir para estudios diversos. Finalmente, el *Corpus Lingüístico da Universidade de Vigo (CLUVI)* es un corpus paralelo de más de 22 millones de palabras que contiene muchos subcorpus de distintas especialidades y que casi siempre combinan el gallego con otras lenguas.

Otros corpus se han construido con objetivos concretos como el de investigar aspectos léxicos, morfológicos o sintácticos de la lengua española. Algunos ejemplos de ello son el corpus desarrollado dentro del proyecto *SISCOOR (Sistemas de Consultas Coordinadas)* en la Universidad Politécnica de Valencia. En cuanto a entidades privadas, editoriales como SM Diccionarios y Bibliograf S.A. también han compilado corpus para fines lexicográficos.

Algunos corpus han sido compilados con el propósito de comprobar una base de datos construida anteriormente, como es el caso de *CorVerifSDGEE (Corpus de Verificación del Sistema de Diccionarios y Gramáticas Electrónicos del Español)* de la Universitat Autònoma de Barcelona. El objetivo de otros ha sido, por ejemplo, la validación de programas de software, como es el caso de *LAN*, un corpus terminológico elaborado por la compañía Micro Focus S.A., con el fin de validar etiquetadores automáticos para el español. En esta misma línea están el *UAM Treebank*, el *LexEsp (Léxico informatizado del español)*, el *CLiC-TALP*, el *BDS (Base de datos sintácticos del español)* y el *Cast3LB*, todos ellos elaborados y anotados por distintas universidades españolas para su aplicación en áreas como la desambiguación automática, entre otras.

La investigación dentro del campo de la adquisición de primera y segundas lenguas también se ha beneficiado de la construcción de varios corpus. Por ejemplo, existe un corpus disponible en la Universidad de Granada para la investigación en la adquisición y desarrollo de la primera lengua que consiste en textos producidos por niños de entre 6 y 8 años, y el *CHILDES*, sistema para el estudio del lenguaje infantil, tiene un componente de español. Por otro lado, el corpus bilingüe de español y francés que está en proceso de construcción en la Universidad de Sevilla permitirá llevar a cabo análisis contrastivos y estudios dentro de la adquisición de segundas lenguas.

Por último se encuentran otros corpus restringidos a un tema o incluso un texto como es la Biblia. Ejemplos de ello son el *BEC (Biblioteca Electrónica Cristiana)*, el *Bible of University of Maryland Parallel Corpus Project* o la *Polyglot Bible* de Mark Davies. En cuanto a lo literario, las obras completas de autores se han puesto a disposición en Internet para poder estudiarlas en línea, como es el caso de las obras de Berceo, por mencionar un ejemplo.

2.2. Corpus orales

2.2.1. Corpus generales

El corpus oral más importante que se encuentra disponible para el estudio general del español oral es el *Corpus Oral de Referencia de la Lengua Española Contemporánea (CORLEC)*, de la Universidad Autónoma de Madrid. Este corpus es digno de mención por la gran variedad de situaciones de habla que contiene, si bien es cierto que ya está en marcha un nuevo y más ambicioso proyecto, el *C-Oral-Rom*, que es un corpus oral

de referencia de cuatro lenguas romances. Por otro lado, también existen otros corpus menores representativos de variantes orales del español⁴.

2.2.2. *Corpus específicos*

Los corpus incluidos en esta sección son aquellos que se han construido bien con el objetivo de describir aspectos concretos de la lengua oral, o bien de caracterizar las variaciones entre el español oral de distintas regiones o grupos de edad. Por último, otros corpus se han construido con el fin de desarrollar determinados sistemas de procesamiento del lenguaje natural.

El corpus *DIES-RTP* (*Difusión internacional del español por radio, televisión y prensa*) se construyó para el análisis léxico del español oral y contiene material extraído y transcrito de diferentes medios de comunicación. El *DIES-RTP* está disponible en la Universidad de Alcalá de Henares, donde otro corpus, el *ACUAH*, se ha desarrollado para el estudio de la conversación. El estudio del discurso está cubierto hasta cierto punto con el corpus *ADPA* (*Análisis del discurso público actual*), desarrollado en la Universidad de La Coruña, y que contiene material audiovisual en gallego, portugués y español.

También se han compilado varios corpus para la descripción de distintas variantes regionales del español dentro de España. Ejemplos de ello son el *VUA* (*Variedades urbanas andaluzas*, Universidad de Granada - Universidad de Málaga), el *ILSE* (*Investigaciones Lingüísticas del Sudeste Español*, Universidad de Almería), el *ALMECOR* (Universidad de Granada) y el *FAE-Esp Can* (*Fonética Acústica y Experimental del Español de Canarias*, Universidad de La Laguna). Otros centros de enseñanza también han mostrado interés en el estudio de la variación de la lengua en base a la edad de los hablantes, lo que ha llevado a la compilación de varios corpus para tal efecto. En cuanto al estudio de la lengua oral en base a la edad del hablante, el *Corpus López Ornat* destaca por la recogida sistemática de la producción oral de una niña durante más de dos años con el fin de investigar el proceso de gramaticalización.

Algunos corpus se han diseñado y compilado específicamente para el desarrollo, entrenamiento y evaluación de sistemas de procesamiento del lenguaje natural, especialmente de sistemas de reconocimiento de voz. Estos corpus pueden contener

⁴ Para más información sobre corpus orales del español, consultar http://liceu.uab.es/~joaquim/language_resources/spoken_res/Corp_leng_oral_esp.html

materiales diversos que van desde cifras y palabras o frases aisladas hasta textos completos con una función específica. En esta línea están el *CALLHOME Spanish Transcripts* y el *Hub-5 Spanish Transcripts*, ambos compuestos de grabaciones telefónicas. Otros corpus dignos de mención dentro del campo de las tecnologías del habla son el *EUROM.1 (European Languages Speech Database)* y el *ROARS2 (Robust Analytical Speech Recognition System)*, que ya se han completado, y el *ALBAYZÍN*, un proyecto en el que están implicadas varias universidades.

En cuanto a aplicaciones más específicas en el campo del reconocimiento del habla está *TANGORA* (IBM España), que es un proyecto basado en corpus diseñado para reconocer el habla continua para su aplicación a programas de dictado automático. Otros corpus como el *SPEECHDAT* (Universidad Politécnica de Cataluña) y el *CEUDEX* (Telefónica I+D) han sido construidos para desarrollar sistemas de reconocimiento del habla para uso telefónico. Además, los corpus *NÚMERO*, *TELÉMACO* (Universidad Politécnica de Cataluña) y *VESTEL* (Telefónica I+D) se han aplicado al reconocimiento automático de números de teléfono y órdenes relacionadas. *SPATIS* (Telefónica I+D) es otro corpus desarrollado para facilitar información de vuelos por vía telefónica. En cuanto a la disponibilidad de los corpus mencionados, el acceso al *TANGORA* está restringido para uso interno dentro de IBM, y el acceso al *CEUDEX*, el *SPATIS* y el *VESTEL* está restringido a Telefónica. El resto de los corpus mencionados están disponibles para usos no comerciales, aunque el *NÚMERO* y el *TELÉMACO* solicitan que se firme un acuerdo previo al acceso.

2.3. Tesis doctorales relacionadas con la lingüística de corpus en traducción

Al buscar tesis doctorales en el campo de la lingüística de corpus en departamentos o programas de doctorado de traducción (e interpretación) en la base de datos *TESEO* se pone de manifiesto que, hasta la fecha y por lo que refleja dicho recurso, la mayor parte del trabajo efectuado a este nivel académico en España con esta metodología se ha dado en el último lustro. Sin embargo, quizás se podría aventurar que seguramente existen trabajos en este campo de los cuales no se ha informado y que muchos otros estarán en proceso de elaboración. Otra fuente de datos donde descubrir trabajos realizados con esta metodología en traducción ha sido *TDR*.

De entre las tesis realizadas, las que más nos interesan son, por afinidad, las que han aplicado la lingüística de corpus y están relacionadas con la traducción escrita y especializada (excepto la literaria) o la terminología aplicada a la traducción. Siguiendo un orden cronológico, en la de Rodríguez Medina (1999) se observan manuales de informática traducidos para extraer conclusiones sobre los anglicismos, estudio que en este caso tiene como novedad el centrarse en los de tipo sintáctico, muchísimo menos investigados que los léxicos. Otro tema y otro género es el que trata Martínez Sánchez (1999), la traducción de los “realia” o nombres de referentes específicos de una comunidad lingüística y cultural determinada, que la autora investiga con un corpus de folletos turísticos. López Rodríguez (2000) se centra en el estudio de la tipología textual y la cohesión en la traducción del inglés al español de textos biomédicos, estudio para el que utilizó un corpus comparable y paralelo sobre oncología. La investigación de Williams White (2001) consiste en un estudio contrastivo sobre la estructura temática de artículos también del campo de la biomedicina en inglés y traducidos al español. Candel Mora (2002) trabaja con un corpus bilingüe y otro de referencia para estudiar la terminología aplicada a la traducción técnica de informes sobre la gestión de la I+D y producir una base de datos terminológica del campo. Pejović (2002) aborda el tema de las colocaciones y la fraseología, tan clave en la enseñanza de lenguas y traducción. La tesis de Sánchez-Gijón (2003) trata el proceso de documentación del traductor científico apoyándose en un corpus comparable multilingüe de textos especializados recogidos de Internet. Con un interés terminográfico Vargas Sierra (2004) se aproxima también a un lenguaje de especialidad, pero esta vez con el objetivo de elaborar un diccionario traductológico. De la Universitat Pompeu Fabra (UPF), por su excelencia en el campo de la lingüística aplicada, han salido varias tesis que incluyen el trabajo con corpus. Una de ellas es la de Oliver del Olmo (2004) que estudia de forma contrastiva un aspecto del discurso médico, la atenuación retórica; otra tesis es la de Márquez Rojas (2005), sobre los anglicismos terminológicos en español en el ámbito de la informática; la de Fernández Sánchez (2005) se centra en un género muy concreto, el folleto de cursos de idiomas para extranjeros y su análisis contrastivo con corpus. Seguiri Domínguez (2005) ofrece un riguroso y extenso trabajo en el que no sólo incluye un corpus trilingüe de seguros turísticos que ya forma parte de un corpus de turismo más extenso, el *Turicor*, sino que entra en aspectos de evaluación, catalogación, diseño y representatividad de los textos y el corpus. La tesis más reciente registrada es la de Ribas Bruguer (2006), de la UPF, en la que se estudian diferencias discursivas entre los textos jurisprudenciales catalanes y castellanos, aprovechando el estudio también para tratar problemas que plantea la alineación de textos.

Para concluir este apartado, se puede observar un creciente interés en la lingüística de corpus y el trabajo con estos recursos en una diversidad de campos dentro de España. Sin embargo, también es cierto que aún hay áreas, sobre todo dentro de los estudios sobre la traducción, donde esta metodología no ha tenido una gran repercusión, ya sea por desconocimiento o por falta de medios. Un caso así es el de la aplicación de corpus a la enseñanza de lenguas y traducción que aún no ha dado en España sino algunos casos puntuales.

2.4. Utilización de corpus en el mundo profesional: agencias de traducción

Con el fin de probar la utilidad pedagógica de los corpus para traducir hemos creído conveniente observar primero su uso real en el mundo laboral, que es para lo que se prepara a los estudiantes de traducción. En este sentido, se diseñó un sencillo cuestionario dirigido a agencias y empresas de traducción y se envió por correo electrónico a unas veinte, sobre todo a las pertenecientes a la ACT (Asociación de Empresas de Traducción). El cuestionario consta de dos partes, una para obtener información de las agencias sobre su uso de corpus y otra sobre su producción de traducciones. Centrándonos ahora sólo en la primera parte, que se puede ver más abajo (Documento 1), las preguntas tienen el fin de llegar a saber tres datos concretos:

- Si las empresas de traducción utilizan corpus.
- Si recogen textos originales y textos traducidos, textos paralelos o todo ello.
- Con qué herramientas aprovechan los textos que recogen.

CUESTIONARIO SOBRE EL USO DE CORPUS ¹ ELECTRÓNICOS EN TRADUCCIÓN DIRIGIDO A EMPRESAS O AGENCIAS DE TRADUCCIÓN	
CIUDAD:
Por favor, <u>subraye</u> sus respuestas.	
1) ¿Su empresa ha comprado el acceso a algún corpus electrónico de alguna institución (p.ej. Bank of English, British National Corpus)?	
● Sí. Especifique:	
● No	
2) ¿Cuentan con alguna herramienta informática de análisis de corpus (p.ej. WordSmith Tools, SystemQuirk, ParaConc, etc.)	
● Sí. Especifique:	
● No	
3) ¿Utilizan otros corpus on-line como el de la Real Academia Española, los accesos on-line al British National Corpus o el Bank of English, etc.?	
● Sí. Especifique:	
● No	
4) ¿Recopilan y alinean textos (texto original en inglés o español + traducción) en formato electrónico para reutilizarlos en futuros encargos?	
● Sí. Especifique qué programa/s utilizan para aprovechar los textos originales y sus traducciones:	
.....	
● No	
5) ¿Recopilan y guardan textos paralelos (textos originales en inglés y/o textos originales en español) en formato electrónico para después extraer terminología, crear bases de datos o glosarios, o simplemente para reutilizarlos en futuros encargos de un tema similar?	
● Sí. Especifique qué programa/s utilizan para aprovechar los textos paralelos:	
.....	
● No	

Documento 1. Cuestionario dirigido a empresas de traducción (1ª parte).

En octubre de 2004 se obtuvieron resultados completos de 16 empresas, sobre todo sitas en Barcelona. Los datos de las respuestas son los siguientes:

- 1) ¿Su empresa ha comprado el acceso a algún corpus electrónico de alguna institución (p.ej. Bank of English, British National Corpus)?
 - * Sí = **1**
 - * No = **15**
- 2) ¿Cuentan con alguna herramienta informática de análisis de corpus (p.ej. *WordSmith Tools, SystemQuirk, ParaConc*, etc.)
 - * Sí = **2**
 - * No = **14**
- 3) ¿Utilizan otros corpus on-line como el de la Real Academia Española, los accesos on-line al British National Corpus o el Bank of English, etc.?
 - * Sí = **6**
 - * No = **10**
- 4) ¿Recopilan y alinean textos (texto original en inglés o español + traducción) en formato electrónico para reutilizarlos en futuros encargos?
 - * Sí = **16**
 - * No = **0**
- 5) ¿Recopilan y guardan textos paralelos (textos originales en inglés y/o textos originales en español) en formato electrónico para después extraer terminología, crear bases de datos o glosarios, o simplemente para reutilizarlos en futuros encargos de un tema similar?
 - * Sí = **11**
 - * No = **5**

Documento 2. Respuestas al cuestionario dirigido a empresas de traducción (1ª parte).

Sin poder decir que estos datos son concluyentes debido a la limitación del estudio (16 agencias), sí podría decirse que apuntan una tendencia y proporcionan una idea sobre lo que ocurre en el mundo de la traducción a nivel de empresa. Así, se puede observar que prácticamente ninguna invierte en la compra de corpus aunque algunas sí utilizan los gratuitos accesibles en Internet. Por otro lado, se ve clarísimamente que las empresas sí invierten tiempo y esfuerzo en almacenar textos originales y traducidos, es decir, corpus paralelos, y, en menor medida, textos paralelos, es decir, corpus comparables. Las herramientas utilizadas para aprovechar los textos almacenados son básica y lógicamente memorias de traducción, programas que aumentan en gran medida la productividad y la calidad de las traducciones. Entre ellos destacan el *Déjà Vu* y el paquete Trados y, en menor medida, el *Transit*, e incluso se menciona la aplicación *Access*.

3. CAMPOS DE APLICACIÓN

Hasta la fecha, muchas disciplinas dentro del estudio de la lengua se han aprovechado de un enfoque basado en corpus, es decir, de una aproximación empírica que ha permitido al lingüista hacer afirmaciones que son más objetivas y están basadas en datos que ocurren de forma natural, a diferencia de los juicios basados en la intuición del lingüista y ejemplos prefabricados.

Se debe apuntar que todas las ramas de los estudios sobre la traducción han recogido los frutos de la adopción de la metodología de corpus, siendo la rama descriptiva la que se ha desarrollado más hasta el momento. Seguidamente ofrecemos una descripción y ejemplos de las áreas más destacadas dentro del estudio de la lengua que hasta ahora han adoptado un enfoque basado en corpus en mayor o menor grado, aunque también existen trabajos puntuales en áreas muy específicas como la detección de plagios (Atwell *et al.* 2003), por poner un ejemplo. Asimismo, se facilitarán referencias de estudios o aplicaciones relacionados con la traducción en aquellas áreas donde existan.

Como se podrá observar, la mayoría de los ejemplos están relacionados con la lengua inglesa, que es en la que más se ha trabajado con corpus, si bien es cierto que esta situación está cambiando gradualmente a medida que se construyen y se hacen disponibles corpus en otras lenguas. Actualmente, podemos incluso destacar el hecho de que existe un corpus de lenguas “en peligro” que incluye textos en indoeuropeo y otros, y está a cargo de un grupo de lingüistas rusos. Por último, es menester apuntar también que, obviamente, lo que sigue no es una relación exhaustiva de todo el trabajo realizado con corpus hasta la fecha; para tal fin baste referir al lector a alguna publicación de referencia (Olohan 2004; Saldanha 2008a) o a la información resultante de congresos, algunos de ellos periódicos, relacionados con esta metodología, como *Corpus Linguistics* (CL '01, '03, '05, '07), o encuentros de asociaciones como la *Association for Computational Linguistics* (ACL '02, '03, etc.) o la *British Association for Applied Linguistics* (BAAL '06, '08), entre otros. Otros medios, como las listas de distribución, también pueden ofrecer un buen panorama del trabajo que se está llevando a cabo con corpus en diversas áreas. Probablemente el mejor ejemplo es la Corpora List.

3.1. Lexicografía y descripciones de la lengua

Según Kennedy, “a major reason for compiling linguistic corpora is to provide the basis for more accurate and reliable descriptions of how languages are structured and used” (1998: 88). En este sentido, un área que se ha beneficiado considerablemente de un enfoque basado en corpus es la lexicografía en general y la producción de diccionarios en particular.

Se ha afirmado que el uso de corpus en lexicografía data ya de principios del siglo XVII (Francis 1992) y que en el siglo XVIII, Samuel Johnson, un conocido erudito inglés, ilustró los significados y usos del vocabulario inglés contenido en su *Dictionary of the English Language* con citas provenientes de “escritores de primera fila” que previamente había recogido en trozos de papel. De la misma forma, en la construcción del *Oxford English Dictionary (OED)* del siglo XIX se utilizaron citas en papel para estudiar e ilustrar el uso de las palabras, una tarea de compilación que se completó en 71 años. Así, en el pasado y en el presente los corpus han proporcionado definiciones más completas y precisas a los diccionarios que, en su concepto tradicional, se antojaban inexactos ya que dependían en gran medida de la experiencia del lexicógrafo. Corpus grandes como el del proyecto Cobuild, posteriormente convertido en un corpus monitor (*Bank of English*), han supuesto un nuevo concepto de aproximación y descripción de la lengua y han jugado un papel clave en aspectos como la identificación de neologismos, distintas acepciones y usos de elementos lingüísticos concretos y sus frecuencias relativas, de lo que se desprende que los corpus también se han utilizado en el campo de la semántica. En español ya se empieza a contar con diccionarios que también se sustentan en parte o totalmente en datos extraídos de corpus. El diccionario de la RAE es uno de los casos más destacables y útiles por su accesibilidad en Internet, pero también contamos con varios diccionarios de la editorial Vox, como el *Lema Diccionario de la lengua española* o el *Diccionario de uso del español de América y España* y, a punto de salir, un diccionario de frecuencias del español dirigido por Mark Davies, profesor de lingüística en la Brigham Young University que ha creado el *Corpus del español* para diversos propósitos.

Aún dentro del campo de los estudios sobre el léxico, se debería resaltar que los corpus se han utilizado para otros objetivos no lexicográficos, como es el caso de la extracción de palabras con frecuencias de aparición altas para fines pedagógicos. Thorndike (1921) llevó a cabo un estudio léxico-estadístico que estableció la frecuencia relativa de cada palabra en su corpus de 4,5 millones de palabras, lo cual propiciaría cambios

importantes en la enseñanza del inglés como primera y segunda lengua. Su trabajo también ejercería una gran influencia sobre el principio conocido como el “*vocabulary control*”, que subrayaba la importancia de enseñar el vocabulario más frecuente, una tendencia que se extendió entre los años 20 y 50. Otros ejemplos de estudios sobre el léxico se han centrado en la comparación de estadísticas de frecuencia entre el uso de palabras de contenido y funcionales, o el papel de las colocaciones (palabras que aparecen normalmente al lado o cerca de otras). Un ejemplo loable del último caso es el análisis de Cruden (1769) de la Biblia y la identificación de coocurrencias repetidas en determinadas palabras.

Parafraseando la cita de Kennedy al principio de este apartado, el objetivo principal de la compilación de corpus es, según el autor, mejorar las descripciones de la lengua. En este sentido, la gramática, entendida en un sentido amplio, verdaderamente se ha beneficiado de la disponibilidad de corpus. El corpus *Survey of English Usage (SEU)* (Quirk 1968) es, por ejemplo, el corpus de la era preelectrónica más importante construido específicamente para la descripción de la gramática. Y dentro de la gramática, la aproximación basada en corpus a la morfología y la sintaxis ha dado lugar a un gran número de estudios que mencionaremos brevemente. Dentro de la morfología se han llevado a cabo estudios con corpus para investigar, por ejemplo, el uso de los verbos modales y frasales, el uso distintivo del subjuntivo en inglés británico y americano, la relevancia y uso de determinadas preposiciones y conjunciones, etc. Los estudios gramaticales centrados en las oraciones han arrojado resultados sobre procesos sintácticos como la modificación del sustantivo, o se han centrado en aspectos como la condición, la causa o la negación, entre otros.

Dentro de la rama de los estudios descriptivos sobre la traducción basados en corpus se ha buscado, entre otras cosas, caracterizar la lengua de la traducción por sí misma, es decir, se han dado estudios que parten de la idea de que en los textos traducidos hay elementos comunes independientemente del par de lenguas implicadas. En este sentido, la traducción se entiende como un acto de comunicación diferente a otros y el texto traducido como “shaped by its own goals, pressures and context of production” (Baker 1996: 175), lo cual da pie a suponer que tal unicidad debe reflejarse de alguna forma en la superficie, tal y como académicos como Toury (1991) o Even-Zohar (1979) han apuntado. Esta manera de manifestarse puede venir en forma de normas o en forma de lo que se conoce como “universales de la traducción”. Éstos son patrones de comportamiento lingüístico generados por el proceso de mediación que se da en la traducción, independientemente del par de lenguas que esté implicado. Si se llegara a

tener las pruebas empíricas suficientes de que dichos rasgos universales existen, éstos permitirían “the elucidation of the nature of translated text as a communicative event” (Baker 1993a: 243), que es, desde el punto de vista de Baker, la tarea principal de los estudios sobre la traducción, una tarea que se puede llevar a cabo utilizando las técnicas de corpus y, especialmente, el tipo de corpus compuesto por textos originales en una lengua y traducciones a esa misma lengua. En su artículo de 1993, Baker esbozó varias hipótesis que podrían entenderse como universales de la traducción potenciales. Según estas hipótesis, los textos traducidos (1) tienen una tendencia hacia la explicitación, es decir, a incluir información adicional sin que esto venga motivado por la diferencia de lenguas o culturas; (2) normalmente son menos ambiguos que los textos originales; (3) son más simples y más convencionales en cuanto a gramática se refiere; (4) tienden a exagerar rasgos de la lengua meta; (5) pueden mostrar una distribución inusual de elementos léxicos, y (6) pueden mostrar lo que Baker (1996) llama “levelling out”, que se refiere a que los textos traducidos tienden a ser más homogéneos o similares entre sí que los textos originales.

A pesar de la dificultad de compilación del tipo de corpus ideal para la observación de los universales de la traducción, varios estudios ya han visto la luz e incluso han dado lugar a una conferencia internacional centrada en el tema y a una monografía resultante (Mauranen y Kujamäki 2004).

3.2. Lingüística computacional

Junto a la lexicografía, la lingüística computacional es el área que más se ha beneficiado del desarrollo de un enfoque basado en corpus. De hecho, como ya afirmaba Kennedy hace algún tiempo “computational linguistics is one of the most challenging and potentially most significant fields contributing to and making use of corpus-related research” (1998: 276).

Con el tiempo se observó que los paradigmas basados en reglas creadas a partir de la intuición del lingüista no eran suficientes para explicar el lenguaje natural sin restricciones de una forma estricta y exhaustiva. En este sentido, por lo general la lingüística computacional ha hecho uso de los corpus para extraer principios probabilísticos con el fin de contribuir a estudios en áreas como el análisis morfológico y sintáctico, la desambiguación de significados, la resolución de anáforas, etc. Entre éstas, el etiquetado morfológico junto con cuestiones de semántica y análisis sintáctico son cruciales para aplicaciones de procesamiento del lenguaje natural tales como el

reconocimiento de voz, la traducción automática, la extracción de información o el reconocimiento de términos, que se han beneficiado ampliamente de los cálculos estadísticos realizados a partir de corpus. Sin embargo, no deberían ignorarse las dificultades inherentes al procesamiento del lenguaje natural, tanto si está basado en reglas como en principios probabilísticos. Los aspectos pragmáticos, discursivos o semánticos de la lengua ponen límites a la eficacia de los ordenadores y la lingüística computacional, por no mencionar otros aspectos de la lengua como son la elipsis, la anáfora y las anomalías lingüísticas que pueden resultar muy difíciles de tratar para un ordenador. Allá donde el trabajo es más mecánico, valga la expresión, la tecnología sí ha sabido sacar el máximo provecho a los corpus, como es el caso de la construcción o alimentación de memorias de traducción que son, al fin y al cabo, corpus paralelos, como se da en el trabajo descrito en Jacob *et al.* (2003).

3.3. Investigación del habla oral

Según McEnery y Wilson (1996), los corpus orales son importantes para la investigación del habla en tanto que normalmente proporcionan una muestra amplia de habla en términos de variables como la edad de los hablantes, el género y la clase, y en diversos géneros o tipos de actividades como conversaciones, lectura de noticias, lectura de poesía, etc.

Básicamente hasta ahora se han dado tres tipos de trabajo con corpus en la investigación del habla: (1) sobre la naturaleza de la prosodia, (2) sobre la transcripción prosódica, y (3) sobre el diseño de tipologías textuales desde una perspectiva prosódica. Un ejemplo de un corpus anotado prosódicamente que ha sido diseñado y utilizado básicamente para la investigación en prosodia es el *Lancaster/IBM Spoken English Corpus (SEC)* (Knowles, Taylor y Williams 1992). En cuanto a la producción de corpus orales también hay que reconocer los esfuerzos comerciales del Linguistic Data Consortium (LDC) que cuenta con corpus en una diversidad de lenguas.

3.4. Pragmática y análisis del discurso

La investigación basada en corpus en los campos de la pragmática y el análisis del discurso depende de aspectos contextuales como que, por ejemplo, el corpus ofrezca información extralingüística. En ocasiones, este hecho dificulta las investigaciones, de lo que nos da cuenta Magalhães (2001) cuando comenta las dificultades encontradas a

la hora de recabar datos sobre los traductores de un diario brasileño con el fin de poder llevar a cabo estudios discursivo-textuales. En la misma publicación, y a caballo entre el análisis del discurso y la literatura, Pagano (2001b) también presenta el uso de corpus, un corpus paralelo de novelas policíacas en este caso, dentro de los estudios historiográficos en traducción. En su centro de Brasil, Magalhães y Pagano trabajan en análisis del discurso basándose en la gramática sistémica funcional de Halliday y utilizando corpus creados expresamente para facilitar esta investigación (*CORPRAT – Corpus Processual para Análises Tradutórias*; *CORDIAL – Corpus Discursivo para Análises de Língua e Literatura*), de donde ya han surgido varios estudios (Mauri 2006; Minelli y Pagano 2006; Jesús 2006, entre otros).

De la pragmática existen ejemplos en áreas relacionadas con el apartado anterior, es decir, la lingüística computacional, como por ejemplo el procesamiento automático del diálogo. No abundan los estudios sobre el discurso oral llevados a cabo con corpus, siendo el *London/Lund Corpus (LLC)* el más importante y utilizado. Stenström (1984, 1987), por ejemplo, lo utilizó para investigar el uso de elementos discursivos en inglés como “well”, “sort of” y “you know”, o la función de expresiones que denotan seguimiento en una conversación como “right” y “all right”. Aijmer (1984), por otro lado, se centró en el estudio de rodeos y atenuaciones como “*sort of*” y “*kind of*”, que son elementos que ayudan a suavizar un segmento en un discurso. Otra referencia a tener en cuenta sería la de Zorzi (2001) y su trabajo con corpus en el aula para enseñar el uso y traducción de marcadores del discurso en italiano. Finalmente, el reciente libro de Adolphs (2008) se centra en la investigación de funciones pragmáticas, como las relacionadas con los actos de habla, en el discurso oral a través de diversos corpus orales entre los que se encuentra el *Cambridge and Nottingham Corpus of Discourse in English (CANCODE)*.

Dentro del análisis del discurso hay varios ejemplos de investigación basada en corpus. Ya en 1996, Stubbs muestra cómo utilizar corpus de millones de palabras para desvelar patrones relevantes de uso de la lengua en cuanto a colocaciones, modalidad, transitividad, y cómo estos patrones pueden expresar actitudes, presuposiciones y puntos de vista. Una década después y con una perspectiva similar, Baker (2006) examina diversas formas de aproximarse al análisis del discurso utilizando técnicas basadas en la lingüística de corpus, tales como las colocaciones, las concordancias, la dispersión, etc. Desde un punto de vista más teórico, Hoey (2007) se centra en la lengua como un fenómeno social y, a través del uso de corpus, trae a la superficie las actitudes y creencias expresadas por los miembros de una comunidad discursiva.

Por último, y en lo que concierne específicamente a la lingüística de corpus en la investigación de la ideología y, en este caso, también la traducción, queremos comentar un ejemplo del estudio de elementos ideológicamente marcados en textos escritos, que son los trabajos presentados por la misma autora de esta tesis (Rodríguez Inés 1999, 2000). En éstos, la parte aplicada se centra en la observación de elementos con una potencial carga ideológica (“Castilian”, “gay”, “queer”, etc.) en un corpus sólo de traducciones al inglés (*Translational English Corpus - TEC*) y en un corpus de inglés original (*Cobuild Corpus*). La posibilidad de utilizar otro término no marcado en vez de los citados en los textos observados despierta la hipótesis de que existe una motivación detrás de dichas elecciones. En el caso de “gay” y “queer”, utilizados en las traducciones al inglés con sus acepciones más obsoletas, se sospecha que detrás de su uso hay una agenda política del traductor, Lawrence Venuti. En el caso del abundante uso de “Castilian” en los textos traducidos al inglés frente a su casi inexistente aparición en inglés original hace mirar a las obras de donde han salido estos casos y ver que provienen sobre todo de un autor catalán, J. Goytisolo, lo que nos lleva a pensar que el traductor bien se dejó llevar por la presencia de la palabra “castellano” en vez de “español” en las obras originales, bien quiso reflejar una diferencia o darle un gusto extranjerizante a su traducción.

3.5. Sociolingüística

Como la pragmática y el análisis del discurso, el análisis sociolingüístico requiere que se recoja y codifique información extralingüística, lo que no es una práctica aún demasiado extendida. Además, la selección o muestreo de los textos que se lleva a cabo para la compilación de corpus no suele estar motivada sociolingüísticamente, sino más bien textualmente o al azar, lo que dificulta el hecho de que un corpus sea representativo de fenómenos sociolingüísticos. En cualquier caso, existen algunos ejemplos de estudios de este tipo basados en corpus, de los cuales la mayoría están relacionados con el género. Un ejemplo es el estudio de Kjellmer (1986), quien utilizó los corpus Brown y LOB para examinar los prejuicios masculinos en inglés americano y británico observando los casos de elementos léxicos como *man/men* y *woman/women*. Otro es el estudio de Fauth y Schmid (2003) en el que llegaron a extraer de un corpus oral rasgos lingüísticos característicos en conversaciones mantenidas entre personas del mismo sexo para después contrastarlas con las del sexo opuesto, permitiéndoles llegar a una metodología para la detección del género de un hablante en una conversación. También con un corpus oral, Beeching (2003) estudió la variación

sincrónica y diacrónica del francés. Finalmente, Saldanha (2003) destaca por su uso de un corpus de textos traducidos para investigar algunos rasgos lingüísticos relacionados con el género.

3.6. Lingüística histórica

La lingüística histórica también se ha beneficiado del acceso a datos facilitados por corpus diacrónicos, si bien este tipo de corpus debe tratarse con precaución. McEnery y Wilson (1996) han advertido de que los corpus históricos se han construido normalmente sin seguir unos criterios de diseño estrictos, a veces ni siquiera con el objetivo de ser representativos. En la misma línea, Rissanen (1989) aconseja no fiarse demasiado de este tipo de corpus contruidos de tal forma a la hora de entrenar a investigadores o de extraer conclusiones generalizables para todo el periodo de lengua tratado. Asimismo, también llama la atención sobre la fiabilidad estadística de los corpus históricos, sobre todo cuando se deberían tener en cuenta muchas variables como las diferencias entre géneros textuales y su estatus o entre periodos, clases sociales y niveles de alfabetización.

Hay un gran número de corpus históricos que, juntos, cubren todos los periodos de la lengua inglesa desde sus comienzos hasta el siglo XX en diversos géneros. Los corpus históricos de inglés más conocidos son el *Helsinki Corpus*, que cubre el periodo entre el año 850 DC y el 1710, y su corpus complementario, *ARCHER*, que empieza donde acaba el anterior y sigue hasta la actualidad. Sumados a éstos, existen además muchos otros corpus contruidos con el objetivo de representar un periodo de tiempo en particular y un género (p.ej. *Corpus of Early English Correspondence*; *Century of Prose Corpus*), o un periodo de tiempo y una variedad de lengua (p.ej. *Helsinki Corpus of Older Scots*), etc. El corpus histórico más importante con el que contamos en español es el *Corpus diacrónico del español (CORDE)*, que en estos momentos cubre todas las épocas y lugares en que se habló español desde los inicios del idioma hasta el año 1975.

Los estudios basados en corpus en lingüística histórica han proporcionado resultados en una diversidad de temas relacionados como el cambio histórico y el uso de la lengua inglesa en hombres y mujeres (Biber y Burges 2000) o las diferencias de género en la evolución del inglés estándar (Nevalainen 2000), para dar sólo algunos ejemplos. En la aplicación de corpus a la enseñanza de lingüística histórica (Curzan 2000) de nuevo tenemos que mencionar a Mark Davies, que ha enseñado historia de la lengua española con un corpus histórico creado por él mismo.

3.7. Dialectología y variación lingüística

En el siglo XIX encontramos ya ejemplos de trabajo basado en corpus sobre la variación léxica en *The English Dialect Dictionary* (Wright 1898-1905) y *The Existing Phonology of English Dialects* (Ellis 1889). Más recientemente, la construcción de varios corpus que contienen material de diversas variantes regionales del inglés ha dado lugar a una serie de estudios comparativos, así como también ha servido como campo de pruebas para dos teorías sobre la variación de la lengua: (1) la hipótesis del *common core* (Quirk *et al.* 1985), que sugiere que todas las variantes del inglés tienen propiedades fundamentales en común y que difieren en lo cuantitativo más que en lo cualitativo, y (2) la sugerencia de Braj Kachru de que las variantes nacionales del inglés constituyen muchos y únicos “ingleses”. En este sentido, parece que las comparaciones de las variantes regionales del inglés realizadas con los corpus Brown y LOB, y los corpus Kolhapur del inglés de la India, el Wellington de Nueva Zelanda y el Macquarie de Australia, han venido a apoyar la primera hipótesis al haberse encontrado menos diferencias de las que se había esperado. Estos últimos corpus pertenecen al proyecto conocido como el *International Corpus of English* (ICE), que se compone de subcorpus de distintas variantes geográficas del inglés en el que trabajan 15 equipos (otros subcorpus ya terminados son el componente de Singapur o el de Filipinas).

En cuanto al español, una vez más, Mark Davies destaca con la impartición de cursos basados en corpus, en este caso sobre variación lingüística. El libro de Parodi (2007) recoge contribuciones también sobre este tema, centrándose en aspectos diferenciales de registro o modo en variantes del español de Sudamérica.

3.8. Estilística y estudios literarios

La estilística en general y la basada en corpus se han centrado normalmente en el estudio de autores individuales, lo que necesita de que se hagan comparaciones dentro del propio trabajo del autor, entre su trabajo y el de otros, o entre su trabajo y las normas de la lengua para extraer el idiolecto de dicho autor. Como ya hemos mencionado previamente (cfr. supra 3.1.), Cruden (1796) llevó a cabo un análisis de la Biblia, de la que extrajo listas y concordancias de palabras con el objetivo de mostrar si varias partes del libro mantenían la consistencia entre sí. Más adelante se han extraído concordancias de autores como Shakespeare para fines diversos y se han realizado análisis de autores en particular como el poeta portugués Manuel Alegre y su trabajo,

que ha sido posible gracias a software específicamente desarrollado para el manejo de corpus, en este caso, el programa *Hyperbase* (Brunet 1989). Más reciente es el trabajo que se está llevando a cabo en la Lancaster University, donde académicos centrados en la estilística como, por ejemplo, Semino y Short (2004), destacan por la envergadura de sus estudios sobre la representación del discurso oral, el escrito y el pensamiento en textos narrativos. Estas autoras, gracias al uso de un corpus creado y anotado para sus propósitos, pudieron contrastar sus resultados con modelos teóricos ya existentes y observar múltiples correlaciones entre las funciones retóricas y las elecciones estilísticas. El interés en combinar los estudios literarios y estilísticos con la metodología de corpus va en aumento, como se puede apreciar en algunos eventos (p.ej. Taller “Corpus approaches to the language of literature”, previo a la conferencia *Corpus Linguistics* 2005), y existen varios corpus con contenido exclusivamente literario, como el *Literature Online English Corpus (LION)* en Glasgow o el *Project Gutenberg*, que funciona como repositorio de textos.

Aparte de los autores originales, los traductores también han sido objeto de estudio mediante la metodología de corpus (Baker 2000; Saldanha 2008b) con el fin de extraer las características de su estilo, si se da el caso.

Dentro de la estilística los corpus también se han aplicado al estudio de la variación entre géneros textuales y canales. En cuanto al género, Simpson (2000), por ejemplo, se ha centrado en comparaciones transversales de disciplinas en un corpus oral de inglés académico con el objetivo de extraer las diferencias lingüísticas en la lengua oral en géneros de disciplinas diferentes. Sin embargo, la atención se ha centrado sobre todo en el estudio de las diferencias entre la lengua escrita y oral, lo que ha propiciado estudios sobre, por ejemplo, la construcción de la fórmula de causa-efecto (Altenberg 1984) o las estrategias de negación (Tottie 1991).

Aparte de utilizarse para los propósitos descritos, los corpus también se han utilizado para poner en duda tipologías textuales existentes. El ejemplo más destacable es sin duda la tipología textual multidimensional de Biber (1988) que está basada en los rasgos lingüísticos que se encuentran en los textos.

3.9. Psicolingüística y psicología social

Tanto en la psicolingüística como en la psicología social los corpus han representado una importante fuente de datos naturales. En lo que a estudios basados en corpus se

refiere, la psicología social ha utilizado estos recursos para el estudio de las explicaciones, es decir, cómo y por qué la gente intenta explicar cosas y cómo consideran su entorno. Para estudiar las emociones en la lengua y en distintos registros, es decir, cómo las personas hablan (en inglés) de sus emociones y de las de otros, Bednarek (2008) utiliza un subcorpus del *BNC* y produce un trabajo que bebe de disciplinas como la lingüística cognitiva.

Mientras tanto, la psicología ha hecho un gran uso de los corpus como fuente de datos de la que extraer materiales para experimentos de laboratorio con el fin de estudiar los errores de habla en conversación natural en inglés, analizar las patologías de la lengua e investigar el desarrollo de la lengua. Un ejemplo de corpus para fines relacionados con la psicología es el *NEXING 1.0* que incluye transcripciones de datos orales extraídos durante un experimento de razonamiento silogístico.

3.10. Adquisición de la lengua

Como ya se ha mencionado anteriormente (cfr. supra 1.3.1.), la investigación en la adquisición y desarrollo de la lengua no puede llevarse a cabo sólo a través de la introspección, sino que requiere muestras reales de lengua producidas por niños. Ejemplos de corpus compilados específicamente para la investigación en adquisición de la lengua son el *Child Language Data Exchange System* (CHILDES), que contiene muchos tipos de datos diferentes relevantes al tema, y el *Polytechnic of Wales* (POW), que se ha compilado para el propósito concreto de investigar la adquisición y el desarrollo de las estructuras sintáctico-semánticas en niños.

Por otro lado, no sólo se han creado corpus para el estudio de la adquisición de la primera lengua, sino también de la adquisición y el aprendizaje de segundas lenguas, como es el caso de los corpus de estudiantes (cfr. infra 3.11.3.5.), que se mencionan más adelante.

3.11. Didáctica de lenguas

“There is every reason to believe that language corpora will have a role of growing importance in teaching” (Leech 1997: 1).

La aplicación de corpus a la didáctica de lenguas tiene un interés especial en el marco de esta tesis ya que precede en el tiempo y supera con creces en producción bibliográfica a la aplicación de corpus a la didáctica de la traducción, constituyendo un saber de fondo que no se puede ignorar. Asimismo, entre los autores que se han dedicado al uso de corpus en la enseñanza de lenguas, algunos también han probado esta metodología en clases de traducción o han al menos recomendado su uso y aportado ideas.

3.11.1. Antecedentes

Como se ha mencionado antes (cfr. supra 1.3.), ya se encuentran datos del uso de corpus con fines pedagógicos en la primera mitad del siglo XX en el trabajo de Thorndike (1921) y el “vocabulary control movement” (Fries y Traver 1940). En este sentido, dentro de la aproximación a la didáctica de la lengua basada en corpus se ha prestado mucha atención a la distribución y frecuencia del vocabulario y al estudio de lo probable y lo improbable. Un ejemplo de ello es el trabajo de West (1953) consistente en la creación de una lista de las palabras más frecuentes del inglés y que fue ampliamente utilizada en el diseño de temarios para enseñar inglés como lengua extranjera. En la misma línea, Willis y Willis (1989) también desarrollaron materiales didácticos basándose en el corpus Cobuild y en un “lexical syllabus” (Sinclair y Renouf 1988; Willis 1990), es decir, una propuesta que recomienda al estudiante de una lengua centrarse en las palabras más frecuentes, sus patrones de uso más importantes y sus combinaciones más típicas. Muchos son los que han apoyado la idea de que las estadísticas de frecuencia en una lengua, si bien no son el único criterio relevante a la hora de decidir prioridades de enseñanza, sí representan unos datos muy importantes que se deben tener en cuenta (Leech 1997; Aston 2000a, 2001b).

“Whatever the imperfections of the simple equation ‘most frequent’ = ‘most important to learn’, it is difficult to deny that the frequency information becoming available from corpora has an important empirical input to language learning materials” (Leech 1997: 16).

Otros, por su parte, han recomendado cautela a la hora de enseñar solamente aquellas palabras de mayor frecuencia ya que éstas pueden no ser siempre las más interesantes que aprender o investigar.

En relación con este tema se encuentra el debate mantenido a lo largo de algunos años ya entre lo que es descriptivamente, e incluso, prescriptivamente adecuado o

científicamente interesante en el estudio de una lengua, y lo que es pedagógicamente relevante en su enseñanza (Kennedy 1992). Partiendo de una perspectiva lingüística, se piensa que una mejor descripción de la lengua gracias a la investigación con corpus tendrá un efecto positivo en la pedagogía a través de mejores materiales de referencia y temarios. Si se parte del contexto pedagógico la creencia es que la base fundamental está en los procesos de enseñanza y aprendizaje, y que los descubrimientos en lingüística habrán de evaluarse en función de estos procesos. Desde esta perspectiva, llevar los métodos de la lingüística de corpus a la pedagogía no tiene un éxito asegurado, y el hecho de que las concordancias hayan resultado muy útiles a los lingüistas a la hora de formular generalizaciones no garantiza que esto pueda aplicarse en el aula (Aston 1995). En este sentido, Aston (2001b) se posiciona y sostiene que la pedagogía tiene mucho que ganar de una mejor descripción de la lengua si ésta está basada en corpus ya que así se asegura que proviene de una observación sistemática y no de una intuición sobre lo que es posible o probable. Asimismo, deja claro que la aproximación descriptiva y la pedagógica no tienen por qué tener los mismos métodos e intereses, empezando por el hecho de que los estudiantes no son lingüistas que analizan la lengua por su propio interés, sino en tanto que les sirva de ayuda para poder comunicarse.

Un punto en el que sí coinciden lingüística de corpus y pedagogía de la lengua es en ver la competencia y el uso de la lengua como algo idiomático y esquemático, observando que tanto en la producción como en la interpretación del discurso se hace uso de bloques prefabricados y memorizados (Aston 1995). Este hecho, comprobado empíricamente gracias a la lingüística de corpus moderna, ha revolucionado no sólo el modo de trabajar con la lengua, sino también las propias unidades que la conforman.

“[...] evidence of the ‘patterned’ quality of language performance [...] has not only accentuated the mistrust of intuition and elicitation as sources of linguistic insights, but has also contributed to shaping our conception of language as a highly idiomatic system where memory plays a central role” (Bernardini 2000b: 106).

Otras áreas como la psicología, la psicolingüística o la antropología coinciden en aceptar que una gran parte de la actividad lingüística está relacionada con el manejo de grupos de palabras más o menos fijos, lo que también tiene implicaciones directas para la traducción y su enseñanza. Según Sinclair, lo que subyace a este fenómeno de patrones recurrentes es el “idiom principle”, que postula que “a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices [...]” (Sinclair 1991: 110 y ss). Este principio no sólo se aplica a las

expresiones idiomáticas, refranes, frases hechas y colocaciones obvias, sino también a frases que normalmente no se consideran idiomáticas. Otro punto destacable en el trabajo basado en corpus de Sinclair es la afirmación de que forma y significado están estrechamente ligados en el sentido de que “each meaning [of a word] can be associated with a distinctive formal patterning” (ídem: 7). Como consecuencia de su trabajo con corpus y las pruebas obtenidas, Sinclair ha redefinido y ampliado la unidad de significado teniendo en cuenta tanto las asociaciones contextuales como las funciones pragmáticas de las palabras. Dicho de otro modo, el análisis de una unidad de significado se empieza por la palabra-nodo y se expande hacia los lados buscando regularidades léxicas (colocaciones), gramaticales (patrones gramaticales o *colligation*), de familia de palabras (preferencia semántica) y de connotación (prosodia semántica). Un ejemplo del propio Sinclair ilustra estos conceptos: respecto a la expresión inglesa “naked eye”, que podría traducirse al español como “a simple vista”, el análisis llevado a cabo con un corpus de inglés general de grandes dimensiones revela que “naked eye” es una colocación en sí misma que presenta una regularidad gramatical en el sentido de que suele venir precedida de las categorías adjetivo + preposición (p.ej. “invisible to the naked eye”). Asimismo, esta expresión tiene una preferencia semántica por el campo de la visibilidad que se manifiesta en palabras circundantes como “visible”, “invisible”, “discernible”, “unnoticed”, etc. y una prosodia semántica o discursiva que hace referencia a algo dificultoso o fallido, lo que se refleja en prefijos como “in-”, “un-” y otras palabras como “barely”, “almost”, “just”, etc.⁵

Hoey, otro reputado lingüista coetáneo de Sinclair, también ha trabajado en esta línea utilizando corpus. Hoey, sin embargo, habla de *priming* o preferencia en lo que se refiere a todo aquello que rodea y se asocia con una palabra o expresión. En sus propias palabras, “By primed, I mean that as the word is learnt through encounters with it in speech and writing, it is loaded with the cumulative effects of those encounters such that it is part of our knowledge of the word that it co-occurs with other words” (Hoey 2004). Este autor distingue entre varios tipos de *lexical priming*, a saber:

- *Collocational priming* o lo que llamaremos “preferencia colocacional”, que es la preferencia que una palabra o expresión muestra por la presencia de otras palabras determinadas a su alrededor. Es un rasgo que puede cambiar a lo largo del tiempo (p.ej. la palabra “pesetas”, hoy en día, tiene una preferencia

⁵ La autora ha observado que la expresión equivalente a “naked eye” en portugués, “olho nu”, muestra una prosodia semántica opuesta. Es decir, introducida la expresión en un corpus extenso de portugués general (<http://www.corpusdoportugues.org/>), se observa que las palabras que suelen acompañar a “(a) olho nu” reflejan frecuentemente la capacidad de ver, y no la dificultad (p.ej. *até mesmo, mas ainda, visível, visíveis, observam-se*, etc.). Hecha la comprobación en el CREA de la expresión “a simple vista” no se puede afirmar que los resultados sean concluyentes para el español peninsular.

colocacional por el adjetivo “antiguas”), y también es sensible al campo temático donde se encuentre la palabra (p.ej. el término “interés”, en el campo temático de los préstamos, tiene una preferencia colocacional por adjetivos como “fijo”, “variable”, etc.).

- *Colligation* o lo que llamaremos “coligación”, que Hoey en la misma publicación define como:
 - o El acompañamiento que una palabra lleva o evita en cuanto a categoría gramatical.
 - o Las funciones gramaticales que una palabra o grupo de palabras prefiere o evita.
 - o La posición que una palabra prefiere o evita ocupar en una secuencia, que es lo que el propio autor llama “textual colligation”.

Las implicaciones que lo expuesto tienen para la enseñanza de la lengua se plasman en que “when we acquire a lexical item, it becomes primed for collocation, grammatical category, semantic associations and colligation, and it is not properly acquired unless it has all this priming” (ídem), lo que conduce otra vez a esa nueva visión de la lengua construida con el uso y a base de unidades amplias e interrelacionadas que se pueden abordar con la lingüística de corpus. El conocimiento de todas estas preferencias en una lengua o la lengua de un campo, entre otras cosas, significará parte de su dominio.

“Mastery of a subject is, again amongst other things, mastery of the collocations, colligations and semantic associations of the vocabulary of the discipline – mastery, in fact, of the domain-specific and genre-specific primings” (ídem).

En relación con la naturaleza repetitiva de la lengua, un punto importante en el trabajo de Aston (1995) es la insistencia en la necesidad de exponer al estudiante a esta repetición. Sin embargo, asumiendo que el aprendizaje del estudiante consiste en una aproximación progresiva hacia la lengua extranjera siguiendo procesos de reestructuración del conocimiento, Aston sugiere que igual de importante que es la exposición a la repetición lo es la exposición a la variación, para que esta reestructuración del conocimiento pueda tener lugar. Esta idea está estrechamente relacionada con el concepto de esquemas de Skehan (1998), definidos por Widdowson como “cognitive constructs that allow for the organisation of information in long-term memory and which provide a basis for prediction” (Widdowson 1983: 34). Según Bernardini (2000), actividades en las que se ayuda al estudiante a observar regularidades e irregularidades en el uso de la lengua, así como a proponer y probar

hipótesis y relacionar esquemas formales locales con esquemas culturales más generales, propician una reestructuración del conocimiento. Este concepto se ha aplicado también a la adquisición de la competencia traductora en el modelo dinámico del grupo PACTE basado en la reestructuración y desarrollo, en este caso de las subcompetencias que forman la competencia traductora.

3.11.2. Aproximaciones al uso de corpus

En otro plano, según Aston (1997b), hasta la fecha ha habido dos formas clave de aproximarse al uso de corpus en la enseñanza de lenguas:

- 1) El enfoque “Cobuild”, en la que el corpus es un recurso para el diseñador de materiales didácticos y para el profesor ya que permite la creación de materiales de referencia y el diseño de temarios y materiales didácticos.
- 2) El enfoque “data-driven learning” o aprendizaje guiado por los datos, en la que el corpus es un recurso para el estudiante, bien en la forma de concordancias impresas o como concordancias que él mismo extrae del corpus con ayuda de software.

Bernardini nos dice en su libro (2000b: 132) que se han tomado estas dos aproximaciones para identificar dos posturas o actitudes de cara al aprendizaje de la lengua:

- 1) Posicionamiento débil, que reivindica que los materiales de referencia basados en corpus pueden ofrecer a estudiantes y profesores descripciones de la lengua más adecuadas y útiles que los materiales ya existentes. En este sentido, el análisis de corpus ha venido a demostrar que materiales pedagógicos y de referencia no basados en corpus son descriptivamente inexactos (gramáticas, diccionarios, etc.).
- 2) Posicionamiento fuerte, que defiende que los corpus, las concordancias y su software deberían llevarse al aula y utilizarse como fuentes de actividades de aprendizaje.

Johns (1991b), considerado como el padre del “data-driven learning” aboga, junto a otros autores como Leech (1997), Gavioli (1997, 2001) o Bernardini (2000), por el segundo posicionamiento y el acceso directo del estudiante al corpus para que éste pueda participar en su aprendizaje de una forma también más directa.

“Data-driven learning” (Johns 1991b) es, pues, una aproximación al aprendizaje de la lengua basada en corpus en que dicho aprendizaje viene a estar dirigido por los datos. Se fundamenta en un aprendizaje por descubrimiento (Widdowson 1990) en el que la tarea del estudiante es descubrir la lengua extranjera y la tarea del profesor es proporcionar un contexto en el que el estudiante pueda desarrollar estrategias para descubrir o “aprender a aprender”. El ordenador es un tipo especial de informante que da pie a un aprendizaje que a su vez facilita la reestructuración del conocimiento.

“The informant, [...] is passive - and silent - until a question (answer unknown) is asked by the learner. The informant responds to that question as best he (or she) can: and the learner then tries to make sense of that response (possibly asking other questions in order to do so) and to integrate it with what is already known” (Johns 1991b: 1).

“[...] the language-learner is also, essentially, a research worker whose learning needs to be driven by access to linguistic data – hence the term “data-driven learning” (DDL) to describe the approach” (ídem: 2).

Si bien es cierto que el “data-driven learning” describe procedimientos deductivos en los que el estudiante ha de ir de las generalizaciones sobre la lengua a las concordancias, el aprendizaje es, sobre todo, inductivo, con actividades que implican la identificación, clasificación y generalización de patrones de la lengua partiendo de las concordancias. Flowerdew (1996) explica que la propuesta que hace Johns va incluso un paso más allá de lo que normalmente se considera inductivo.

“Noting that the inductive learning of the “examples first, rules second” type is really a case of “rules hiding” rather than the encouragement of truly inductive thinking—because the teacher knows all along where the learner is being led—Johns proposes a role for the learner in which the teacher does not in fact know where the analysis will lead and in which the linguistic data is primary. In this “data-driven” learning approach, the initial problem, instead of being teacher-initiated, is likely to come from a student query” (Flowerdew 1996: 108).

Aquí, la mediación del profesor entre el estudiante y el corpus no se considera necesaria, tomando el profesor un papel de mero facilitador o asesor (Leech 1997; Aston 2001b; Tognini-Bonelli 2001) y el estudiante un papel de investigador independiente (Leech 1997; Gavioli 1997, 2001) que formula y prueba sus propias hipótesis sobre el uso de la lengua. Sin embargo, para que esto sea posible, es necesario dotar al estudiante de conocimientos y habilidades técnicas y metodológicas para que

pueda utilizar corpus en sus investigaciones de una forma eficaz y cada vez más autónoma.

“Proposing corpora and the concordancer as instruments for autonomous research involves more than simply teaching learners a new technology. In the first place, they need to acquire appropriate methods of analysis. They need to appreciate that corpora and concordances provide samples rather than examples, and that analysis should aim at identifying recurrences and inferring patterns which appear in some way typical of certain contexts” (Gavioli 2001: 129).

Concretamente, respecto a la metodología que acompaña al uso de corpus en el aprendizaje de una lengua, Aston (2001b) sugiere que los estudiantes necesitan ser capaces de identificar problemas, seleccionar corpus adecuados para interrogar, diseñar búsquedas de calidad y ajustadas a los objetivos y, finalmente, manipular e interpretar los datos para obtener respuestas fiables y utilizables. Tognini-Bonelli (2001), por su parte, propone enseñar a los estudiantes a leer verticalmente y de forma fragmentada y a centrarse en aspectos recurrentes como lo son las colocaciones, las categorías gramaticales y otros aspectos de la lengua a nivel semántico y funcional. Sin embargo, además de enseñar al estudiante cómo puede un corpus mostrar regularidades de uso, Aston (1997a) llama también la atención sobre la necesidad de que el estudiante entienda que dichas regularidades son relevantes para su aprendizaje y que no necesariamente se corresponden con otras generalizaciones contenidas en diccionarios y gramáticas a las que el estudiante estará más acostumbrado.

“By making learners aware of such regularities and of their importance in language use and learning, and by showing them how they can be highlighted in corpora through the use of retrieval, selection and sorting procedures, we also move in the direction of independent use of corpora as resources for autonomous learning” (Aston 1997a: 61).

El potencial del aprendizaje autónomo es uno de los mayores atractivos del uso de corpus en la pedagogía de lenguas (Aston 2001b), sin hablar de las posibilidades que abre para la enseñanza a distancia tanto de lenguas como de lingüística o traducción. Estas posibilidades aún no están bien exploradas en la práctica en la enseñanza de la traducción, en la que entraremos en detalle en el capítulo 4., pero sí es cierto que existen ideas al respecto en las que estamos trabajando (Rodríguez Inés 2004). En la citada comunicación, la autora defiende que una combinación de uso de recursos electrónicos y una metodología de aprendizaje por descubrimiento son los ingredientes idóneos para enseñar a traducir con corpus a distancia, dado que se dan las siguientes circunstancias:

- Acceso electrónico a textos y corpus.
- Acceso en línea a programas de análisis de corpus.
- Posibilidad de trabajo autónomo (descubrimiento).
- Posibilidad de interacción entre alumnos (discusión).
- Posibilidad de auto-evaluación del proceso.
- Posibilidad de repetición del proceso si el resultado no es satisfactorio (nueva selección de textos/corpus, nuevas búsquedas, etc.).

Aspectos que favorecen el uso de corpus en la enseñanza a distancia (Rodríguez Inés 2004).

La autonomía es, además, junto a la motivación, la autenticidad y el trabajo por tareas, uno de los aspectos principales dentro del enfoque comunicativo en la enseñanza de lenguas. El enfoque comunicativo, centrado en el estudiante, el significado y el proceso, ha sido estudiado por Bernardini (2000) en relación con el uso de corpus para la enseñanza de una lengua extranjera (el inglés) a estudiantes avanzados de traducción. La puesta en práctica de un ejercicio de final abierto basado en el uso directo y autónomo de un corpus grande (*BNC*) por parte de los estudiantes y siguiendo una pedagogía del descubrimiento, ha llevado a la autora a concluir, entre otras cosas, que el aprendizaje de lenguas se beneficiaría de la introducción de tareas que incorporasen ciertos elementos que, en este caso, el corpus puede facilitar.

“Available evidence suggests that language learning would profit from the introduction of tasks incorporating consciousness-raising reasoning-gap activities, which require the search for a solution to authentic problems, with public and private communication activities” (Bernardini 2000b: 76).

En otras palabras, la autora argumenta cómo se puede utilizar un corpus como fuente de actividades de aprendizaje exploratorio y como plataforma de navegación o búsqueda y cómo, metafóricamente, el estudiante puede tomar el papel de un viajero sin rumbo fijo que va adquiriendo conocimiento y habilidades a medida que avanza en su viaje, mientras que el profesor sirve de guía experto ayudando al estudiante a recorrer el camino, sobre todo en sus primeros pasos. Bernardini defiende el uso de corpus grandes generales como el *BNC* o el *Bank of English* en la enseñanza de lenguas por la riqueza de posibilidades que ofrecen. Según Bernardini, los corpus grandes generales no son comparables a los pequeños y especializados ya que son completamente diferentes en cantidad y calidad, es decir, no solamente en tamaño, sino en variedad de géneros incluidos, tipos de análisis que permiten llevar a cabo y filosofía subyacente.

“The richness and range of the information available, both linguistic and cultural; the flexibility of the instrument, which lends itself to all sorts of investigations, learning styles and personal interests; the picture of the language it reveals, with patterns observable at various levels of specificity/generality and from various viewpoints (lexico-syntactic, semantic, and pragmatic); all these require us to treat large corpora as very different from small ones, calling for different kinds of training and satisfying different kinds of pedagogic priorities” (Bernardini 2001: 247).

Sin embargo, la situación real es que la mayor parte del trabajo con corpus que se ha llevado a cabo hasta la fecha en enseñanza de lenguas ha sido con corpus pequeños, especializados y, en la mayoría de los casos, contruidos ad hoc para la ocasión. Aston (1997b), Tribble (1997) y Gavioli (2001) son sólo algunos de los autores que abogan explícitamente por el uso de corpus pequeños especializados. Aston, por su parte, comenta las ventajas de estos corpus frente a corpus grandes generales desde un punto de vista “data-driven learning”, ventajas que podrían hacerse extensivas a la enseñanza de la traducción. Según Aston (1997b: 55 y ss), los corpus pequeños especializados:

- 1) Son más fáciles de manejar.
- 2) Son analizables en su totalidad.
- 3) Es más fácil familiarizarse con ellos, lo que tiene un efecto de mejora en la precisión a la hora de interrogar el corpus.
- 4) Son más fáciles de interpretar ya que la unicidad del género o tema reduce drásticamente las posibilidades de interpretar una concordancia, lo que también ocurre a la hora de interpretar la posición o función de una concordancia dentro de su texto.
- 5) Son más fáciles de construir, permitiendo que los mismos estudiantes los compilen y así se familiaricen desde un principio con los textos y con cuestiones de diseño de corpus.
- 6) Son más fáciles de reconstruir, permitiendo la posibilidad de dividir un corpus en subcorpus aún más específicos como, por ejemplo, de las distintas partes que conforman los textos.
- 7) Es más fácil identificar los patrones (colocaciones, patrones gramaticales y connotaciones) debido a la homogeneidad.
- 8) Los límites están más claros o son más obvios, si bien es cierto que habrá que cuidar que el estudiante no haga generalizaciones que vayan más allá del tipo de texto contenido en el corpus.

Según Aston, además, un corpus pequeño y especializado es un recurso a partir del cual es posible generar hipótesis que después se pueden comprobar en un corpus grande, al que el estudiante se aproximará ya con algún tipo de criterio en mente a la hora de

elaborar sus preguntas y clasificar los datos. Gavioli (2001), por su parte, recomienda el uso de corpus pequeños y homogéneos de textos cortos y sencillos especialmente en los primeros pasos de identificación de recurrencias en concordancias. En una publicación posterior, Aston (2002) apostilla que, si bien la construcción de un corpus propio aporta elementos de control y seguridad sobre los textos incluidos, da pie a la creatividad y la comunicación y desarrolla el sentido crítico, también existen razones de peso para utilizar corpus ya “empaquetados”. En este segundo caso, las aportaciones del corpus se centran en una mayor fiabilidad de los contenidos y, por ende, de los resultados, una mejor documentación de los textos, la conveniencia que supone el ahorro de esfuerzo de creación de un corpus y que, además, en ocasiones estos corpus vienen ya con un software especialmente diseñado para su interrogación. Asimismo, entre las opciones de construir un corpus ad hoc y adquirir un corpus diseñado por otros, existe también un camino intermedio representado por la opción de crear un subcorpus a partir de un corpus más grande. Estos subcorpus aportan todas las ventajas descritas anteriormente, si bien suelen adolecer de especificidad y tamaño suficientes.

3.11.3. Usos de los corpus

Dentro de la didáctica de lenguas, el área que más se ha beneficiado de la aplicación de corpus, y en concreto de corpus pequeños especializados contruidos ad hoc, es la enseñanza del inglés para fines específicos. Vocabulario especializado, colocaciones (formas específicas de combinar términos u ordenar información) y rasgos estilísticos propios son las características de una lengua para fines específicos (Bowker y Pearson 2002). Cada vez más, profesores en centros universitarios de una variedad de países diseñan cursos basados en corpus para enseñar el inglés de dominios concretos en los que, por lo general, los intereses y necesidades de los estudiantes son fáciles de definir y se pueden construir corpus especializados de acuerdo con ellos. Sólo a modo de ejemplo, citaremos el trabajo de autores, en su mayoría lingüistas y profesores, como el de Yang (1985) y el corpus *JDEST* de ciencia y tecnología, o Zhu (1989) y el corpus *GPEC* sobre la industria del petróleo, ambos contruidos en China y que permitieron la creación de materiales didácticos hechos prácticamente a medida para los estudiantes de estos sectores. Otros ejemplos más recientes son el trabajo de Jabbour (1997), relacionado con el desarrollo de materiales didácticos para un curso de lectura para médicos a partir de un corpus de artículos de investigación médica, y el de Narita, Kurokawa y Utsuro (2002), que describe *BEAR (Building English Abstracts by Ricoh)*, una herramienta basada en un corpus paralelo inglés-japonés diseñada para ingenieros

informáticos para la redacción de resúmenes en inglés. Aston (2000a) menciona que Collins utiliza un corpus de 400 cartas comerciales para diseñar un curso de correspondencia comercial en inglés. Tribble (1997) propone la creación de corpus a partir de recursos de fácil adquisición como CD-ROM de publicaciones; en su caso, a partir de la creación de un corpus sobre comercio, produce ejercicios de rellenar blancos con el objetivo de que el estudiante aprenda hiperónimos de palabras frecuentes en este campo. Foucou y Kübler (2000), desde Francia, no sólo diseñan un corpus de documentos procedimentales e informativos sobre informática para impartir un curso de inglés dentro de este campo, sino que van más allá desarrollando un entorno basado en Internet que permite la generación automática de ejercicios lingüísticos dirigidos al estudiante. Brodine (2001), por su parte, no describe su trabajo práctico en el aula, pero sí la utilización que hace de dos corpus comercializados junto con otro construido ad hoc para investigar precisamente la manera en que el trabajo con corpus y concordancias puede contribuir a un curso de inglés para fines académicos. En España, si bien la aplicación de corpus a la enseñanza de lenguas no está aún muy extendida, sí destacan algunos nombres como Curado Fuentes (Curado Fuentes y García Berzosa 1999; Curado Fuentes y Edwards Rokowski 2002), de la Universidad de Extremadura, que utiliza corpus con el fin de enseñar inglés en el campo de la informática y de la economía, o Bravo, catedrático de Filología, de la Universidad de Valladolid.

Otras aplicaciones de corpus en el aula se encuentran en la enseñanza de distintos aspectos del inglés general, sobre todo como lengua extranjera. Así, Kettemann (1997) aporta ejemplos de cómo utilizar concordancias para enseñar o reforzar algunos aspectos lingüísticos de especial dificultad para sus estudiantes de inglés alemanes. Frankenberg-García (2002) explota concordancias paralelas para enseñar inglés a sus alumnos portugueses y, en España, Pérez-Paredes y Cantos Gómez (2002), de la Universidad de Murcia, enseñan inglés con corpus entendiéndolo como una fuente de actividades. Osborne (2000) describe las bases de un curso de gramática y lingüística inglesas para alumnos franceses en un centro de auto-aprendizaje en el que las concordancias tienen enlaces a ejercicios y apuntes. En relación con el español como lengua extranjera, Davies (2002) destaca no sólo por la compilación de un corpus diacrónico extenso del español, sino también por el diseño de cursos en línea de historia del español y variación lingüística basados en corpus, como ya se ha apuntado en apartados precedentes. Fuera del aula, Rundell (1997), por ejemplo, hace uso de corpus para estudiar el fenómeno de la ironía y la prudencia en inglés y comprobar cómo lo utilizan los estudiantes. Lenko-Szymanska (1997), aparte de utilizar corpus

para mejorar sus clases de inglés como lengua extranjera, ha descubierto el potencial que estos recursos tienen para el propio profesor no nativo, asistiéndole no sólo a la hora de preparar materiales, sino también a la hora de corregir errores cometidos por los estudiantes en sus redacciones escritas y para cuya evaluación sería necesaria una competencia cercana a la de un profesor nativo. Con el énfasis puesto en otros aspectos, Gavioli (1997) describe ejercicios para hacer al estudiante consciente de la relación entre el texto y los datos de frecuencias, estimular su curiosidad e implicarlo en el proceso de investigación, mientras que en otros ejercicios (2001) persigue demostrar la naturaleza repetitiva de algunos tipos de texto, lo que no es óbice para que también exista opción a la creatividad.

En resumen, el uso que se ha hecho de corpus y concordancias en la enseñanza de lenguas hasta el momento ha sido amplio y variado, aunque se pueden establecer categorías (no excluyentes) según la finalidad pedagógica que se les ha dado: para diseño de materiales didácticos y temarios, como complemento de materiales de referencia ya existentes, para el aprendizaje autónomo y como fuente de actividades de aprendizaje.

3.11.3.1. Diseño de materiales didácticos

Es probablemente el uso más obvio y abundante que han tenido los corpus en la enseñanza de lenguas. Entre los primeros ejemplos se encontraría Honeyfield quien en 1989 ya desarrolló una tipología de ejercicios basados en materiales extraídos de concordancias. En todos los casos de esta categoría el profesor selecciona y, en ocasiones, adapta material extraído de corpus para enseñar o ilustrar distintos aspectos lingüísticos o metodológicos. Para mencionar sólo algunas ideas, los corpus pueden servir de ayuda en la creación de materiales para, por ejemplo, identificar patrones contextuales recurrentes (p.ej. colocaciones, patrones gramaticales, preferencia semántica y prosodia semántica), distinguir sinónimos (p.ej. “couple”/“pair”), o más bien comparar rasgos lingüísticos diferentes en palabras fáciles de confundir (p.ej. “see”/“watch”), comparar rasgos lingüísticos o palabras en diferentes géneros o registros (p.ej. “interests” en arte o en comercio) (Aston 2001b), identificar recurrencias tanto dentro de un texto como entre distintos textos, estudiar expresiones idiomáticas, la metáfora, la referencia o la sintaxis (p.ej. “for”/“since”) (Partington 2001), mejorar la lectura de textos académicos (p.ej. a través de actividades centradas en el léxico, la gramática o el discurso) (Brodine 2001), identificar y extraer términos

especializados, aprender sobre la gramática y el estilo de una lengua para fines específicos, identificar neologismos (Bowker y Pearson 2002), etc.

3.11.3.2. *Diseño de temarios*

Apoyados en datos y pruebas arrojadas por los corpus, algunos temarios de cursos de enseñanza de lenguas han experimentado cierta reestructuración con el fin de ajustarse más a la realidad de la lengua para incluir contenidos y usos de la lengua más frecuentes, mejor descritos y pedagógicamente más relevantes. En este sentido, y a pesar del debate ya apuntado entre lo que es científicamente interesante y lo que es pedagógicamente útil, parece que se camina hacia una pedagogía que parte de los intereses del estudiante, tanto en el sentido de lo que le preocupa como lo que es relevante para su aprendizaje.

3.11.3.3. *Complemento de materiales de referencia ya existentes*

Si hay algo que los corpus, especialmente los grandes y generales, han revelado son las deficiencias de diccionarios y gramáticas a la hora de describir la lengua. Muchos son los autores que han comentado sobre las limitaciones de los materiales de referencia en distintas áreas (Lenko-Szymanska 1997; Tognini-Bonelli 2001; Zorzi 2001) reconociendo a la vez los beneficios que aporta el uso de corpus. Bowker y Pearson (2002: 14 y ss), por ejemplo, afirman que, frente al uso de diccionarios, documentos en papel, consulta a expertos e intuición, un corpus puede ofrecer una pléyade de ventajas:

- Puede ser más extenso en tamaño.
- Es más fácil de actualizar.
- Es más fácil de consultar.
- Los datos se pueden buscar mucho más rápido, de manera más eficaz y completa.
- Contiene documentos auténticos que reflejan el uso real de la lengua.
- Con la ayuda de herramientas de análisis, pueden encontrarse patrones significativos con más facilidad.
- Aporta información sobre lo que no se usa.
- Aporta información estadística de forma objetiva y consistente.
- Ofrece un acceso ilimitado.
- Contiene información de más de un experto.
- Proporciona soporte documental a la intuición.
- Contiene información léxica y no léxica (p.ej. estilo, puntuación, gramática, registro).
- Aporta información lingüística y conceptual.

- Es más divertido utilizar corpus que fuentes documentales tradicionales, además de que permite aprender otras cosas a la vez.

Ventajas que ofrece un corpus en comparación con otras fuentes de documentación (traducido de Bowker y Pearson 2002: 14 y ss).

3.11.3.4. *Aprendizaje autónomo*

Sugerido anteriormente (cfr. supra 3.11.2.) a propósito del enfoque comunicativo y su énfasis en este aspecto, los corpus son, según Leech (1997: 3) el mejor ejemplo del paradigma del aprendizaje por descubrimiento centrado en el estudiante, quien tendrá que desarrollar estrategias para poder resolver problemas tanto durante su aprendizaje como después. La mayoría de los ejemplos expuestos en este capítulo responden a autores que creen en el posicionamiento fuerte respecto al uso de corpus, es decir, sostienen que el estudiante debe interactuar con el corpus de una forma no mediada, bien como un investigador, bien como un viajero, pero independiente, autónomo. El profesor aquí hace un papel de asesor o guía y el corpus sirve de estímulo a la investigación o al viaje.

3.11.3.5. *Fuente de actividades de aprendizaje*

Relacionado con el apartado anterior, Bernardini (2000a), por ejemplo, comenta las posibilidades que ofrece un corpus grande y general no ya como herramienta de referencia, sino como entorno de búsqueda y creación de actividades de aprendizaje exploratorio. Aston (2000a) y Gavioli (1997), por su parte, aportan la idea de los corpus como fuentes de actividades comunicativas en las que el corpus es, de nuevo, un estímulo a la lectura, a la discusión, al conocimiento y a la comparación.

En relación con todos los usos que se les han dado a los corpus en pedagogía de lenguas, los autores que han puesto en práctica actividades basadas en corpus en sus aulas han comentado sobre sus ventajas, si bien es cierto que algunos no se adscriben ciegamente a esta metodología y llaman la atención sobre aspectos que no deben descuidarse. Entre las ventajas comentadas, Aston (1997a) sostiene que el uso de corpus enriquece el entorno de aprendizaje porque propicia la oportunidad de utilizar el inglés, es decir, sirve de estímulo comunicativo, como se apuntaba en el párrafo anterior. Además, parece que el estudiante reduce su dependencia del profesor, y éste su dependencia del libro de texto. Jabbour (1997) centra su atención en las ventajas que el uso de corpus trae a la organización de la clase en el sentido de que, según él, se

aprovecha mejor el tiempo porque las explicaciones requeridas son mínimas dado que los estudios con corpus son suficientemente reveladores y flexibles. La lengua se ve y estudia en su contexto, siempre relacionando significado y uso. La combinación de ordenador y corpus le parece a Leech (1997) una fuente de ventajas como, por ejemplo, que las búsquedas y organización de los datos son automáticas, se promueve una aproximación a la enseñanza centrada en el estudiante, ofrece una provisión ilimitada de datos, permite que el proceso de aprendizaje sea a medida del estudiante, y además permite la auto-evaluación. Para Gavioli (2001), las concordancias suponen ya una ventaja por el hecho de incluir una serie de casos que a menudo muestran más de un patrón de uso, pero incluso, dice la autora, aún si sólo se muestra un patrón, el estudiante todavía tendrá que hacer un análisis considerable para identificarlo, implicándole así en el proceso de investigación. Por su parte, Kettemann (1997) expresa la más apasionada exposición de las razones por las que se deberían usar concordancias en la enseñanza de lenguas.

“As an extremely powerful hypothesis testing device on vast amounts of data, the computer allows controlled speculation, makes hidden structures visible, enhances at the same time imagination and checks it by inductivity, thus making higher degrees of objectivity possible. The students’ work becomes more exploratory and thus motivating and highly experiential” (Kettemann 1997: 63).

Dentro de la investigación sobre la enseñanza de lenguas extranjeras basada en corpus, las interlenguas y el análisis de errores también han atraído una cierta atención. La investigación en estas áreas se ha beneficiado de la disponibilidad de corpus de estudiantes, es decir, “electronic collections of spoken and written texts produced by foreign/second language learners” (Granger 1999: 191). Un ejemplo de este tipo de corpus es el *ICLE (International Corpus of Learner English)*, que se divide en varios subcorpus que contienen textos producidos por estudiantes de inglés de diversas nacionalidades. En algunos corpus de estudiantes se ha llevado a cabo una anotación de los errores realizados por los estudiantes, lo que ha resultado ser especialmente conveniente para aquellos especialistas que se dedican a la preparación de materiales de enseñanza. Sin embargo, a pesar de las ventajas vislumbradas del uso de corpus en este área, debemos reconocer que la aproximación basada en corpus a la enseñanza de lenguas en general no ha tenido tanta influencia como era de esperar en, por ejemplo, la producción de libros de texto. Hasta cierto punto, se podría decir que la aplicación de corpus en el aula ha estado obstaculizada por la reticencia de algunos profesores a abandonar una pedagogía centrada en el significado, y otras razones prácticas como la falta de equipamiento y conocimientos informáticos por parte de alumnos y profesores.

3.11.4. Situación actual y necesidades

La década de los 90 ha sido testigo de un creciente interés en las aplicaciones pedagógicas de los corpus, como puede comprobarse por las series de conferencias TALC (Teaching and Language Corpora) y CULT (Corpus Use and Learning to Translate), celebradas regularmente en Europa, aparte de otras relacionadas con otros usos de corpus como PALC (Practical Applications of Language Corpora), LREC (Language Resources and Evaluation), CL (Corpus Linguistics) o ICAME (International Computer Archive of Modern and Medieval English), ésta última con orígenes en 1979. Sin embargo, el trabajo con corpus de autores como Sinclair se remonta a los años 60, trabajo que después tendría sus efectos en la pedagogía de lenguas a través, sobre todo, de la producción de materiales de referencia basados en corpus. Leech empieza su trabajo con corpus a finales de los 70, mientras que Johns dedica los años 80 y 90 a aplicar corpus directamente en el aula, por poner algunos de los ejemplos más visibles. En el transcurso de estos años los corpus han visto, entre otras cosas, un desplazamiento de su uso del laboratorio al aula para ser accedidos directamente por los estudiantes (Leech 1997; Gavioli 2001). Si bien es cierto que la posibilidad de acceder a los corpus en entornos académicos ha abierto las puertas a numerosas aplicaciones prácticas, también lo es que algunas voces han llamado ya la atención sobre la falta de una base teórica clara fundamentada en la lingüística aplicada, además de la falta de investigaciones empíricas en el campo del aprendizaje de lenguas, tanto mediado como independiente, para comprobar si los corpus son unas buenas herramientas pedagógicas y si compensa el esfuerzo (Aston 2000a; Bernardini 2000b). En este sentido, Brodine (2001) se pregunta sobre la necesidad de utilizar concordancias en el aula.

“Why should concordances be used to create decontextualized examples of reference problems in discourse when learners can encounter all kinds of contextualized instances in their assigned reading materials? One reason is that juxtaposing instances in brief, judiciously selected parallel passages makes it easier to see both regularities and a range of variability” (Brodine 2001: 161).

Respecto a la repetición y la variación, comentadas anteriormente (cfr. supra 3.11.1.), se ha expresado la conveniencia de exponer al estudiante a ambos aspectos para dar lugar a una reestructuración de esquemas y conocimientos (Aston 1995). El generador de concordancias, el instrumento principal de la lingüística de corpus, permitirá al estudiante identificar regularidades contextuales (colocaciones, patrones gramaticales, preferencia y prosodia semánticas) que serán de gran relevancia para su aprendizaje.

Para ser capaz de identificar estas regularidades, hacer observaciones y extraer generalizaciones, el estudiante habrá de ser dotado de conocimientos y habilidades tanto técnicas como metodológicas. Además, el profesor, que actuará de asesor o guía en sus actividades de aprendizaje con corpus, le aportará *feedback* ayudándole a desarrollar y mejorar sus estrategias de investigación y resolución de problemas. Algunos autores han expuesto formas de optimizar el trabajo con corpus del estudiante:

- la implicación del alumno en el proceso de selección de textos para estimular la participación y la motivación del mismo;
- la construcción de corpus a base de textos ya conocidos por el estudiante;
- la utilización de diversos corpus para poder contrastar datos y corroborar o refutar hipótesis;
- la necesidad de hacer un informe escrito u oral sobre el trabajo realizado o, más bien, los procesos seguidos y los descubrimientos hechos, lo que puede llevar a una reflexión crítica sobre los límites de los corpus consultados y el alcance de sus generalizaciones (Bernardini 2000b);
- la entrega de listas de control para asistir al estudiante en sus análisis (Bernardini 2000a, Zorzi 2001);
- la reducción de la dificultad en lo que respecta al corpus, las tareas o los datos;
- la instrucción en la compilación de corpus basados no sólo en criterios temáticos, sino también de género, fuente y público de llegada (Pearson 2000);
- la manera de presentar los datos: sin manipular, seleccionados u ordenados, según el nivel de dificultad que se quiera conseguir en la actividad (Zorzi 2001);
- la guía y el control por parte del profesor sobre todo en las primeras actividades, hasta llegar a una autonomía casi total; y
- el desarrollo de una “mentalidad de corpus” (Aston 2001b: 44), que consiste en enseñar a los estudiantes que los corpus proporcionan fuentes de información de mayor o menor fiabilidad sobre la cultura y la lengua, los textos y los tipos de textos, y además ofrecen toda clase de oportunidades para que la comunicación oral o escrita se desarrolle.

Recopilación de formas de optimizar el trabajo con corpus del estudiante.

Finalmente, como hemos mencionado (cfr. supra 3.11.2.), para la realización de actividades de aprendizaje basadas en corpus y la puesta en práctica de las estrategias que el estudiante ha de ir desarrollando, muchos son los que han abogado por la utilización de corpus pequeños y especializados con textos de un sólo tema o género, y pocos los que apuestan por el uso de corpus grandes y generales. Dentro del segundo grupo, Bernardini (2000a, 2000b, 2000c, 2001) es la que más ha trabajado y abogado por el uso de corpus grandes en el aula de traducción y defiende las posibilidades que ofrecen estos corpus, grandes y heterogéneos, destacando su capacidad para hacer

pensar al estudiante por sí mismo a la hora de interpretar las pistas que dan a sus preguntas, retándolo y estimulando finalmente el desarrollo de un método autónomo de investigación y trabajo. Así, con la concepción mental del aprendizaje como un viaje y del alumno como un explorador, Bernardini expone la puesta en práctica de ejercicios abiertos donde el alumno escoge su punto de partida, su camino y su punto de destino en un viaje que le proporcionará, de seguro, oportunidades para el aprendizaje. Según palabras de la autora (comunicación personal), una clase de lengua y una clase de traducción no difieren tanto sino en el objeto de énfasis, a lo que añade que, en toda clase de traducción siempre hay un componente de enseñanza de la lengua. Sus ejercicios de traducción basados en corpus tienen como objetivo desarrollar diversas capacidades o habilidades de traducción con el fin de que el alumno sea más consciente de lo que implica interpretar y traducir un texto, que sea más reflexivo en su análisis textual y que tenga un mayor dominio de los recursos tanto a la hora de explotar aquellos ya disponibles como a la hora de encontrar otros nuevos, además de desarrollar su capacidad de aprender autónomamente.

Bernardini (2000c: 4) sostiene que los cimientos pedagógicos de la enseñanza de la traducción son ciertamente débiles si la metodología subyacente en un curso de traducción se reduce a simplemente ofrecer a los estudiantes exposición a la traducción, oportunidades para traducir y corrección, como así es en la enseñanza tradicional de la traducción. En este sentido, Bernardini apuesta por una enseñanza basada en actividades abiertas con un componente de impredecibilidad en las que el punto de inicio y final no importen; el camino tampoco estará preestablecido y no tendrá importancia si el alumno llega al destino que tenía en mente mientras que ese camino sea rico en experiencias y encuentros casuales fructíferos lo que, todo junto, supondrá un estímulo a la comunicación.

Finalmente, creemos que el uso de corpus puede aportar grandes beneficios al campo de la enseñanza de la traducción, al igual que lo hace en la enseñanza de lenguas. Sin embargo, hay que tener en cuenta que, como apunta Tognini-Bonelli (2001) llevar pruebas arrojadas por los corpus al aula representa una innovación metodológica y teórica porque es, al fin y al cabo, una nueva forma de enseñar.

4. USO DE CORPUS EN LA DIDÁCTICA DE LA TRADUCCIÓN

En una época en la que la oferta y la demanda de información están al alcance de todos y el acceso a las nuevas tecnologías es cada vez mayor y más fácil, el campo de la traducción profesional no podía quedarse al margen de esta revolución informativa e informática. En un mundo en el que cada vez se mueve un número mayor y más variado de documentos y tienen lugar actos de comunicación de todo tipo, el traductor tampoco podía quedarse impávido ante las exigencias del mercado. Así, en este entorno cada vez más competitivo, el traductor se ve más bien como un profesional de la lengua al que se puede requerir para prestar servicios no sólo como traductor, sino también como terminólogo, documentalista, redactor técnico, localizador, corrector o gestor de proyectos, por poner sólo algunos ejemplos.

Aún si nos centramos exclusivamente en un traductor cuya función principal sea la de traducir, muy probablemente, tendremos que hablar de su capacidad de trabajar en distintos campos de especialidad adaptándose a diferentes géneros, registros, etc. e incluso traduciendo en ocasiones en ambas direcciones (directa e inversa). En estas circunstancias, el profesor de traducción encargado de formar a futuros traductores se encuentra no sólo con limitaciones naturales como la falta de tiempo en el aula, sino también con carencias propias, por ejemplo en cuanto a variedad de conocimientos y experiencia en todas las áreas en las que el alumno de traducción se debería preparar para poder hacer frente al mundo profesional. En otras palabras, el profesor de traducción, como proveedor de información, se encuentra ante la imposibilidad de atender todas las especialidades y variantes a las que aparentemente sería recomendable exponer al alumno. Por otra parte, la impredecibilidad de los textos con los que trabajará el alumno de traducción en su vida profesional hace que la especialización, aunque parezca necesaria, es paradójicamente, desaconsejable (Gavioli 1996).

Ante estos hechos, el mejor papel que puede cumplir el profesor de traducción, desde nuestro punto de vista, no es el de proveedor de información, sino el de facilitador, acompañando al alumno en su proceso de aprendizaje, asistiéndole allá donde lo necesite y estimulando el desarrollo de sus conocimientos operativos. Esto no quiere decir que la adquisición de conocimientos declarativos no sea relevante para un futuro traductor, sino que será mucho más productivo desarrollar en el alumno habilidades

que pueda aplicar en situaciones nuevas a la hora de resolver problemas. En otras palabras, cuando los estudiantes de traducción salgan al mundo real tendrán que tratar con temas y disciplinas muy diversas, teniendo que convertirse en mini-expertos en un corto espacio de tiempo. Es en este marco donde los corpus aparecen como un recurso reutilizable de gran valor y la lingüística de corpus como una metodología aplicable a las distintas fases de la traducción y, como apoyamos aquí, a la didáctica de la traducción. En este sentido, ya son varios los autores que no sólo abogan por la integración de los corpus en las clases de traducción, sino que consideran el manejo de corpus (*corpus management in translation*) como una necesidad e incluso una obligación para el futuro traductor (Maia 1997, 2003; Aston 1999; Friedbichler y Friedbichler 2000, Corpas Pastor 2002; Varantola 2003, etc.).

“[...] the knowledge of how to compile corpora and use them is an essential part of modern translational competence” (Varantola 2003: 55).

4.1. Ventajas de los corpus electrónicos frente a diccionarios y textos paralelos en papel

Es de destacar que la práctica totalidad de los autores que han trabajado con corpus subraya las ventajas del uso de estos recursos y metodología frente al uso de obras de referencia como diccionarios o el uso de textos paralelos convencionales en papel, siendo estas ventajas tanto prácticas como de aprendizaje y aplicándose tanto dentro como fuera del aula de traducción.

Exigencias editoriales hacen que los diccionarios tanto generales como especializados presenten dos problemas principales para un usuario como puede ser un traductor o un estudiante de traducción. En primer lugar, el tiempo que transcurre entre la preparación de una obra de este tipo y su disponibilidad real de consulta provocará un desfase temporal especialmente en campos donde la terminología evoluciona con rapidez (Pearson 1996; Maia 1997). En segundo lugar, por lo general, los diccionarios presentan los términos de una forma descontextualizada y, en ocasiones, acompañados de ejemplos inventados sobre los que no se suele dar información de otro tipo que no sea puramente léxica o semántica. En contraste, y como afirman Corpas Pastor y Postigo Pinazo (2002), un corpus puede dar información de tipo discursivo, cultural, terminológico, sintáctico, combinatorio, semántico, cognoscitivo, fraseológico y estilístico, a lo que se añade también la información estadística que ayuda a interpretar

la relevancia de los patrones hallados. Como así sostiene Bowker (2000: 20), los textos paralelos pueden ayudar a superar las limitaciones de recursos lexicográficos como diccionarios en cuestiones como, por ejemplo, la información contextual. Sin embargo, las ventajas que presentan los textos paralelos convencionales en papel frente a los diccionarios se diluyen cuando se pone de manifiesto el esfuerzo que requiere la recogida de documentos impresos, cuya consulta obliga a la lectura de bastante material irrelevante. Aquí el traductor también se encuentra ante la imposibilidad de recoger y consultar una gama de documentos suficientemente variada como para asegurarse de que todos los conceptos, términos y patrones lingüísticos estarán presentes, a lo que Church *et al.* (1991: 116) añaden, además, la propensión del ser humano a cometer errores a la hora de analizar grandes cantidades de texto e intentar extraer regularidades a ojo, valga la expresión. En este sentido, el corpus electrónico se puede ver como un avance respecto a la tradición de utilizar textos paralelos en traducción, lo que tiene implicaciones para la enseñanza.

“[...] el uso de corpus para la pedagogía de la traducción constituye, hasta cierto punto, un desarrollo lógico si tenemos en cuenta dos factores clave: primero, la creciente ‘informatización’ y ‘automatización’ de la actividad traductora como proceso y como producto; y, segundo, el uso tradicional de textos ‘paralelos’ como fuente de documentación más fiable y completa que los diccionarios al uso” (Corpas Pastor 2002: 5).

Sin hacer aún distinciones de tipo, las ventajas que presenta un corpus electrónico, en general, se pueden resumir en los siguientes puntos:

- Se puede construir en menos tiempo que un corpus convencional en papel.
- Se puede reutilizar.
- Permite la búsqueda rápida y sistemática de información en una gama de documentos más amplia y variada.
- Ofrece un número mayor de ejemplos y contextos.
- Permite ir directamente a aquellas áreas del texto que contienen las palabras clave con la ayuda de una herramienta de análisis de corpus.
- Facilita la detección de información como patrones colocacionales.
- Facilita la detección de datos de frecuencias, lo que permite determinar la relevancia de los patrones encontrados.
- Facilita y acelera la recuperación y el análisis de datos concretos si el corpus está anotado.
- Ofrece seguridad tanto a la hora de tomar decisiones estratégicas como elecciones léxicas.
- Ofrece mayores oportunidades de que se dé el aprendizaje casual.

Como ya se ha apuntado anteriormente (cfr. supra 4.1.), el corpus electrónico ofrece una información más completa y variada que otros recursos. A menudo, los diccionarios no alcanzan a ayudar al traductor en su tarea, especialmente cuando se trata de una terminología muy especializada o de cuestiones de registro, fraseología, etc., razón por la que algunos autores han sugerido la utilización del diccionario en combinación con otros recursos como pueden ser los corpus electrónicos (Fraser 1999; Varantola 2003). En cuanto a la enseñanza de la traducción, Aston (1999), afirma que el corpus electrónico proporciona un entorno de aprendizaje más rico, y Corpas Pastor (2002) destaca los beneficios que el uso de corpus puede aportar, devolviéndonos a la primera cuestión tratada en este apartado.

“Unlike the dictionary, a concordance leaves it to the user to work out how an expression is used from the data. This typically calls for deeper processing than does consulting a dictionary, thereby increasing the probability of learning (Hultsijn 1992). In more general terms, by drawing attention to the different ways expressions are typically used and with what frequencies, corpora can make learners more sensitive to issues of phraseology, register and frequency, which are poorly documented by other tools” (Aston 1999: 295).

“El corpus favorece, pues, el desarrollo de los procesos autónomos de enseñanza-aprendizaje, estableciendo los mecanismos adecuados para la especialización y, al mismo tiempo, la diversificación que el mercado laboral exige al traductor hoy en día” (Corpas Pastor 2002: 6).

Por otra parte, y a pesar de las muchas ventajas aquí expuestas respecto a la utilidad de los corpus electrónicos en la traducción y su enseñanza, también es necesario mencionar que el uso de este recurso y metodología exige una formación previa y un entendimiento de sus implicaciones. Como así mantiene Bowker (2000: 46), el (futuro) traductor necesita que se le forme tanto en las ventajas y las limitaciones de las herramientas y recursos basados en corpus, como necesita que se le forme en los puntos a favor y en contra de los recursos convencionales. En este marco cobra especial relevancia el conocimiento de los distintos tipos de corpus que pueden ser de utilidad para el traductor profesional o traductor en potencia.

4.2. Tipos de corpus más relevantes para la didáctica de la traducción escrita

Como ya se ha indicado al principio de este trabajo, los tipos de corpus que han sido de mayor utilidad en la enseñanza de la traducción son, principalmente, el corpus

paralelo, el corpus comparable bilingüe y el corpus comparable monolingüe o especializado en lengua meta, siguiendo a Bowker (2000: 17 y ss). Ahora repasaremos los usos que se han dado a cada uno de estos tipos de corpus, centrándonos en los beneficios que pueden aportar a la enseñanza de la traducción.

4.2.1. Corpus paralelo

Un corpus paralelo se compone de textos originales (TO) en lengua A alineados con sus traducciones o textos meta (TM) en lengua B, si hablamos de un corpus bilingüe, aunque también existen corpus paralelos multilingües en los que los TOs están alineados con varias traducciones. En general, los corpus paralelos se han explotado en el campo del procesamiento del lenguaje natural (PLN) y la traducción automática (TA) para la mejora de sistemas existentes, la traducción asistida por ordenador (TAO) para su uso como memorias de traducción y la extracción de equivalentes de traducción para traductores humanos con el fin de, por ejemplo, crear glosarios o bases de datos.

En un corpus paralelo se puede observar cómo uno o varios traductores han tratado con problemas concretos y establecido equivalencias en circunstancias determinadas, así como obtener ejemplos de estrategias de traducción (Maia 1997; Zanettin 1998; Pearson 2003). En este sentido, las implicaciones para la enseñanza de la traducción no están en la línea de la observación de las concordancias bilingües alineadas extraídas del corpus como si éstas fueran modelos de perfección a seguir. En otras palabras, en el aula de traducción, el corpus paralelo podría servir como punto de partida para la discusión de estrategias, contextos, connotaciones, diferencias de registro, género, etc. y no como una fuente de respuestas fijas a problemas de traducción. Aston (1999), por su parte, también comenta el impacto positivo que el uso de corpus paralelos puede tener en el aprendizaje. Según el autor, el hecho de que se den una serie de manifestaciones paralelas puede ayudar a los estudiantes a distinguir entre contextos de uso diferentes y reducir su tendencia a pensar en términos de equivalencia unívoca. Además de esto, un generador de concordancias paralelas también se puede utilizar para analizar uno o varios TOs y su traducción o traducciones, lo que puede ser especialmente útil para un traductor o estudiante a la hora de evaluar su propio trabajo, así como entender y comprobar aspectos como la consistencia terminológica, por ejemplo. Precisamente, lo que Bowker y Bennison (2003) han hecho es crear un archivo de traducciones de estudiantes (*Student Translation Archive - STA*) y un sistema para gestionarlo (*Student Translation Tracking System - STTS*). El valor pedagógico de un corpus de traducciones y de éste en concreto se basa en las

posibilidades que abre, por ejemplo, a la detección de áreas comunes de dificultad en las traducciones encargadas, la detección de problemas transversales individuales estén relacionados con el campo temático, el género, el registro, etc., la observación del impacto de la lengua de partida, o el estudio longitudinal del progreso de un estudiante o un grupo, lo que a su vez facilita labores como la evaluación o la autoevaluación.

Sin embargo, los corpus paralelos también presentan algunas desventajas de uso, como en lo que atañe a su creación. La necesidad de encontrar textos originales con sus traducciones y alinearlos ralentiza el proceso de compilación de este tipo de corpus incluso cuando la alineación se lleva a cabo de forma semi-automática. Respecto a la selección de textos, diversos autores (Aston 1999; Friedbichler y Friedbichler 2000; Bowker 2000, Maia 2003), también apuntan la falta de fiabilidad que presentan los textos incluidos en un corpus paralelo. Por un lado, es probable que los TOs hayan sido editados antes de traducirse y, por otro, la autoría de las traducciones o TMs puede ser dudosa ya que, en la mayoría de las ocasiones, no se tiene información sobre la identidad o la competencia del traductor, y tampoco sobre el estatus o la calidad de la traducción. Por esta razón, muchos traductores profesionales desconfían del material traducido como recurso de documentación. En lo que atañe a la fiabilidad, ésta también puede verse limitada si todas las traducciones de un corpus pertenecen al mismo autor, como así pasa a veces con las memorias de traducción. Y en la misma línea, Aston (1999) llama también la atención sobre el riesgo que comporta utilizar traducciones como modelos de TM ya que se pueden reproducir aquellos rasgos peculiares de la traducción, algunos de ellos identificados como universales de la traducción. Por último, Friedbichler y Friedbichler (2000) añaden que, respecto al uso de corpus paralelos, la decisión de cuál o cuáles de las traducciones ofrecidas es la más adecuada para un contexto determinado requiere un dominio considerable de la lengua y el análisis textual. Sin embargo, y contrastando con esta afirmación, Bowker y Pearson (2002: 193) sostienen que identificar equivalentes terminológicos y estudiar el uso (por ejemplo, observando las colocaciones) es relativamente fácil con un corpus paralelo, e incluso aportan un ejemplo en el que se observa el estilo de textos que pertenecen al género de acuerdos colectivos en inglés para concluir que este tipo de textos parece caracterizarse por el uso de la voz pasiva, la 3ª persona del singular y las frases largas, todo ello después de realizar unos ejercicios muy sencillos.

Respecto al uso de corpus paralelos en la enseñanza de la traducción, y frente a los comentarios anteriores, sugeriríamos que el uso de corpus paralelos en el aula de traducción se ha de adaptar, obviamente, al nivel de competencia de los alumnos. Así,

mientras que algunas actividades deberán realizarse en un nivel avanzado, es decir, cuando los alumnos ya tengan una noción y una práctica del análisis de textos tanto monolingüe como contrastivo, otras actividades sí podrán llevarse a cabo perfectamente a niveles de iniciación, teniendo en cuenta en este caso que el control del profesor sobre el corpus y la actividad será mayor que con alumnos de niveles más avanzados. En el caso del nivel inicial, la dificultad podría reducirse, por ejemplo, limitando el corpus paralelo a un solo género, pre-editando los textos o pre-editando las concordancias paralelas bilingües mostradas a los alumnos, lo que reduciría a su vez la gama de interpretaciones posibles que un elemento pudiera tener.

Volviendo a los problemas presentados por los corpus paralelos, sobre todo en lo que respecta a la fiabilidad de los textos incluidos (TOs y TMs), se ha recomendado el uso de estos corpus en combinación con otros bilingües comparables y también monolingües especializados, y ya son varios los autores que abogan por el uso de un nuevo tipo de corpus paralelo, el bidireccional o recíproco (Teubert 1996; Aston 1999; Bernardini, Stewart y Zanettin 2003). Este tipo de corpus se compone de TOs en lengua A y sus traducciones en lengua B, además de TOs en lengua B y sus traducciones en lengua A; algunos autores lo consideran como un corpus “ideal” ya que combina las ventajas de los corpus paralelos y los comparables en el sentido de que no sólo permite establecer relaciones entre TO y TM, sino también entre TOs en las dos lenguas. Ejemplos de este tipo de corpus son el *English-Norwegian Parallel Corpus (ENPC)* de inglés y noruego (Johansson y Hofland 1994), el *COMPARA* de inglés y portugués (Frankenberg-Garcia 2001), y el *CEXI* de inglés e italiano (Bernardini 2003). Un corpus paralelo bidireccional permite, entre otras cosas, realizar estudios lingüísticos contrastivos, investigar cambios discursivos en la traducción, desarrollar lexicografía y terminología bilingüe, refinar programas de traducción automática, investigar normas de traducción o universales y potenciar habilidades en el estudiante de lenguas o traducción en formación.

4.2.2. Corpus comparable bilingüe

Un corpus comparable es el compuesto por TOs en lengua A y TOs en lengua B, comparables a nivel de composición, género y tema, principalmente. Este tipo de corpus se ha utilizado en proyectos de lexicografía bilingüe, todo tipo de estudios contrastivos y estudios sobre la noción de equivalencia (Corpas Pastor 2001); además, el corpus comparable bilingüe también ha probado ser de gran utilidad en la enseñanza de la traducción por las posibilidades contrastivas que permite.

Como así afirma Zanettin (1998), la práctica de recoger textos en distintas lenguas en base a su similitud de tipo, contenido o función ya era común en la investigación y la formación en traducción antes de que existieran los corpus electrónicos. Autoras como Snell-Hornby (1988) o Schäffner (1996), que compilaron corpus no electrónicos, afirman que observando muchos textos que pertenecen al mismo tipo se pueden identificar rasgos prototípicos que pueden proporcionar un “perfil” del tipo de texto que el traductor tendrá que producir en la lengua meta. En este caso, un corpus comparable bilingüe con textos originales en las dos lenguas implicadas puede facilitar la observación de diferencias léxicas entre equivalentes, diferencias textuales y estructurales y puede aportar a los estudiantes de traducción una mejor comprensión no sólo de los textos de llegada, sino también de partida, permitiéndoles comparar terminología, fraseología y convenciones textuales entre lenguas y culturas (Maia 1997; López Rodríguez 2002; Bernardini, Stewart y Zanettin 2003), así como también puede ayudar a confirmar hipótesis de traducción y sugerir posibles soluciones. En esta misma línea, Zanettin (2001) nos ofrece un ejemplo concreto de los beneficios pedagógicos del uso de un corpus comparable bilingüe en italiano e inglés para la traducción inversa de un artículo de prensa especializado. Según el autor, a través de la comparación de algunos elementos (falsos amigos, préstamos, pseudo-sinónimos, expresiones metafóricas, etc.) en los dos corpus monolingües y la comparación de segmentos de texto (estructuras proposicionales y retóricas, información contextual, conectores lógicos, terminología, etc.), el estudiante no sólo pudo mejorar su traducción, sino también encontrarse con una gran cantidad de oportunidades de aprendizaje. Frente a los corpus paralelos, que pueden mostrar usos “no naturales” y por tanto no suponen un recurso totalmente fiable como guía para la producción de textos, la mayor ventaja o el resultado más importante del uso de corpus comparables es la naturalidad que se puede alcanzar en la lengua meta (Gavioli y Zanettin 2000). Según los mismos autores, los corpus comparables representan una fuente valiosa de información sobre estructuras, léxico o género, pudiendo mover al traductor o estudiante a hacer una reestructuración del texto más allá del nivel de las unidades fraseológicas y los patrones colocacionales, consiguiendo una reorganización más completa.

Los corpus comparables bilingües suelen ser pequeños, especializados y, en el ámbito que nos interesa, creados con un propósito pedagógico. La “comparabilidad” de los textos, tanto en una lengua como en otra, se basa en aspectos como el campo de especialidad, la relación autor-lector, el origen del texto y su constitución (si ha sido

creado por uno o varios autores), el grado de tecnicidad, la función comunicativa, etc. (Zanettin 1998). En el contexto de la enseñanza de la traducción, este tipo de corpus presenta claras ventajas pedagógicas (Gavioli y Zanettin 2000). Por ejemplo, la construcción de corpus pequeños especializados se puede considerar como una actividad de aprendizaje valiosa por sí misma. Para empezar, un corpus comparable bilingüe es relativamente fácil de construir, lo que ayuda a que los alumnos se impliquen en este proceso; la tarea de buscar y seleccionar textos para incluirlos en un corpus ofrece a los estudiantes una oportunidad de oro para encontrar mucho material y familiarizarse ya desde un principio con el contenido del corpus. En el caso concreto de los autores mencionados, la actividad llevada a cabo se basó en la recogida de artículos médicos de investigación en inglés e italiano por parte de los alumnos en un campo de especialidad de su elección, guardando cada sección de cada artículo por separado. Más adelante, los archivos guardados se pudieron combinar para formar corpus individuales que trataban temas específicos, o un corpus más general de investigación médica, en ambas lenguas. Zanettin ofrece pruebas concretas del uso de uno de los corpus individuales para analizar expresiones específicas de un campo (p.ej. “liver biopsy”), y del uso del corpus general para analizar expresiones menos restringidas (p.ej. “experienced”), así como de la obtención de información sobre la posición en el texto donde ocurren algunos rasgos concretos (p.ej. subcorpus de resúmenes o *abstracts*), todo ello provocando una reestructuración de las hipótesis de traducción iniciales del estudiante.

Como ocurría con los corpus paralelos, los corpus bilingües comparables no sólo tienen puntos a favor. Uno de los problemas que presenta la compilación de este tipo de corpus es el desequilibrio existente entre la producción textual en lenguas mayoritarias como es el inglés, y otras menos extendidas como el portugués (Maia 2000, 2003), lo que dificulta la búsqueda de textos, especialmente en formato electrónico y pertenecientes a determinados campos de especialidad, aunque esta situación está cambiando paulatinamente. Por otra parte, Aston (2000) añade un comentario referente a los corpus pequeños especializados y que bien podría aplicarse a los comparables bilingües.

“Una più generale difficoltà insita nell’uso di piccoli corpora specialistici durante la fase concomitante riguarda il fatto che essi difficilmente forniranno una risposta a tutti i problemi affrontati dai traduttori” (Aston 2000b: 26).

En este caso, el uso de este tipo de corpus también se puede complementar con el uso de otros corpus como puede ser uno general monolingüe (p.ej. *BNC*, *Bank of English*, *CREA*). Pero quizás la principal desventaja que presentan los corpus comparables bilingües es la falta de conexión entre las lenguas que los componen o, como lo llaman Friedbichler y Friedbichler (2000: 110) “the missing interlingual link” que, si bien se puede solucionar alineando textos paralelos indirectamente a través de una lista de palabras clave bilingüe del campo de especialidad, esto no está exento de problemas. Esta carencia de relación directa entre los TOs en una lengua y los TOs en otra lengua, hace que sea necesario desarrollar estrategias de búsqueda de equivalencias de traducción explotando tanto los conocimientos propios como la información que nos proporciona el corpus, que a su vez habrá que saber interpretar. En este sentido, Zanettin (2001), en un ejercicio concreto en el que sus alumnos utilizaron un corpus comparable bilingüe italiano/inglés, propone investigar dichos equivalentes de traducción a través de la búsqueda y la generación de concordancias de nombres propios, palabras clave y sus equivalentes (p.ej. oro/gold) y cognados (p.ej. podio/podium), algunos de los cuales serán falsos amigos (p.ej. consecutiva/consecutive). En el ejercicio que describe Zanettin, cuyas conclusiones son extrapolables a otros ejercicios de traducción con corpus, este recurso permitió la reflexión sobre cuestiones como la estructura discursiva del texto (p.ej. la función de una palabra o expresión en una posición determinada), o sobre la adecuación estilística (p.ej. la idoneidad o naturalidad de una expresión en un género determinado en la lengua meta), así como también permitió el aprendizaje casual.

“[...] these comparable corpora allowed students to contrast the source and target languages at various levels, from single words and phrases to discourse functions and organizations. By looking for patterns and regularities in the English and Italian texts and comparing the uses of words and expressions which they felt had some kind of relationship either within or across languages, they were able to find evidence to formulate and to support (or to reject) specific translation hypotheses” (Zanettin 2001: 189).

Sin embargo, aunque en los ejercicios que describe Zanettin el estudiante pudiera encontrarse en el corpus con segmentos similares a los de su TO o incluso segmentos similares a su traducción en proceso de elaboración, esto no significa que traducir con corpus se convierta en una actividad de cortar y pegar.

“The learner not only has to select relevant chunks of language according to their perceived meaning, but also to fit them together according to the desired overall outcome” (ídem.: 191).

El estudiante ha de trabajar con unidades fraseológicas que tiene que ajustar e integrar en el co-texto concreto y el contexto más general, lo cual puede ayudarle a desarrollar habilidades de ajuste y control según los propósitos y estilo del discurso. Seleccionar la estrategia de búsqueda de equivalentes, analizar los resultados y seleccionar los segmentos relevantes o comenzar una nueva búsqueda en un proceso de ensayo y error hasta encontrar el equivalente adecuado supone, a nuestro entender, un proceso complejo de toma de decisiones en el que el estudiante aprende a refinar sus propias estrategias e hipótesis. La relación que establece Zanettin (2001) entre buscar en un corpus con un generador de concordancias y navegar en la web se aplica a este contexto ya que el estudiante ha de aprender a ir de sus hipótesis a los corpus y volver, refinando sus estrategias de búsqueda y comprobando tanto si sus intuiciones son buenas, como si la información que le devuelve el corpus es correcta. En otro artículo, Gavioli y Zanettin (2000) llaman la atención sobre el riesgo de que los estudiantes se centren demasiado en el corpus de textos meta y tiendan a ignorar rasgos de la lengua original.

“[...] we believe that corpus work in the translation classroom needs to sensitize students to consider, in their final translation choices, adequacy to source language norms as well as acceptability according to target language ones, and that this may be encouraged by providing them with a variety of types of corpora, where learners may go from one to the other in order to validate their hypotheses” (Gavioli y Zanettin 2000).

Respecto a este último comentario, creemos que, precisamente, el valor pedagógico de un corpus comparable bilingüe descansa en la oportunidad que ofrece al alumno de examinar rasgos y elementos tanto en su lengua materna como en la lengua extranjera, permitiendo un análisis contrastivo sistemático que le aportará, entre otras cosas, conocimientos sobre las convenciones textuales y estructurales diferentes en ambas lenguas, práctica en el análisis de textos, práctica en la toma de decisiones y la selección tanto de estrategias de traducción como de contextos apropiados a la traducción que tiene en proyecto, etc.

“[...] we think it is important that neither parallel nor comparable corpora be seen as offering ready-made solutions to specific translation problems. [...] This, we believe, is what makes them useful in a learning context” (Gavioli y Zanettin 2000).

Sin embargo, a pesar de las ventajas mencionadas, en la práctica profesional la mayoría del trabajo realizado con corpus especializados como ayudas en traducción ha sido con corpus comparables monolingües en lengua meta, lo que encuentra justificación en el

tiempo y esfuerzo que supone compilar un corpus en cada lengua implicada, especialmente si se trata de hacer una sola traducción.

4.2.3. Corpus comparable monolingüe (o corpus especializado en lengua meta)

Un corpus comparable monolingüe es el compuesto por textos originales en lengua meta y de un grado variable de especialidad. Como así afirman Bernardini, Stewart y Zanettin (2003), este tipo de corpus puede dar información sobre “unidades de significado” típicas de la lengua meta o de una sección especializada de la misma. Puede funcionar como herramienta de referencia y ayudar al traductor o al estudiante a elegir expresiones más naturales, especialmente cuando se traduce hacia la lengua extranjera. Por su parte, Aston (1999: 297 y ss) apunta una serie de ventajas del uso de corpus especializados, algunas de ellas también extrapolables a otros tipos de corpus. Según el autor, en general, un corpus especializado es fácil de manejar, es más informativo que otros recursos, proporciona cifras de densidad léxica y repetitividad en textos del tipo incluido en el corpus, es más homogéneo y por lo tanto se reduce la probabilidad de encontrar concordancias “extrañas”, proporciona más ayuda a la hora de formular hipótesis de traducción, facilita el análisis relacionado con la macroestructura textual, proporciona un medio útil de aprender sobre un área en la que el traductor necesita trabajar y sus convenciones textuales y, por último, es más probable que el aprendizaje casual en un corpus especializado sea relevante para la tarea en la que se está trabajando o, si no, al menos puede ser de utilidad para próximas traducciones de naturaleza similar.

Bowker y Pearson (2002: 198 y ss) han identificado otros usos complementarios para el corpus monolingüe. Según las autoras, el traductor puede utilizarlo para identificar términos importantes o para verificar su intuición, es decir, para confirmar una hipótesis o para eliminar un equivalente incorrecto, lo que ya es un paso adelante. Puede ayudar a escoger entre múltiples opciones y sinónimos o pseudo-sinónimos al ofrecer información contextual y estadística que facilita la distinción de usos y significados. Permite investigar el uso y el estilo (p.ej. fórmula de tratamiento o uso de las contracciones en inglés en un género determinado) observando las colocaciones y las frecuencias. Permite buscar contextos explicativos allá donde se necesite aprender sobre los conceptos propios del campo de especialidad (p.ej. se pueden utilizar patrones de búsqueda como “es un/a”, “es el/la”, “tipo/clase de”, “parte de”, “hecho de”, “consiste en”, “utilizado para/en”, etc.). Por último, con un corpus monolingüe también

se pueden buscar equivalentes “desconocidos”, para lo que será necesario aplicar determinadas estrategias que se explican más adelante (cfr. infra 4.2.5.).

Apuntada la idea por Aston (1999) y llevada a la práctica por Pearson (1999) y, de una forma más elaborada, por Bowker (1999) los corpus monolingües también se han utilizado para la evaluación de traducciones por parte de los docentes en el contexto de la enseñanza de la traducción. En su artículo, Pearson comenta sobre el reto que supone para el profesor de traducción especializada ser capaz de justificar sus recomendaciones o correcciones, lo que, en el caso concreto que describe, le hace recurrir a la construcción de un corpus comparable en lengua meta (inglés, lengua nativa) compuesto por artículos extraídos del *New Scientist* para comparar las elecciones hechas por los estudiantes en su traducción de un texto de una revista científica popular, *La Recherche*, con aquellas elecciones hechas por autores duchos en la materia. El estudio de Pearson le lleva a confirmar que un corpus comparable le proporciona al profesor las pruebas que necesita para apoyar sus decisiones, si bien es cierto que es ésta una forma de evaluación que requiere más tiempo que otros métodos tradicionales, lo que a su vez le conduce a sugerir que sean los propios estudiantes los que evalúen sus traducciones utilizando el recurso descrito. Respecto a la evaluación de la calidad en traducción, Bowker apuesta igualmente por un recurso que ayude al profesor de traducción no sólo a corregir, sino también a proporcionar *feedback* constructivo. En este contexto, la autora describe el diseño de un corpus de evaluación en el campo del *I2O*, un nuevo sistema operativo. *EvalCorp*, que así se llama el corpus de evaluación diseñado por Bowker, contiene tres subcorpus principales, *QualCorp*, *QuanCorp* e *InapCorp*. *QualCorp* es un subcorpus de alta calidad que contiene textos seleccionados cuidadosamente, sobre todo por su contenido conceptual, y cuya función principal es la de permitir al profesor adquirir conocimiento sobre el tema. *QuanCorp* es un subcorpus diseñado para proporcionar una muestra más grande y representativa de la lengua especializada en cuestión y cuya función es la de permitir al profesor verificar si las elecciones terminológicas, fraseológicas y estilísticas del alumno son apropiadas. Y, finalmente, *InapCorp*, que es un subcorpus de textos paralelos inapropiados que difieren del TO en términos de tipo de texto, fecha de publicación, etc. con lo que su función principal es permitir al profesor mostrar pruebas concretas de la falta de adecuación de las elecciones del estudiante. Sobre los resultados del estudio, la autora comenta que, aunque probada a pequeña escala, esta aproximación a la evaluación de traducciones es prometedora y permite ofrecer un tipo de *feedback* ante el que los estudiantes se muestran más receptivos.

4.2.4. Corpus ad hoc, virtual, desechable o de fabricación propia

La terminología propia del campo de la metodología de corpus aún no está firmemente establecida, lo que en ocasiones conduce a solapamientos en las denominaciones. En este caso, tanto los corpus comparables monolingües como los bilingües pueden verse también como “corpus ad hoc” si atendemos a la naturaleza de su creación.

“La compilación de un corpus ad hoc persigue una finalidad específica: reunir en el mínimo espacio de tiempo la mayor cantidad de documentos concretos. Este tipo de corpus será, por definición, eminentemente desequilibrado y limitado en cuanto a su extensión, pero, al mismo tiempo, tremendamente homogéneo” (Corpas Pastor 2001: 164).

En esta publicación Corpas Pastor habla de un *corpus ad hoc de endocrinología* conocido como *CAHDE*, monolingüe, especializado, documentado y en formato sólo texto, que utiliza para la enseñanza de la traducción inversa al inglés, en este caso, trabajando sobre un artículo científico escrito en español. Al igual que Aston (2000b), que defiende el papel de los corpus antes, durante y después de la traducción, la autora aboga por el uso de un corpus como el *CAHDE* tanto en la fase de pretraducción como en la de redacción y revisión (Corpas Pastor 2001: 167 y ss). De esta forma, antes de empezar a traducir, el corpus puede servir para ilustrar las características de los textos científico-técnicos, así como para ayudar al alumno a familiarizarse con este tipo de sublenguaje (términos y pseudo-términos, fraseología, abreviaturas, siglas y acrónimos, falsos amigos, concisión sintáctico-estilística del artículo médico especializado, normas textuales y culturales de estos artículos, etc.). Después, en las fases posteriores, el corpus también ofrece grandes ventajas para el alumno de traducción.

“En esta fase de redacción técnica, el corpus constituye un aliado indiscutible: a) ayuda al alumno a decidirse entre distintos equivalentes de traducción posibles; b) le ofrece modelos lingüísticos, estilísticos y discursivos a modo de guía; y c) arroja luz sobre aspectos culturales y las normas textuales que determinan la producción de este tipo de textos en la cultura meta” (ídem.: 169).

Para Varantola (2003), los corpus ad hoc comparables se caracterizan, además, porque no tienen como objetivo principal formar parte de un corpus permanente, sino que son “virtuales” en el sentido de Ahmad, Holmes-Higgin y Abidi (1994) y desechables,

concepto éste último que ha dado lugar a una nueva denominación de corpus, “el corpus desechable”.

“These [disposable corpora] are small, specialized corpora created ad-hoc to serve the needs of the translator for a specific translation project, and their value lies not only in their analysis but even more so in their creation” (Zanettin 2002: 239).

Respecto a la creación de corpus, Maia (1997, 2000) es una de las autoras que más énfasis pone en los beneficios que la construcción de un corpus propio puede traer al proceso educativo. También en nuestra opinión la búsqueda de textos apropiados para su inclusión en un corpus ad hoc que coincidan en tema, género, registro, etc. con el texto original, que provengan de fuentes cuya fiabilidad habrá de ser evaluada por el alumno, y cuya localización se encuentre quizás en diferentes recursos es una actividad susceptible de facilitar al alumno una oportunidad de aprender o de mejorar estrategias de documentación y análisis textual, desarrollar su espíritu crítico o capacidad de evaluación y ampliar sus conocimientos sobre el tema y género en cuestión.

En esta línea, y para ayudar en el proceso de construcción de un corpus especializado en lengua meta, Bowker (1996) proporciona una relación de criterios y requisitos que debe tener un corpus de estas características si se pretende que éste sea útil para la traducción especializada.

Criterios	Requisitos para la traducción especializada
Tamaño del corpus	Amplia selección de textos (para saber si el uso está establecido o es idiosincrásico).
Tamaño de los textos	Textos completos (para que ejemplos de uso o explicaciones de conceptos no estén cortados).
Tipo de textos	Mezcla de textos instructivos, de expertos y populares (para ayudar a los traductores a comprender el tema y ver diferentes registros de uso).
Fecha de publicación	Sobre todo textos recientes, pero también algunos textos más antiguos (los antiguos son útiles porque los conceptos están mejor explicados cuando aparecen por primera vez; los nuevos se necesitan para reflejar el estado actual del campo temático).
Autor	Textos de una variedad de autores (para determinar si el uso está establecido o es idiosincrásico).
Lengua	Preferiblemente textos escritos por autores que escriben en su lengua materna (para mostrar el uso idiomático).
Cultura	Textos escritos por autores con diferente bagaje cultural, p.ej. británico/norteamericano (para mostrar un uso regional apropiado).

Tabla 2. Requisitos de un corpus para su uso en la traducción especializada (traducido de Bowker 1996).

A pesar de las recomendaciones ofrecidas sobre la planificación de un corpus, suele darse el hecho de que si éste ha sido construido de forma individual o en el contexto de una clase de traducción, es decir, por estudiantes aún en formación, o si toca un tema o género en los que es difícil acceder a documentos relevantes, lo más probable es que dicho corpus sea también desequilibrado o poco representativo. Sin embargo, Atkins *et al.* (1992) no ven en este hecho un obstáculo para trabajar con corpus, sino un dato que debe tenerse en cuenta a la hora de interpretar los datos extraídos.

“...we have found any corpus – however “unbalanced” – to be a source of information and indeed inspiration. Knowing that your corpus is unbalanced is what counts. It would be short-sighted to wait until one can scientifically balance a corpus before starting to use one, and hasty to dismiss the results of corpus analysis as “unreliable” or “irrelevant” simply because the corpus used cannot be proved to be balanced” (Atkins *et al.* 1992: 6).

En otro orden de cosas, si bien es cierto que no han de desdeñarse los recursos documentales convencionales en papel, la compilación de un corpus electrónico ad hoc monolingüe o bilingüe se simplificará si los textos paralelos seleccionados ya se encuentran en formato electrónico, como es el caso de aquellos contenidos en CD-ROM especializados o en Internet. A este respecto, Varantola (2003) nos recuerda que Internet nos proporciona un acceso ilimitado a textos electrónicos, por lo que el problema ya no está en la disponibilidad de datos en este formato, como así era el caso años atrás; ahora el problema, según la autora, está en la relevancia, adecuación, fiabilidad y análisis de este tipo de material. Como comentan Friedbichler y Friedbichler (2000), algunas de las desventajas de Internet es que una parte importante de los documentos que contiene no están editados o no han sido revisados a fondo o son (malas) traducciones, a lo que se podría añadir la escasez de documentos en determinadas lenguas. Según Fletcher (2001), el número de páginas web de acceso público llegó a 4 billones en el 2000, de las cuales el 80% correspondería a páginas en inglés. Sin embargo, un estudio del mismo autor revela que la cantidad de usuarios de Internet cuya primera lengua no es el inglés está aumentando, lo que se refleja en que el factor de crecimiento de las páginas en otros idiomas es más alto que el de las páginas en inglés. Por ejemplo, el número de páginas en español fue el que más aumentó entre el 1996 y el 2000, pasando de 104 a 1894 millones de palabras, por delante de las páginas en alemán, francés, italiano, portugués, noruego, finlandés e inglés, si bien es cierto que el inglés pasó de contar con 6.082 millones de palabras en Internet en 1996 a 48.064 millones en el 2000. Una publicación más reciente, la de O'Neill *et al.* (2003), muestra que el español contaba en el año 2002 con un 3% de

representación en Internet en lo que a páginas públicas se refiere, sólo detrás de otras lenguas como el inglés, el alemán y el japonés.

La relativa facilidad con la que un usuario puede crear un corpus extrayendo documentos de Internet ha provocado también la acuñación de otro término, “el corpus DIY” (Zanettin 2002). Un corpus de fabricación propia, como podríamos llamarlo, es, según el autor, un conjunto de páginas web en formato HTML creado ad hoc para la traducción de un texto en concreto; es un corpus abierto, es decir, se puede añadir más material si se considera necesario; es desechable (Varantola 2000) o virtual (Ahmad, Holmes-Higgin y Abidi 1994) ya que no pretende formar parte de un corpus permanente; no es necesario solicitar licencias de derechos de propiedad intelectual y puede ser bilingüe o monolingüe en lengua meta.

Por otra parte, Internet puede funcionar no sólo como fuente de documentos para la creación de un corpus, sino también como corpus mismo. En este caso estaríamos hablando de un macro-corpus con infinidad de documentos de muy diverso tipo al que tendríamos que abordar bien a través de motores de búsqueda generales (p.ej. *Google* o *Copernic*) bien a través de otras aplicaciones especialmente diseñadas para las necesidades profesionales de la lengua (p.ej. *WebCorp* o *KwicFinder*) (Zanettin 2002). Estas herramientas facilitarán la tarea al alumno de traducción en tanto que agilizarán sus búsquedas documentales, por ejemplo permitiéndole limitar el dominio geográfico de las páginas que le interesan (p.ej. .es, .uk, .fi), limitar sus búsquedas por palabras clave o disminuir el ruido en los resultados eliminando palabras que no sean de interés para la búsqueda.

En este sentido, aunque el trabajo con textos en formato electrónico y la disponibilidad de herramientas informáticas para buscar y analizar estos textos permitan al alumno o usuario compilar corpus extensos, es menester recordar que el tamaño del corpus, en este caso, no ha de ser un criterio absoluto de calidad. La importancia de la construcción y trabajo con corpus en didáctica de la traducción no ha de residir en la mera acumulación de textos con el fin de obtener un corpus grande, sino en hacer que todo el trabajo se enmarque en un proceso educativo en el que los alumnos perciban que su esfuerzo tiene sentido, que está bien orientado y que además tiene una recompensa de aprendizaje; en otras palabras, que el uso de corpus es beneficioso.

“La compilación de corpus *ad hoc* para la docencia en traducción e interpretación resulta especialmente útil, por cuanto supone una fuente de documentación rápida, económica y fiable para conocer la materia, las unidades terminológicas y fraseológicas de un ámbito especializado dado, así como el índice de variación formal y conceptual de las unidades terminológicas de la especialidad de acuerdo con la situación comunicativa (cf. Cabré Castellví, 1999). En este sentido, el corpus *ad hoc* constituye una *macrofuente* de documentación, en tanto fuente gramatical y discursiva, lexicográfica, terminológica y cognitiva-especializada” (Corpas Pastor 2001: 173).

Como podemos observar, los corpus comparables tienen una gran riqueza potencial en lo que a información se refiere, siendo la forma de llegar a ella la cuestión que queda pendiente. Trabajos como el de Sánchez-Gijón (2003) han venido a ofrecer, entre otras cosas, vías para extraer de los corpus comparables información tanto conceptual como factual. Para Bowker (2000: 22), al utilizar un corpus monolingüe en lengua meta, el mayor reto está en diseñar un método para buscar términos desconocidos allá donde el término en lengua original no se puede usar como punto de acceso.

4.2.5. Estrategias de búsqueda de equivalentes de traducción en un corpus especializado monolingüe en lengua meta

A continuación se recogen algunas de las técnicas o estrategias que hemos encontrado más útiles con el fin de, por un lado, llegar a crear un corpus especializado en lengua meta y, por otro, extraer información relevante para la traducción. Para ilustrar algunas estrategias, se ofrecen ejemplos extraídos de nuestra experiencia docente, tanto de traducción directa como inversa.

OBJETIVO	ESTRATEGIA	EXPLICACIÓN DETALLADA
<p>Obtener términos clave en lengua meta para poder buscar TPs y construir un corpus especializado (p.ej. un corpus de fabricación propia).</p>	<p>Utilizar la terminología y el vocabulario presentes en el TO.</p>	<p>Extraer del TO una lista de palabras por orden de frecuencia; sacar concordancias o agrupaciones de aquellos términos frecuentes o clave. Observar el cotexto de los términos del TO, identificar las palabras de cuya traducción se puede estar seguro e introducir estas palabras en un buscador especializado de textos (p.ej. quizás no se encuentre fácilmente la traducción al inglés del término “habón”, pero en el TO donde aparece puede que también se halle información relacionada con él, como que es una lesión de la piel, con relieve y relacionada con otros términos como “edema”, “urticaria” o “hipersensibilidad”. La traducción de estos últimos términos es clara y se podría introducir como búsqueda para localizar TPs que hablen del “habón” en inglés).</p>
		<p>Hacer una primera búsqueda de TPs (mismo género o tema que el TO) pero en lengua meta; después, extraer las palabras clave de los textos descargados para obtener nuevas palabras clave y hacer una segunda búsqueda de textos.</p>
	<p>Recopilar los textos en LM citados en la bibliografía del TO, si viene incluida.</p>	<p>Recoger los documentos en lengua meta sobre los que está basado el TO porque ahí se encontrará la terminología y fraseología claves e incluso citas completas en la lengua meta.</p>
<p>Fijarse en la bibliografía del TO, si viene incluida.</p>	<p>Utilizar palabras o frases en lengua meta que aparecen en los títulos de la bibliografía para hacer búsquedas de otros textos para resolver las primeras dificultades conceptuales o terminológicas.</p>	

Identificar términos importantes.	Fijarse en los términos más frecuentes en el campo.	Extraer una lista de palabras por orden de frecuencia para hacerse una idea de las palabras más comunes o importantes en el campo de especialidad.
	Fijarse en los términos propios del campo.	Extraer una lista de palabras clave comparando el vocabulario del corpus especializado con otro corpus general para llegar a saber cuál es el vocabulario propio del campo de especialidad.
Investigar una hipótesis lingüística o conceptual.	Utilizar los cognados.	Extraer concordancias de palabras parecidas morfológicamente en la lengua meta (p.ej. “stipulate” y “estipular”).
	Combinar la palabra de hipótesis con otras que están en el contexto.	Extraer concordancias introduciendo como búsqueda un término traducido del que se esté seguro en combinación con otro para llegar al equivalente de un término compuesto (p.ej. “trim kit: dimensions” → se podría buscar “kit” con “dimensiones” como palabra de contexto).
Seleccionar entre múltiples opciones.	Probar variantes ortográficas.	Extraer concordancias de una palabra tanto con guión como sin guión, o en mayúsculas y minúsculas (p.ej. introducir la búsqueda “real*time ultrasonography” para comprobar que aparece más frecuentemente con guión).
	Utilizar palabras de contexto en la búsqueda.	Extraer concordancias de dos palabras o expresiones que parecen ser sinónimos introduciendo una de ellas como palabra de contexto para comprobar si ambas pueden aparecer en el mismo contexto y cómo se usan.
	Utilizar el asterisco para truncar las búsquedas.	Extraer concordancias introduciendo una búsqueda truncada para comprobar las diversas posibilidades y si hay diferencias entre ellas (p.ej. buscar “chromosom*” para observar los casos de “chromosome/-al/-ic aberration”).

	Comprobar el uso de sinónimos y pseudo-sinónimos.	Extraer concordancias de sinónimos o lo que pueden parecer sinónimos para comprobar si significan exactamente lo mismo, funcionan igual o mantienen una relación jerárquica (p.ej. “stock exchange” y “bourse” son sinónimos en cuanto a significado pero no son 100% intercambiables).
	Comprobar el orden de palabras.	Combinar las palabras de búsqueda cambiando su orden, o introduciendo una palabra como búsqueda de contexto.
Localizar fraseología propia del género textual.	Fijarse en las agrupaciones.	Extraer las agrupaciones de palabras o expresiones frecuentes o propias del género (p.ej. extraer las agrupaciones con “información” en un corpus de manuales de instrucciones nos dará frases como “Para más/mayor información... // Si desea/s / necesita/s más información, contacte/contacta con ...”).
Determinar el registro.	Fijarse en las palabras más frecuentes.	Extraer una lista de palabras por orden de frecuencia para observar los pronombres personales (de sujeto y posesivos) con el fin de identificar el trato más frecuente que se da al lector en un determinado género (p.ej. “usted”, “tu”/“su”); fijarse también en la persona de los verbos (p.ej. “encienda”) y elementos que nos den la clave (p.ej. “si” condicional para observar casos como “si se introduce // si introduce/s”) para luego confirmar el uso extrayendo concordancias de esos elementos.
Investigar la ideología y la prosodia semántica.	Observar el cotexto y contexto de la palabra o frase que se quiere estudiar.	Extraer concordancias de la palabra o frase para observar si lo que rodea a una palabra clave en una concordancia tiene connotaciones positivas o negativas (p.ej. “reconversión”, aunque significa modernización, suele implicar despidos laborales). Extraer las colocaciones de la palabra o frase para observar si el conjunto de palabras que rodea a una palabra clave tiene connotaciones positivas o negativas.

Recuperar posibles equivalentes para términos “desconocidos”.	Fijarse en la bibliografía del TO, si viene incluida.	Extraer concordancias de palabras que aparecen en la bibliografía en lengua meta y observar si en el contexto aparecen otras palabras relacionadas que sean posibles equivalentes a otros términos del TO o al menos útiles para nuevas búsquedas.
	Introducir búsquedas truncadas con asterisco.	Realizar búsquedas truncadas con asterisco (p.ej. introducir “prefijo/raíz de la palabra + asterisco”, como en el caso de buscar “pre*” en un corpus de textos médicos para saber si se suele hablar de “prebirth” o “prenatal”; también se puede introducir “asterisco + última cifra de un número”, como en el caso de buscar “*oo” en un corpus de textos técnicos sobre electrodomésticos para saber si en español se habla de “X Watt” / “X Watts” / “X vatios” / “X W.”) para expresar la potencia).
	Introducir como búsquedas términos conocidos o partes de términos que se encuentren en el contexto.	Extraer concordancias, colocaciones o agrupaciones de términos conocidos o partes de términos que se encuentren en el contexto, como buscar una parte de un término de la que se esté seguro (p.ej. para el término “longitud cráneo-nalga”, en inglés médico, se puede buscar “length” y observar los patrones que aparezcan; se puede estar más seguro del equivalente de “longitud” que de los de “cráneo” o “nalga”).
	Utilizar el conocimiento sobre estructuras sintácticas.	Buscar pensando no en palabras, sino en la posible estructura equivalente (p.ej “ * a * ” → “ * by * ” : “paso a paso” → “step by step”).
	Buscar abreviaturas, siglas o acrónimos y su expansión.	Buscar el acrónimo, sigla o abreviatura o introducir como búsqueda alguna de las palabras de su forma extensa; en algunos casos se mantienen igual que en el original. Otra fórmula es buscar paréntesis en el corpus.
	Buscar nombres propios y de productos.	Buscar los nombres tal cual ya que, dependiendo del género, es probable que se mantengan en el idioma original.

Tabla 3. Estrategias de búsqueda de información en un corpus monolingüe en lengua meta.

4.3. Pruebas de eficacia

Aparte de los procesos de observación, análisis y generalización propios del trabajo con corpus, la no muy extensa bibliografía en el campo muestra la falta de una metodología establecida en lo que se refiere al uso de corpus en la enseñanza de la traducción, probablemente debido a la breve trayectoria que tiene esta aplicación. Entre los trabajos existentes, la mayoría explican casos anecdóticos de puesta en práctica de ejercicios en el aula y resultados obtenidos, pocos teorizan o realizan experimentos encaminados a aportar pruebas empíricas de los beneficios del uso de corpus y establecer una metodología y muy pocos comentan explícitamente sus implicaciones pedagógicas. Sin embargo, prácticamente todos los artículos y comentarios respecto al uso de corpus recogen aspectos positivos.

“In this climate of empirical uncertainty, arguments for and against the use of corpora in translator training must be of a theoretical nature, and can resort at best to anecdotal evidence. Where available and accessible, appropriate corpora appear able to provide better and faster solutions to many of the translator’s problems in a unified environment, with positive effects on learning. They make possible more idiomatic, native-like interpretations of source texts and a use of more idiomatic, native-like strategies in target texts” (Aston 1999: 310).

Entre los escasos estudios empíricos existentes considerados como experimentos por sus autores destacan sobremanera los de Bowker (1998; 2000) y, después, si cabe, los de Zanettin (2002) y Varantola (2003), si bien es cierto que estos últimos se podrían considerar más bien actividades con corpus. El primero de los experimentos o, en palabras de la autora, estudio piloto, hipotetiza que un corpus especializado monolingüe en la lengua meta puede ser un recurso útil en la formación en traducción directa, especialmente en lo que atañe a la comprensión del tema y la competencia lingüística nativa especializada. El experimento, que consistió en la traducción del francés al inglés de dos extractos de un artículo sobre escáneres ópticos, fue llevado a cabo por 14 estudiantes, la mitad de los cuales tradujo un texto 1 con recursos convencionales y un texto 2 con diccionarios bilingües generales, un corpus compilado a base de textos extraídos del CD-ROM *Computer Select* y el programa *WordSmith Tools*, y la otra mitad de los estudiantes al contrario. Los resultados revelaron que las traducciones hechas con ayuda del corpus fueron de mayor calidad respecto a la comprensión del TO, correcta elección de terminología y expresión idiomática. Los comentarios realizados por los participantes en el experimento confirmaron la falta de satisfacción con los recursos convencionales y la ayuda que el corpus les había

prestado, a pesar del tiempo que requirió su consulta para algunos⁶. Más adelante, la misma Bowker (2000) realiza un experimento similar con sólo dos participantes que hubieron de traducir un texto, cada uno hacia su lengua materna (del francés al inglés), sobre el sistema operativo I2O. De los participantes, uno tradujo un texto A ayudándose de recursos convencionales, y el otro un texto B sirviéndose de recursos lexicográficos convencionales, un corpus especializado en lengua meta de 150.000 palabras también extraídas del *Computer Select* y el *WordSmith Tools*. Al acabar la traducción, se entrevistó a los dos estudiantes y se les pidió que comentasen las elecciones de traducción que habían hecho, las estrategias que habían utilizado y el valor de los recursos que tenían a su disposición, datos con los cuales la autora llegó a las mismas conclusiones que en el experimento anterior. Zanettin (2002), por su parte, comenta un experimento de características muy abiertas consistente en la construcción de un corpus a partir de Internet (un corpus de fabricación propia o DIY) para el que no se establecieron límites de lenguas, tipo o número de textos. Se animó a los estudiantes a trabajar individualmente o en grupo, a utilizar herramientas que ya conocían (paquete Trados, *Déjà Vu*), añadir más material al corpus si así lo creían conveniente y se les pidió redactar un informe sobre las ventajas y desventajas de crear y utilizar un corpus DIY como recurso de traducción. Dichos informes arrojaron información sobre cuestiones tales como la relevancia de los textos y su fiabilidad, desventajas del *WordSmith Tools*, tendencias en las búsquedas, que principalmente fueron de información terminológica, fraseológica y colocacional, inclinaciones a la hora de tomar decisiones como por ejemplo al basarse en los datos de frecuencia cuando existían varios candidatos en la traducción o al crear sólo un corpus en lengua meta, prescindiendo de un corpus en lengua original y, finalmente, opiniones que reiteran la idea de que construir un corpus y analizarlo requiere tiempo, si bien es cierto que aporta más seguridad a la hora de escoger una solución y que la utilidad del corpus es más obvia con encargos más voluminosos. Varantola (2003), por último, explica de forma somera un experimento que pone el énfasis en la definición de criterios de búsqueda de textos relevantes para un corpus con el fin de que los estudiantes lleven a término un encargo real de traducción entre finlandés e inglés. Igual que en el caso anterior, tal actividad proporcionó información sobre cuestiones de accesibilidad y fiabilidad de los textos, además de reflexiones sobre estrategias de compilación y el establecimiento de los grupos meta o grupos receptores de las traducciones.

⁶ La descripción completa del experimento se encuentra en Bowker (1998: 633 y ss).

4.4. Situación actual y necesidades

Si se observa la bibliografía perteneciente a esta sección se puede deducir que la aplicación de corpus electrónicos a la enseñanza de la traducción no tiene aún muchos años de trayectoria. Aún hoy artículos y libros sobre el tema incluyen tipologías de corpus, definiciones y aclaraciones al respecto de la terminología utilizada, lo que es índice de la juventud de la aplicación, aunque también es cierto que muchas de estas publicaciones se dirigen a un público no necesariamente experto en metodología de corpus al que se le quieren mostrar los buenos resultados de su aplicación. Por otro lado, algunos artículos de los 90 en los que se nos explican las dificultades de acceso a Internet, la escasez de documentos en formato electrónico o las limitaciones de capacidad de los ordenadores y los programas de gestión de corpus, parecen aún más antiguos de lo que son en realidad. Sin embargo, esto no significa que haya que olvidar que autores como Lindquist ya utilizaban corpus para estudiar el diseño de manuales de traducción en 1983, o que, como nos recuerda Corpas Pastor (2002), en los congresos de EURALEX ya se señalaban a principios de los 90 las deficiencias de los diccionarios y se reclamaba el uso de corpus efímeros pequeños extraídos de fuentes electrónicas, principalmente de Internet, para la realización de encargos concretos de traducción. Asimismo, al término del conocido proyecto LETRAC (Language Engineering for Translator Curricula), el grupo de expertos participantes en el estudio diseñó un currículo relacionado con las tecnologías de la lengua para adecuar la formación del traductor a las exigencias del mercado. Dicho currículo, diseñado en 1999, ya contemplaba contenidos de lingüística de corpus en cuanto a conocimientos y habilidades básicas de creación, preparación, análisis y explotación de corpus. De cualquier modo, lo que es innegable es el creciente interés que la aplicación de corpus tiene en el campo de la traducción y su enseñanza, de lo que da fe principalmente la serie de conferencias CULT (Corpus Use and Learning to Translate), celebradas en Bertinoro en 1997 y 2000 y en Barcelona en 2004.

Por último, la bibliografía sobre el tema también arroja que los corpus se han empleado casi por igual en la enseñanza de la traducción inversa (Zanettin 1998, 2001, 2002; Gavioli y Zanettin 2000; Corpas Pastor 2001, 2002; Corpas Pastor y Postigo Pinazo 2002; Varantola 2003) y de la traducción directa (Piotrowska 1997; Bowker 1998; Pearson 1999; Maia 2000; Rodríguez López 2002; Zanettin 2002; Kübler 2003; Varantola 2003). En este sentido, el presente trabajo hace aportaciones tanto en una dirección como en otra y a distintos niveles de especialidad, utilizando todos los tipos de corpus más relevantes para la traducción aquí descritos e integrándolos en tareas

diversas con objetivos variados en los que el fin principal no es utilizar el corpus para solucionar problemas de traducción, sino para aprender a traducir reflexionando.

“[...] however paradoxically, corpora can and should be employed to problematise rather than simplify the task of (future) translators. The greatest pedagogic value of the instrument lies, we suggest, in its thought-provoking, rather than question-answering, potential” (Bernardini, Stewart y Zanettin 2003).

La necesidad principal que se presenta en esta área es, a nuestro parecer, la aplicación de una metodología didáctica clara basada en el estudiante y organizada en tareas de traducción que sistematice el uso de corpus en la enseñanza de la traducción y conduzca al desarrollo de la competencia traductora. Presentamos una propuesta para rellenar este vacío en la *Parte II* y la *Parte III* de esta tesis.

PARTE II.

MARCO CONCEPTUAL PARA LA UTILIZACIÓN DE LA METODOLOGÍA DE CORPUS EN LA FORMACIÓN DE TRADUCTORES

“... the pedagogically most suitable key-concepts are those associated with experiencing, exploration, and discovery...”.

(Toury 1992: 69)

“La sociedad del conocimiento es también la sociedad del aprendizaje, que demanda aprendizaje a lo largo de toda la vida.

En este contexto, las personas necesitan ser capaces de utilizar el conocimiento, de actualizarlo, de seleccionar lo que es apropiado para un contexto específico, de aprender permanentemente, y de entender el potencial de lo que aprenden, de tal forma que puedan adaptar el conocimiento a nuevas situaciones que se transforman rápidamente”.

(Yániz y Villardón 2006: 12)

5. LA DIDÁCTICA DE LA TRADUCCIÓN ESCRITA

5.1. La competencia traductora y su adquisición

5.1.1. *La competencia traductora*

El objetivo global de la enseñanza de la traducción es desarrollar una competencia traductora (CT) en el estudiante. Así, dado que lo que se pretende con este trabajo es facilitar el desarrollo de la competencia traductora, cabe dedicar un apartado a lo que hasta la fecha se ha realizado en cuanto a la investigación sobre este tema y cómo, ulteriormente, los resultados de tal investigación podrían aplicarse a la didáctica de la traducción en la forma de mejores planes de estudio, métodos de evaluación y criterios pedagógicos unificados.

Los estudios sobre la competencia traductora o su adquisición son muy limitados en tanto que la mayoría de modelos de la CT no están basados en datos empíricos, a lo que se suma la falta de estudios longitudinales que arrojen luz sobre el proceso de adquisición. La mayoría de estos modelos de la CT propuestos se centran en describir sus componentes, entre los cuales suelen estar presentes los conocimientos lingüísticos, textuales, temáticos, culturales, de documentación, capacidad de transferencia, etc. y, sólo en algunos modelos, el componente estratégico. Algunos modelos existentes del funcionamiento de la competencia traductora son, entre otros, Lowe (1987), Bell (1991), Hewson y Martin (1991), Nord (1991, 1992), Pym (1992), Kiraly (1995), Presas (1996), Hurtado (1996b), Hatim y Mason (1997), Hansen (1997), Risku (1998), Neubert (2000), Kelly (2002), y Gonçalves (2003, 2005). Para la traducción inversa, las propuestas sobre la competencia traductora son incluso más escasas (Beeby 1996, Campbell 1998, Martínez Melis 2001).

Kiraly habla de la distinción entre “translation competence” y “translator competence”. La primera comprendería “[...] specific skills that allow one to produce an acceptable target text in one language on the basis of a text written in another” (2000: 13), y la segunda implicaría “[...] being able to use tools and information to create communicatively successful texts that are accepted as good translations within the community concerned” (ídem: 14), es decir, los aspectos estratégico, comunicativo, instrumental y profesional de un todo que es la competencia traductora. Incluso dentro

de lo profesional, Kiraly va más allá y distingue entre experticidad¹ y profesionalidad² y reivindica más espacio para ésta última, dada su importancia social. Para Kiraly, la mejor manera de conducir a los estudiantes en su aprendizaje es “to provide them with opportunities for interacting with peers and experts to collectively – as well as individually – construct the knowledge of the domain” (idem: 33).

5.1.1.1. La investigación del grupo PACTE

El grupo PACTE (Proceso de Adquisición de la Competencia Traductora y Evaluación), del Departament de Traducció i d'Interpretació de la Universitat Autònoma de Barcelona y al que la autora de esta tesis pertenece desde el año 2000³, destaca con su investigación empírico-experimental sobre la CT. El objetivo principal del grupo es analizar la competencia traductora y su adquisición en traducción escrita.

Según la ha definido PACTE, la competencia traductora:

“is considered to be the underlying knowledge system needed to translate and has four distinctive characteristics: (1) it is expert knowledge and not possessed by all bilinguals; (2) it is basically procedural knowledge; (3) it is made up of various interrelated sub-competencies; (4) the strategic component is very important, as it is in all procedural knowledge” (PACTE 2005: 610).

La competencia traductora es, como decimos, un conocimiento experto y supone una especialización del individuo. Todas las subcompetencias están involucradas en la traducción, con lo que es más acertado hablar de activar, movilizar, incidir, practicar o entrenar distintos aspectos de cada subcompetencia. Además, las subcompetencias se interrelacionan y se compensan entre sí dependiendo del encargo, el texto o las necesidades del traductor. Así, por ejemplo, es de esperar que en la traducción inversa se requerirá activar más la subcompetencia relacionada con la documentación que otras subcompetencias.

1 Según Kiraly (2000: 30), básicamente, la capacidad de producir una traducción de calidad que satisfaga al cliente y respete las normas de la profesión.

2 Según Kiraly (idem: 31), la capacidad de trabajar dentro de los límites sociales y éticos de la profesión.

3 El grupo PACTE se formó en 1997 y desde entonces se le han concedido diversos proyectos financiados: PB96-1125, BFF2001-2510, PHB2002-0010-PC, 2002SGR 00178). PACTE está formado actualmente por A. Beeby, M. Fernández Rodríguez, O. Fox, I. Kozlova, A. Kuznik, W. Neunzig, P. Rodríguez Inés y L. Romero, con Amparo Hurtado como investigadora principal.

El modelo elaborado por PACTE propone que la CT está integrada por cinco subcompetencias y componentes psicofisiológicos:

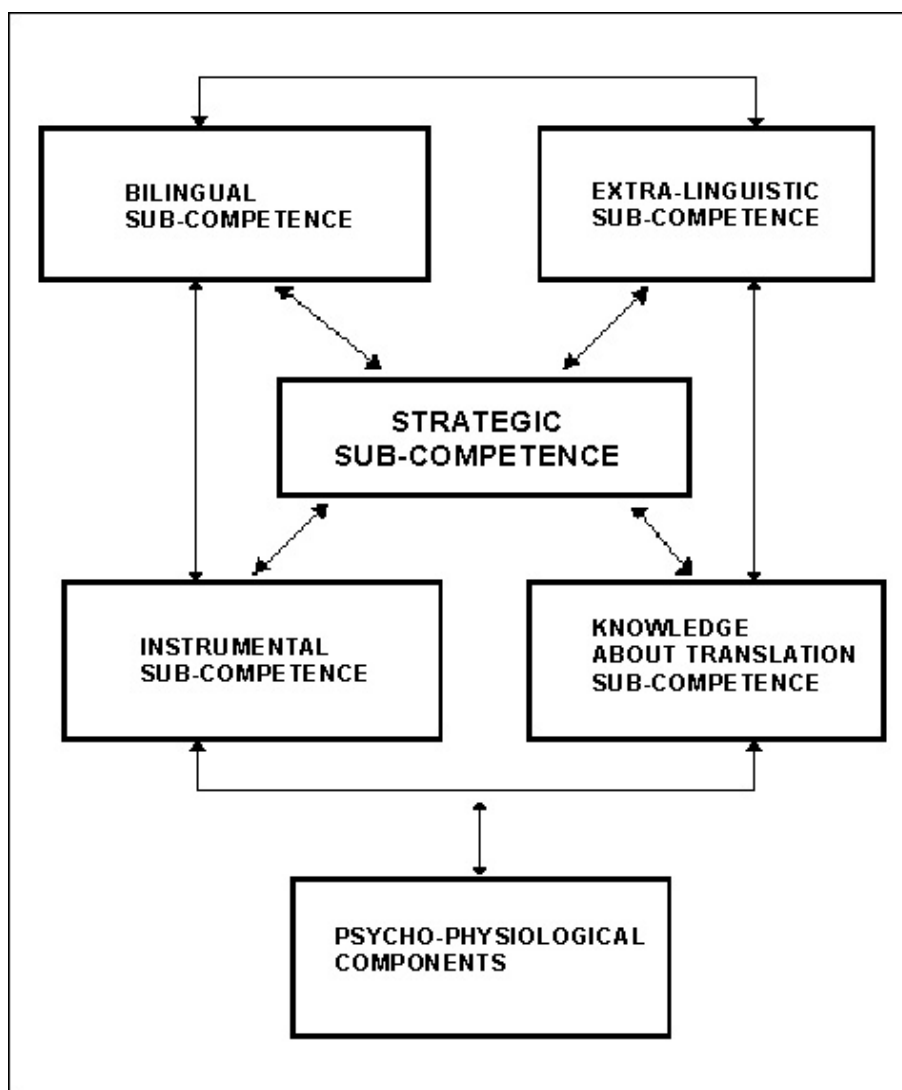


Figura 1. Modelo holístico de la competencia traductora (PACTE 2003: 60).

Las definiciones de las subcompetencias de la CT, según PACTE (2003: 58-59), son:

The bilingual sub-competence. Predominantly procedural knowledge needed to communicate in two languages. It includes the specific feature of interference control when alternating between the two languages. It is made up of pragmatic, socio-linguistic, textual, grammatical and lexical knowledge in the two languages.

Extra-linguistic sub-competence. Predominantly declarative knowledge, both implicit and explicit, about the world in general and special areas. It includes: (1) bicultural knowledge (about the source and target cultures); (2) encyclopaedic knowledge (about the world in general); (3) subject knowledge (in special areas).

Knowledge about translation sub-competence. Predominantly declarative knowledge, both implicit and explicit, about what translation is and aspects of the profession. It includes: (1) knowledge about how translation functions: types of translation units, processes required, methods and procedures used (strategies and techniques), and types of problems; (2) knowledge related to professional translation practice: knowledge of the work market (different types of briefs, clients and audiences, etc.).

Instrumental sub-competence. Predominantly procedural knowledge related to the use of documentation sources and an information and communication technologies applied to translation: dictionaries of all kinds, encyclopaedias, grammars, style books, parallel texts, electronic corpora, searchers, etc.

Strategic sub-competence. Procedural knowledge to guarantee the efficiency of the translation process and solve the problems encountered. This is an essential sub-competence that affects all the others and causes inter-relations amongst them because it controls the translation process. Its functions are: (1) to plan the process and carry out the translation project (choice of the most adequate method); (2) to evaluate the process and the partial results obtained in relation to the final purpose; (3) to activate the different sub-competencies and compensate for deficiencies in them; (4) to identify translation problems and apply procedures to solve them.

Psycho-physiological components. Different types of cognitive and attitudinal components and psycho-motor mechanisms. They include: (1) cognitive components such as memory, perception, attention and emotion; (2) attitudinal aspects such as intellectual curiosity, perseverance, rigour, critical spirit, knowledge of and confidence in one's own abilities, the ability to measure one's own abilities, motivation, etc.; (3) abilities such as creativity, logical reasoning, analysis and synthesis, etc.

Dentro de este modelo, y para el propósito de esta tesis, nos centraremos especialmente en la subcompetencia instrumental. Hablar de trabajo con corpus electrónicos es sinónimo de manejo de documentación e informática y su aplicación en la formación de traductores puede ayudar a mejorar las habilidades de documentación en los estudiantes y hacerles más fácil el uso de las nuevas tecnologías.

5.1.2. La adquisición de la competencia traductora

Los estudios sobre la adquisición de la competencia traductora (ACT) son aún muy escasos (Harris 1973, 1977, 1980; Harris y Sherwood 1978; Toury 1995; Shreve 1997; Chesterman 1997), y los pocos estudios empíricos efectuados no son longitudinales ni estudian la ACT de modo total (Jääskeläinen 1987, 1989; Tirkkonen-Condit 1990;

Jääskeläinen y Tirkkonen-Condit 1991; Seguinot 1991; Lorenzo 1999), con lo que no han aportado mucha luz a la incógnita de cómo se adquiere la CT, bien por defecto de la muestra, bien por el instrumento más utilizado, el *Thinking Aloud Protocol* o *TAP*, bien por otros defectos de diseño del experimento (García Álvarez 2006). Sin embargo, algunas de estas pruebas y las verbalizaciones que incluían sobre el proceso seguido por el estudiante al traducir, dejaron entrever la linealidad de su procesamiento y una total tendencia a centrarse únicamente en lo “microtextual”. En este sentido, estamos de acuerdo con la autora citada en cuanto a que, en parte, las argumentaciones de los estudiantes dejan traslucir la influencia de unas metodologías docentes que fraccionan el saber y no lo integran, que se centran en lo pequeño y eluden la visión general.

El grupo PACTE ha desarrollado un modelo de adquisición de la competencia traductora. Si bien es cierto que la investigación actual del grupo se centra en la caracterización de la competencia traductora, el objetivo ulterior de su investigación es la caracterización del proceso de adquisición de ésta misma para poder luego implementar los hallazgos en la formación de traductores.

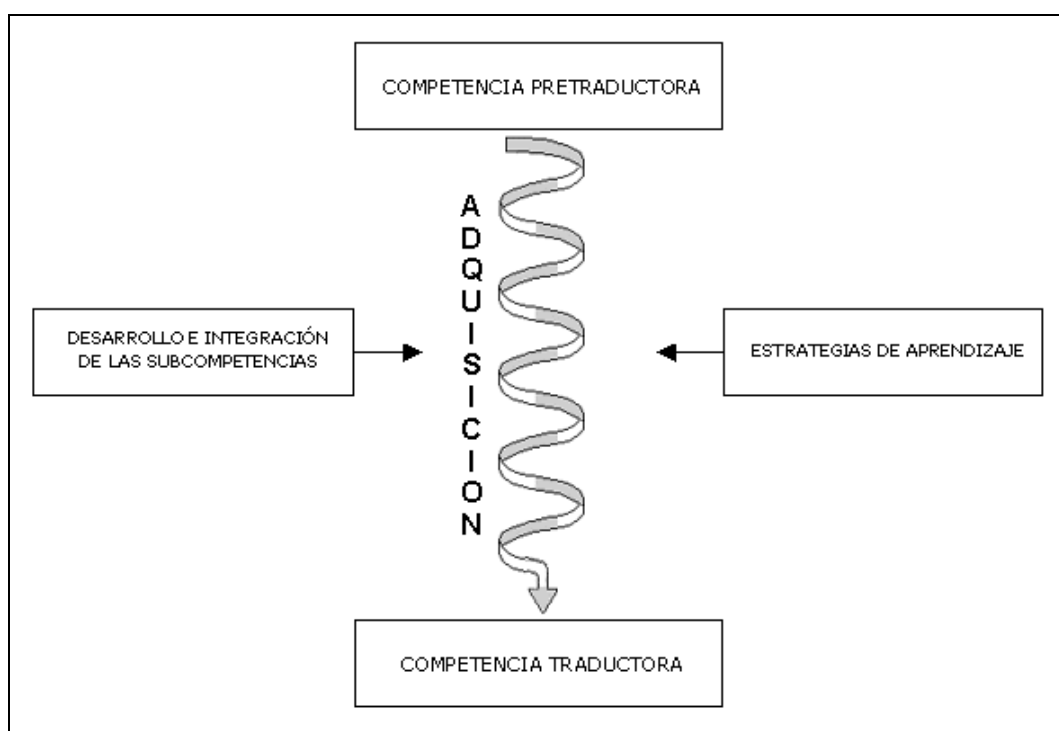


Figura 2. Modelo dinámico de la adquisición de la competencia traductora (PACTE 2000).

Según este modelo, la adquisición de la competencia traductora es un proceso de reestructuración y desarrollo de las diversas subcompetencias de la competencia traductora. Así, la ACT podría definirse como:

- (1) A dynamic, spiral process that, like all learning processes, evolves from novice knowledge (pre-translation competence) to expert knowledge (translation competence); it requires learning competence (learning strategies) and during the process both declarative and procedural types of knowledge are integrated, developed and restructured.
- (2) A process in which the development of procedural knowledge and, consequently, of the strategic sub-competence are essential.
- (3) A process in which the translation competence sub-competencies are developed and restructured.

Definición de la ACT (PACTE 2003: 49-50).

En resumen, el modelo de ACT de PACTE, que suscribimos, hace referencia a un aprendizaje en espiral donde el conocimiento se va reestructurando según se avanza en la adquisición de la CT.

Como ya se ha podido comprobar, la tradición en la investigación en CT y ACT es aún muy corta. Por otro lado, sin embargo, aunque la enseñanza reglada de la traducción como la conocemos hoy en día en Europa se inicia hace unos cincuenta años, esta actividad tiene, obviamente, una historia mucho más larga.

5.2. Traducción pedagógica

El uso pedagógico de la traducción en la didáctica de lenguas merece una mención especial por la gran presencia que ha tenido y tiene aún en las clases de lengua extranjera. Entre los autores que defienden el uso de la traducción como actividad en el aprendizaje de lenguas destaca considerablemente Duff (1989), quien aporta una metodología muy rica y abundantemente ejemplificada; también rechaza la práctica que se da con frecuencia en la enseñanza de la traducción de solicitar al alumno que traduzca un texto sin más instrucción que ésta misma. Duff, adelantándose a enfoques más elaborados que vendrán más adelante, considera que en el diseño de actividades o tareas en las que se utiliza la traducción para aprender una lengua extranjera es importante tener en cuenta cuestiones como preparar al alumno para conseguir el objetivo del ejercicio, establecer o prever el tiempo que el alumno puede invertir en una tarea, seleccionar unos materiales didácticos siguiendo unos criterios claros, o pensar en el alumno como alguien que, individualmente o en grupo, debe estar involucrado en la actividad, y en el profesor como un guía, estimulador o “animador” de la tarea o discusión y no como un espectador pasivo. Sin embargo, hemos de decir aquí que, en nuestra opinión, Duff todavía deja demasiada responsabilidad en manos del profesor

en lo que a conocimientos y evaluación se refiere, y que este peso se podría aliviar de formas que presentaremos más adelante.

Denunciando la falta de una pedagogía actualizada y apropiada de la traducción y que se aleje de la idea de que un traductor no se puede formar, sino que debe nacer con un don especial, está Grellet (1991) que, como Duff, reniega de la situación del momento de la traducción en la enseñanza de lenguas: “Une pédagogie qui (...) peut fréquemment se résumer par les mots « Lisez – Traduisez »” (Grellet 1991: 11). Grellet apuesta por un enfoque comunicativo, unas actividades de traducción más auténticas en el aula y una enseñanza más variada, estructurada y metódica en la que habría una progresión desde los ejercicios de sensibilización hacia la traducción hasta los ejercicios de familiarización con problemas específicos entre lenguas. Es de destacar que en el libro citado cada ejercicio viene precedido de una ficha con el objetivo y los pasos del trabajo de clase.

5.3. Enfoques de la didáctica de la traducción

Entre las corrientes, entendidas de forma general, que han existido en la didáctica de la traducción, podemos listar las siguientes:

- La didáctica tradicional de la traducción.
- Los estudios contrastivos.
- Los tratados esencialmente teóricos.
- La enseñanza por objetivos de aprendizaje.
- El constructivismo social.
- El enfoque por tareas de traducción.
- La formación por competencias.

La obra de Hurtado (1999) nos sirve de apoyo para hacer un breve repaso a las cuatro primeras corrientes mencionadas.

5.3.1. La didáctica tradicional de la traducción

Relacionada directamente con el uso de la traducción en la didáctica tradicional de lenguas, la didáctica tradicional de la traducción gusta de centrarse en problemas sólo de tipo lingüístico y de plasmarse en manuales en los que la única consigna metodológica es la del “lea y traduzca”. Los rasgos que caracterizan esta didáctica son:

- Identificación de la didáctica de lenguas y de la didáctica de la traducción ya que la clase de traducción sirve para ampliar conocimientos de la lengua.
- Concepción literalista de la traducción ya que está basada en la transcodificación de elementos lingüísticos, razón por la que en este tipo manuales se encuentran palabras traducidas, comentarios morfosintácticos, etc.
- Ausencia de objetivos pedagógicos y, especialmente, que el objetivo no es enseñar a traducir, sino el perfeccionamiento lingüístico, para el que se utilizan sobre todo textos literarios.
- Ausencia de criterios metodológicos, que se manifiesta en tres aspectos:
 1. Falta de criterios en la selección de textos y en la progresión en la enseñanza.
 2. Falta de consignas metodológicas ya que no se proporciona al alumno un método de trabajo para que no cometa otra vez un error del mismo tipo.
 3. Falta de diferenciación entre traducción directa e inversa en cuanto a objetivos de aprendizaje y planteamientos metodológicos.

Rasgos característicos de la didáctica tradicional de la traducción (Hurtado 1999: 16 y ss).

Finalmente, en la didáctica tradicional de la traducción también se da una polarización de resultados, es decir, se suele ofrecer una única solución a un problema con lo que se corre el peligro de que el alumno la fije para otras ocasiones.

En definitiva, esta metodología va de la mano con las ideas de que el profesor es el único transmisor de información, así como de que a traducir se aprende sólo traduciendo, unas ideas que, hasta cierto punto, aún hoy persisten.

5.3.2. Los estudios contrastivos

Como afirma Hurtado (1999: 18) “de todos los estudios comparados entre lenguas, los que con mayor fuerza se han planteado explícitamente como un método de enseñanza de la traducción son las Estilísticas comparadas” y, se podría decir que, dentro de éstas, lo más representativo son los procedimientos de traducción de Vinay y Darbelnet (1958), que son el préstamo, el calco, la traducción literal, la transposición, la modulación, la equivalencia y la adaptación.

Las estilísticas comparadas presentan insuficiencias teóricas a la hora de explicar la traducción y configurar su enseñanza ya que, por un lado, se centran en el resultado y no explican el proceso (o proceder) para llegar al mismo; por otro lado, también presentan comparaciones descontextualizadas ya que, generalmente, se comparan unidades aisladas; y por último, se produce una fijación de la comparación efectuada al

proponer una única equivalencia. Estas insuficiencias, afirma Hurtado (ídem: 19), llevan a otras pedagógicas ya que los objetivos de aprendizaje de la traducción quedan limitados a cuestiones de diferencias entre las dos lenguas y la metodología se limita a ejercicios en torno al uso o detección de esas diferencias.

Un avance hacia una textología comparada lo representan los enfoques contrastivos textuales que dejan de centrarse sólo en lo lingüístico para fijarse en lo textual, si bien es cierto que todavía presentan insuficiencias teóricas y metodológicas, ya que siguen polarizándose los resultados y no se explica cómo se ha de recorrer el proceso traductor para llegar a una buena solución de los diferentes problemas que puede plantear el texto original.

Sólo con una parte contrastiva, encontramos el libro de López Guix y Minnet Wilkinson (1997), que es un manual de traducción del inglés al español en el que los autores dedican una parte a la resolución de problemas específicos de este par de lenguas de tipo sintáctico, morfológico y de puntuación, y otra parte a aspectos más teórico-descriptivos donde se detienen a ejemplificar los principales procedimientos de traducción que se han listado hasta el momento (préstamo, calco, traducción literal, transposición, modulación, equivalencia, adaptación, expansión, reducción y compensación).

En la serie de manuales de traducción que otorgan gran importancia a la contrastividad entre distintos pares de lenguas de Hervey y Higgins, que comienza en el 1992, los autores también se unen a la idea de que es posible enseñar a traducir y fijan sus objetivos generales en desarrollar las habilidades de traducción de los alumnos y mejorar la calidad de su trabajo en general. Sin embargo, los contenidos y la forma de presentar las actividades que proponen ciertamente carecen, a nuestro parecer, de una progresión más razonada y de una explicación más detallada, respectivamente.

5.3.3. Los tratados esencialmente teóricos

La confusión entre teoría de la traducción y didáctica de la traducción o entre el saber teórico sobre la traducción y el saber traducir viene representada por obras a las que les falta orientación pedagógica, obras que incluyen una introducción sobre metodología y una parte de textos para traducir u obras que incluyen una parte teórica sobre la traducción con alguna que otra aplicación práctica (Tatilon 1986; Newmark 1988). Esta forma de aproximarse a la enseñanza de la traducción presenta dos problemas

fundamentales que son, por un lado, la gran importancia que se da al conocimiento teórico para aprender a traducir y, por otro, el mecanicismo que se ve en la relación entre teoría y práctica.

Hurtado resume que, a pesar del avance experimentado en la didáctica de la traducción, existen en estos enfoques muchas insuficiencias pedagógicas:

- Falta de definición de objetivos de aprendizaje: se plantean cuestiones de tipo lingüístico, teórico y de método de trabajo, pero no objetivos relacionados con las dificultades que supone aprender a traducir.
- Polarización en los resultados: no hay indicaciones suficientes de cómo ha de ir armándose al estudiante para que consiga adquirir un método de trabajo y capte los principios que le permitan aprender a resolver los problemas de traducción.
- Falta de un marco pedagógico propio: criterios de selección de textos, actividades que le enseñen a traducirlos correctamente, progresión, evaluación...

Insuficiencias pedagógicas en la didáctica de la traducción (Hurtado 1999: 20).

5.3.4. La enseñanza por objetivos de aprendizaje

El primero en aplicar a la traducción la enseñanza por objetivos de aprendizaje fue Delisle en 1980, para quien un objetivo de aprendizaje es una descripción de la intención que persigue una actividad pedagógica y que precisa los cambios a largo plazo que deberá efectuar el estudiante (Delisle 1993). Así, con su libro de 1980, Delisle empieza a poner un mayor orden en la enseñanza de la traducción en cuanto a metodología y objetivos se refiere. Como afirma Hurtado:

“El gran paso en la Didáctica de la traducción lo da Delisle en *L'analyse du discours comme méthode de traduction* (1980), al quejarse de la falta de sistematización en la enseñanza de la traducción y buscar estrategias pedagógicas, planteando la necesidad de una pedagogía heurística que haga descubrir al estudiante los principios que ha de seguir para efectuar un correcto desarrollo del proceso traductor” (1999: 20).

La obra de Delisle, que propone un método de iniciación a la traducción “profesional” y no ya sólo pedagógica, deja bien sentadas en un solo párrafo las bases de lo que debe contener un método de enseñanza. Estos elementos básicos que compondrán las unidades didácticas y les darán un sentido de integración se recogen, como se podrá comprobar más adelante, en la parte aplicada del presente trabajo.

“Une méthode d’enseignement doit clairement délimiter la matière à transmettre, sérier les difficultés, fixer des objectifs d’apprentissage, préciser les moyens permettant de les atteindre, établir une progression dans la formation et, en fin, prévoir des modalités d’évaluation des performances observables” (Delisle 1993: 15).

Delisle plantea que el objeto propio del aprendizaje de la traducción es el desarrollo de dos competencias generales, que son la comprensión y la reexpresión, y tres aptitudes, que son disociar las lenguas, aplicar los procedimientos de traducción y dominar las técnicas de redacción. Entre los términos clave de su propuesta están, por supuesto, los objetivos de aprendizaje y la jerarquía que se establece dentro de los mismos. A continuación ofrecemos las definiciones de estos términos:

- **Objetivo general:** Breve enunciado de intenciones, formulado en términos más o menos precisos, que indica los resultados a que debe conducir un proceso de aprendizaje en el marco de un programa de estudios o de un curso (Delisle 1993).
- **Objetivo específico:** Enunciado, formulado en términos de comportamientos observables, que describe, del modo más preciso posible, los resultados a que deben conducir una o varias actividades pedagógicas en el marco de un programa de estudios o de un curso (Delisle 1993).
- **Objetivo intermedio:** “Enunciado que describe las nociones y los comportamientos que favorecen la consecución de un objetivo específico” (Hurtado 1999: 248).

Entre los objetivos generales que propone Delisle para su método de iniciación, y que luego se desglosan en 56 objetivos específicos, queremos destacar el objetivo referente a la documentación básica del traductor por su relevancia para este trabajo; relacionado con este objetivo general el autor comenta los tipos de obras básicas con las que debería contar el traductor. En este sentido, y si bien no menciona específicamente el uso de corpus electrónicos como tal, sí habla de recoger textos para documentarse y opina que las obras contextualizadas o las enciclopedias son más útiles que las lexicográficas, como se ve en las limitaciones que presentan los diccionarios bilingües.

En resumen, y parafraseando de nuevo a Hurtado (1999: 23), el diseño de objetivos de aprendizaje persigue establecer de un modo sistemático los contenidos y la progresión de la enseñanza. Con esta intención, la misma autora (1996b, 1999) propone cuatro categorías de objetivos de aprendizaje inherentes a la traducción y que se han de ver plasmadas en su enseñanza. Cada una de dichas categorías está relacionada con distintas subcompetencias de la competencia traductora, si bien todas están

relacionadas entre sí y juegan un papel en la adquisición de la misma. Estos cuatro tipos de objetivos son (Hurtado 1996b):

- **Objetivos metodológicos:** están relacionados con los principios y estrategias que se tienen que controlar para recorrer el proceso de traducción y llegar a soluciones apropiadas.
- **Objetivos contrastivos:** están relacionados con la búsqueda de soluciones a divergencias fundamentales entre las dos lenguas implicadas en el proceso.
- **Objetivos profesionales:** están relacionados con el conocimiento del mercado de trabajo, el uso de recursos de documentación y herramientas útiles para el traductor (en papel y en formato digital).
- **Objetivos textuales:** están relacionados con la resolución de problemas de traducción variados que dependen del texto.

Para terminar, sólo cabe reconocer la importancia del diseño de los objetivos de aprendizaje debido a que son “el punto de partida de cualquier didáctica [...], ya que constituyen el esqueleto en torno al cual se confecciona la enseñanza” (Hurtado 2001: 166) y, aún más, son también el punto de llegada que se espera que el estudiante alcance. Los objetivos de aprendizaje son, en definitiva, la salida y la meta de un camino que es el propio proceso de aprendizaje.

5.3.5. Los enfoques profesionales

La integración de la perspectiva del mundo profesional en la formación de traductores, sin llegar a ser una corriente didáctica en el sentido estricto del término, entró con fuerza de la mano de Nord (1991) y con base en el funcionalismo. La premisa de la que se parte es la de que la formación debe emular la práctica profesional de la traducción y que las tareas propuestas por el docente deben ser no sólo realistas sino factibles y no desmoralizadoras para el estudiante a la hora de ir adquiriendo gradualmente la competencia traductora.

En la misma línea profesionalizadora, Vienne (1994) propone el enfoque situacional, basado en la idea de que las actividades formativas deberían ser tareas de traducción llevadas a cabo por los docentes en el ámbito profesional como traductores para poder dar a dichas actividades un tono todavía más realista. Gouadec (1994), por su parte, propone introducir en la formación encargos de traducción reales, no sólo realistas, con clientes y circunstancias reales.

5.3.6. El constructivismo social

De Kiraly (2000), lo más destacado es su defensa del constructivismo social aplicado al aprendizaje y que básicamente postula que éste último no es sólo un proceso individual, sino social y que se construye en la interacción con los demás y el entorno. Así, Kiraly ha aplicado este constructivismo social y el trabajo colaborativo a la traducción centrándose en la importancia que tiene el hecho de que el traductor desarrolle un concepto de sí mismo y se integre en la comunidad profesional de traductores a través de la práctica real de la traducción y la interacción con los profesionales a cuyo grupo pretende pertenecer.

5.3.7. El enfoque por tareas de traducción

5.3.7.1. El enfoque por tareas en la enseñanza de lenguas

El enfoque por tareas en la enseñanza de lenguas hunde sus raíces en el enfoque comunicativo y subsana algunas deficiencias de éste último. Mucho más integrador, el enfoque por tareas “va a ser el motor de una metodología comunicativa específica, que preconiza una convergencia entre los objetivos de aprendizaje y los medios de aprendizaje, sin efectuar distinción entre el diseño de los contenidos educativos y la metodología [...]. El Enfoque por tareas es, pues, un marco metodológico, pero también de diseño curricular.” (Hurtado 1999: 47). En otras palabras, este enfoque no sólo cubre el diseño de tareas, sino también la identificación y análisis del contexto, la selección de objetivos y contenidos y la evaluación, integrándolo todo y dotando al estudiante de un papel más activo.

Dentro de este enfoque se considera que las tareas (y no los contenidos) son las unidades organizadoras del proceso de aprendizaje. La tarea es, basándonos en Zanón (1990), “una unidad de actividades en el aula, representativa de procesos de comunicación, que se dirige intencionalmente al aprendizaje y que está diseñada con un objetivo concreto, una estructura y una secuencia de trabajo” (Hurtado 1999: 49). Uno de los teóricos fundamentales del enfoque por tareas, Nunan (1989), distingue entre tareas de ensayo y tareas pedagógicas, siendo las primeras las que proporciona el mundo real y las segundas las que vienen inspiradas por la investigación en aprendizaje. Nunan también ofrece indicaciones respecto a los factores que se deben tener en cuenta a la hora de realizar la gradación de las tareas, y que son los contenidos,

el alumno y la actividad en sí. Teniendo esto presente, una agrupación coherente de tareas dará lugar a una unidad didáctica.

“La secuenciación e integración de tareas para formar una unidad didáctica puede efectuarse siguiendo varios principios: por simple encadenamiento de tareas, en torno a la resolución de un problema, por el contenido, por escenarios diferentes... De todos modos, lo importante es que exista una buena articulación entre las tareas, y que éstas estén bien delimitadas y secuenciadas” (Hurtado 1999: 49).

Para el diseño de la unidad didáctica completa Estaire y Zanón proponen los siguientes apartados.

1. Elección del tema o área de interés.
2. Especificación de objetivos.
3. Programación de tarea/s final/es que demostrarán la consecución de los objetivos.
4. Especificación de los contenidos [...] necesarios para la realización de las tareas.
5. Programación de tareas posibilitadoras y tareas de comunicación necesarias para la consecución de las tareas finales [...].
6. Especificación de la evaluación de cada tarea, que se incorpora así al proceso de aprendizaje.

Marco para el diseño de la unidad didáctica (Estaire y Zanón 1994, en Hurtado 1999: 49).

5.3.7.2. *El enfoque por tareas de traducción*

Hurtado inicia la aplicación del enfoque por tareas a la traducción ya en 1992. Partiendo de la caracterización de “tarea” de Zanón (1990) antes citada, Hurtado define “tarea de traducción” como “una unidad de trabajo en el aula, representativa de la práctica traductora, que se dirige intencionalmente al aprendizaje de la traducción y que está diseñada con un objetivo concreto, estructura y secuencia de trabajo” (1999: 56). Las tareas de traducción, pensadas para lograr los objetivos propuestos y secuenciadas de forma coherente, darán lugar a las unidades didácticas. El enfoque en general y las unidades didácticas en particular se caracterizarán por su gran flexibilidad de aplicación; se podrán adaptar y modificar a medida que se ponen en práctica y vayan surgiendo nuevas necesidades, como sería el caso si se observara que los alumnos necesitan la introducción de pretareas o postareas con el fin de preparar el terreno o de consolidar objetivos, respectivamente.

Los elementos fundamentales que deben indicarse en una unidad didáctica, según la concepción de Hurtado (1999), son el título, el nivel de la misma, las materias en las que se puede poner en práctica, los objetivos y la secuenciación de las tareas de la unidad. De forma esquemática, la estructura general de una unidad didáctica es la siguiente:

TÍTULO DE LA UNIDAD: NIVEL: MATERIA(S): OBJETIVO(S) ESPECÍFICO(S):
ESTRUCTURACIÓN DE LA UNIDAD TAREA 1: TAREA 2: TAREA ... TAREA FINAL

Tabla 4. Estructuración de la unidad didáctica (Hurtado 1999).

Cada tarea tiene uno o varios objetivos que se persiguen y que pueden ser de distinto tipo. También se plantean los materiales que se utilizan (textos para traducir, textos paralelos, fichas de ejercicios, recursos electrónicos, etc.), el desarrollo de la tarea (actividades, dinámica, duración estimada, etc.), la evaluación (tipo, medios), si cabe, ya que no es necesario que todas las tareas se evalúen, y comentarios sobre el desarrollo de la tarea, antes o después de su puesta en práctica (anotaciones, modificaciones, conveniencia de la realización de pre- o postareas).

ESTRUCTURACIÓN DE CADA TAREA
OBJETIVO(S): MATERIALES: DESARROLLO DE LA TAREA: (EVALUACIÓN): COMENTARIOS:

Tabla 5. Estructuración de la tarea (Hurtado 1999).

Finalmente, otra autora que posteriormente ha aplicado el enfoque por tareas a la didáctica de la traducción es González Davies (2003, 2004), que también habla de la secuenciación de las tareas y sus distintos tipos.

“[...] nuestras propuestas se articulan alrededor de unas *tareas* cuyas *cadena de actividades* relacionadas y coherentes inciden en el *proceso* de adquisición y en la mejora de la competencia traductora y, por tanto, en el *producto* de los alumnos. Estas actividades pueden ser tanto *pedagógicas* como *de mundo real* [...] el primer tipo de actividades sirve de apoyo, introducción, refuerzo o concienciación del segundo tipo: las actividades que imitan encargos de traducción del mundo real” (González Davies 2003: 13).

Aparte del enfoque por tareas, esta autora destaca por la importancia que da a la interacción en la formación, basándose en el anteriormente citado constructivismo social, así como por perseguir la progresiva autonomía del alumno.

5.3.8. La formación por competencias en traducción

5.3.8.1. La formación por competencias

La formación por competencias (FPC) es una corriente pedagógica que ha surgido recientemente, si bien es cierto que nace de la evolución de modelos anteriores.

Competence-based training (CBT) is a logical continuation of objectives-based learning , the latter developed since the 1960s. I believe that the fundamental advances represented by CBT compared to this forerunner are: greater transparency of professional profile in study programmes, greater emphasis on the outcome of learning, more flexibility and a greater integration of all aspects of a curriculum (Hurtado 2007: 165).

Una competencia, según Lasnier, se define como:

“Un savoir-agir complexe résultant de l’intégration, de la mobilisation et de l’agencement d’un ensemble de capacités et d’habilités (pouvant être d’ordre cognitif, affectif, psychomoteur ou social) et de connaissances (connaissances déclaratives) utilisées efficacement, dans des situations ayant un caractère commun” (2000: 32).

La integración, como se acaba de comprobar, destaca como palabra clave en este modelo de formación. Se habla de integración de los componentes de la competencia, integración de las diferentes competencias que configuran una titulación e integración de las actividades de aprendizaje y evaluación diseñadas. Se integran ahora con más fuerza en el diseño curricular y de actividades los conocimientos previos del alumnado, y se diseñan los objetivos en términos de resultados de aprendizaje. Dichos objetivos o resultados de aprendizaje están relacionados con competencias y expresan lo que el

estudiante será capaz de hacer cuando acabe el proceso de aprendizaje (por ejemplo, al acabar la unidad didáctica X / el módulo Y / el curso Z, el estudiante será capaz de ... identificar.../ analizar.../ aplicar.../ desarrollar.../ interpretar.../ obtener.../ formular.../ evaluar.../ interactuar.../ redactar.../ etc.).

El mismo autor nos da y define brevemente los principios de la FPC y que deberían tenerse en cuenta con vistas a la adquisición de una competencia:

- Globalité: éléments analysés à partir d'une situation globale
- Construction: activation des acquis antérieurs, élaboration de nouveaux apprentissages, organisation des informations
- Alternance: global → spécifique → global
compétence → composantes → compétence
tâche intégratrice → activité d'apprentissage spécifique → tâche intégratrice
- Application: apprentissage par l'agir
- Distinction: entre contenu et processus
- Signifiante: situations signifiantes, motivantes pour l'élève
- Cohérence: la relation cohérente entre les activités d'enseignement, les activités d'apprentissage, les activités d'évaluation et la compétence
- Intégration: les éléments étudiés sont liés entre eux et liés à la compétence; l'apprenant développe la compétence en utilisant les composantes de façon intégrée
- Itération: l'apprenant est soumis plusieurs fois à un même type de tâches intégratrices en lien avec la compétence et à un même contenu disciplinaire
- Transfert: transfert d'une tâche source à une tâche cible. Utilisation, dans un autre contexte, de connaissances et de capacités apprises dans un contexte donné

Principios de la FPC (Lasnier 2000: 158 y ss).

La FPC, en contraste con la etapa anterior de formación por objetivos, destaca en que aporta, según Hurtado:

- plus de transparence des profils professionnels dans les programmes d'études ;
- plus d'incidence sur les résultats d'apprentissage ;
- plus de flexibilité de la planification curriculaire ;
- plus d'intégration de tous les aspects curriculaires.

Aspectos destacados de la FPC respecto a la etapa anterior (Hurtado 2008).

Dentro de la formación por competencias se distingue entre competencias genéricas (o transversales) y competencias específicas (o disciplinares):

“Se entiende por *competencias genéricas* aquellas que constituyen una parte fundamental del perfil profesional y del perfil formativo de todas o de la mayoría de las titulaciones. Están relacionadas con cualidades que se asocian a la formación universitaria e incluyen un conjunto de habilidades cognitivas y metacognitivas, conocimientos instrumentales y actitudes consideradas valiosas en la sociedad del conocimiento” (Yániz y Villardón 2006: 24).

“Con el término *competencias específicas* se describen las competencias propias de un perfil exclusivamente o compartidas por un escaso número de ellos. Generalmente, se expresan a través de conocimientos relacionados con las disciplinas o habilidades específicas de las prácticas profesionales más comunes en el perfil definido” (idem: 25).

Respecto a la identificación de competencias genéricas deseables en los egresados tras su formación universitaria, el proyecto Sócrates-Erasmus *Tuning Educational Structures in Europe*, más conocido como “Proyecto Tuning”, ha aportado, entre otras cosas, un listado de dichas competencias clasificadas en tres categorías, instrumentales, interpersonales y sistémicas (González y Wagenaar 2003: 81-84).

Las competencias instrumentales son las que “tienen una función instrumental” (idem: 81).

- Capacidad de análisis y síntesis.
- Capacidad de organizar y planificar.
- Conocimientos generales básicos.
- Conocimientos básicos de la profesión.
- Comunicación oral y escrita en la propia lengua.
- Conocimiento de una segunda lengua.
- Habilidades básicas de manejo del ordenador.
- Habilidades de gestión de la información (habilidad para buscar y analizar información proveniente de fuentes diversas).
- Resolución de problemas.
- Toma de decisiones.

Competencias instrumentales (idem: 83).

Las competencias interpersonales son “las capacidades *individuales* relativas a la capacidad de expresar los propios sentimientos, habilidades críticas y de autocrítica” (idem: 82).

- Capacidad crítica y autocrítica.
- Trabajo en equipo.
- Habilidades interpersonales.
- Capacidad de trabajar en un equipo interdisciplinar.
- Capacidad para comunicarse con expertos de otras áreas.
- Apreciación de la diversidad y multiculturalidad.
- Habilidad de trabajar en un contexto internacional.
- Compromiso ético.

Competencias interpersonales (ídem: 84).

Las competencias sistémicas “son las destrezas y habilidades que conciernen a los *sistemas como totalidad*” (ídem: 82).

- Capacidad de aplicar los conocimientos en la práctica.
- Habilidades de investigación.
- Capacidad de aprender.
- Capacidad para adaptarse a nuevas situaciones.
- Capacidad para generar nuevas ideas (creatividad).
- Liderazgo.
- Conocimiento de culturas y costumbres de otros países.
- Habilidad para trabajar de forma autónoma.
- Diseño y gestión de proyectos.
- Iniciativa y espíritu emprendedor.
- Preocupación por la calidad.
- Motivación de logro.

Competencias sistémicas (ídem: 84).

El movimiento hacia la formación por competencias ha venido impulsado por cambios en la sociedad, la economía y la empresa y la necesidad de que la universidad tenga una comunicación más cercana con dicha sociedad. Debido a la globalización de los mercados existe una demanda de trabajadores más integrales y flexibles, trabajadores que sean capaces de aprender de forma autónoma y puedan adaptarse a distintos, no ya puestos de trabajo, sino funciones o roles, es decir, trabajadores con competencias diversas, generales y específicas, que puedan aplicar de acuerdo con las circunstancias que les rodeen, y todo esto dando por supuesto su buen manejo de las nuevas tecnologías.

Con el objetivo de aumentar, entre otras cosas, la empleabilidad y la movilidad de los trabajadores potenciales (los estudiantes) en la Unión Europea, cabe mencionar que

algunos países pertenecientes a ésta y otros, se lanzaron a crear el Espacio Europeo de Enseñanza Superior (EEES), que acoge la FPC como marco pedagógico y de planificación curricular.

Para dicho diseño curricular, se plantean los siguientes pasos generales con el fin de desarrollar en los estudiantes capacidades para integrarse en la sociedad como ciudadanos y trabajadores:

1. Identificar y desarrollar el perfil profesional.
2. Identificar los resultados de aprendizaje (objetivos).
3. Identificar cómo se facilita ese aprendizaje y el desarrollo de las competencias.

5.3.8.2. La formación por competencias en traducción

5.3.8.2.1. La importancia del perfil profesional

Según Fernández *et al.* (2002, en Yániz y Villardón 2006: 18),

“un perfil profesional define la identidad profesional de las personas que llevan a cabo una determinada labor y explica las funciones principales que dicha profesión cumple, así como las tareas más habituales en las que se plasman esas funciones” (Fernández *et al.* 2002, en Yániz y Villardón 2006: 18).

Al perfil profesional se llega planteándose qué tiene que *saber* el trabajador (conocimientos teóricos), qué tiene que *saber hacer* (conocimientos prácticos) y cómo tiene que *saber estar y actuar*. A su vez, el perfil profesional, una vez definido, ayudará a, como dice Hurtado, “identifier la formation nécessaire, c’est-à-dire les compétences nécessaires, qui facilitent à leur tour l’identification des éléments formatifs” (2008).

Desde el ámbito académico, Kelly es una de las pocas autoras que ha reflexionado sobre este tema para llegar a identificar un perfil básico al que se le unirían habilidades específicas (Kelly 2005: 23-28). Kuznik (2007), desde una perspectiva laboral, estudia la identificación de las tareas del traductor e intérprete en este ámbito.

Otro esfuerzo en cuanto a la identificación del perfil profesional del traductor e intérprete es el Libro Blanco de Traducción e Interpretación elaborado por la Conferencia de Centros y Departamentos Universitarios de Traducción e Interpretación

(CCDUTI) bajo el auspicio de la ANECA (Agencia Nacional de Evaluación de la Calidad y Acreditación). Seis son los perfiles profesionales identificados en este estudio:

1. Traductor profesional “generalista”.
2. Mediador lingüístico y cultural.
3. Intérprete de enlace.
4. Lector editorial, redactor, corrector, revisor.
5. Lexicógrafo, terminólogo y gestor de proyectos lingüísticos.
6. Docente de lenguas.

Para formar profesionales con alguno o varios de estos perfiles se tienen que fijar objetivos acordes con las competencias que se desea que el estudiante desarrolle. En el Libro Blanco, son los datos resultantes de las encuestas realizadas a tres grupos de sujetos, egresados, empleadores y docentes, de 18 centros donde se imparten estudios superiores de traducción e interpretación de Europa, los que señalan las tendencias en cuanto a la importancia de las competencias.

5.3.8.2.2. La formación por competencias en traducción

Si recordamos, el Libro Blanco de Traducción e Interpretación antes citado identifica 3 categorías de competencias: transversales, interpersonales y sistémicas. En el estudio realizado, respecto a las competencias transversales, los encuestados primaron las siguientes⁴ en sus respectivas categorías:

- Instrumentales
 1. Comunicación oral y escrita en la lengua propia
 2. Conocimiento de una 2ª lengua extranjera
- Personales
 1. Compromiso ético
 2. Razonamiento crítico
- Sistémicas
 1. Motivación por la calidad
 2. Aprendizaje autónomo
- Otras
 1. Capacidad de aplicar los conocimientos a la práctica
 2. Capacidad de trabajo individual

Competencias genéricas más valoradas (ANECA 2004: 81-83).

⁴ Para no extendernos con las listas de competencias identificadas, que se pueden consultar en el propio *Libro Blanco de Traducción e Interpretación* < http://www.aneca.es/activin/docs/libroblanco_traduc_def.pdf >, aquí sólo se recogen las dos competencias más valoradas en cada categoría.

En cuanto a las competencias específicas, éstas se clasifican en disciplinares (saber), profesionales (saber hacer) y académicas. A continuación siguen las competencias específicas más valoradas por los encuestados, si bien en el Libro Blanco se advierte de que estos datos son “fiables aunque menos informativos” (ANECA 2004: 17) debido a que se pudieron apreciar tendencias a pesar de que las respuestas fueron un tanto irregulares.

1. Dominio de lenguas extranjeras.
2. Conocimiento de culturas y civilizaciones extranjeras.
3. Dominio de la lengua propia, escrita y oral.
4. Dominio de técnicas y terminología de la traducción especializada.
5. Manejo de herramientas informáticas.
6. Dominio de técnicas de traducción asistida/localización.
7. Destreza para la búsqueda de información/documentación.
8. Conocimiento de los aspectos económicos y profesionales.
9. Capacidad de trabajo en equipo.
10. Capacidad de diseñar y gestionar proyectos.
11. Poseer una amplia cultura.

Competencias específicas más señaladas (las que definirían a un buen traductor), según el *Libro Blanco de Traducción e Interpretación* (ANECA 2004: 87-88).

Sin embargo, a pesar de los datos obtenidos de los tres grupos de sujetos consultados (profesores, licenciados y empleadores), el estudio carece de un mayor detalle en cuanto a, por ejemplo, la definición del universo de encuestados; en el caso de los empleadores el informe no concreta sus características con lo que los resultados podrían variar dependiendo de si éstos se refieren a empresas, agencias de traducción, particulares, etc. Por otro lado, si bien se aprecia el esfuerzo, el potencial del estudio y las tendencias que apunta, no se puede afirmar que la muestra sea muy amplia ya que, por ejemplo, sólo se obtuvieron datos de 50 empleadores. Finalmente, las competencias identificadas, sobre todo las específicas, están ciertamente poco definidas, como por ejemplo, la competencia en “buscar información y documentación”, que resulta muy poco valorada por los empleadores frente a otras. En este sentido, la interpretación de los datos es bastante superficial y no ofrece argumentos.

Hurtado (2007, 2008) es pionera en realizar una propuesta de aplicación de la FPC a la traducción, identificando varias categorías de competencias específicas en la formación de traductores y operacionalizando dichas competencias. Para esto, la autora parte de su investigación precedente en la formación por objetivos de aprendizaje y tareas de

traducción (Hurtado 1983, 1984, etc.). Así, las categorías de competencias que ha identificado están relacionadas con las ya mencionadas (cfr. supra 5.3.4.) categorías de objetivos de aprendizaje propios de la traducción con las que trabaja la autora, que son los objetivos metodológicos, contrastivos, profesionales e instrumentales, y textuales (Hurtado 1996b, 1999a), cada uno de los cuales se relaciona con las diferentes subcompetencias de la competencia traductora que define el grupo PACTE (cfr. supra 5.1.1.). En sus últimas publicaciones Hurtado (2007, 2008) plantea una redefinición de las cuatro categorías de objetivos de aprendizaje para pasar a enunciarlos en términos de competencias y acabar proponiendo seis categorías de competencias específicas: metodológicas y estratégicas, contrastivas, extralingüísticas, laborales, instrumentales y textuales. A continuación, mostramos estas seis categorías de competencias relevantes para la traducción que propone y define Hurtado.

METHODOLOGICAL AND STRATEGIC COMPETENCES

- Competences related to the principles and strategies that have to be applied in order to work through the translation process and to be able to reach the appropriate translation solution for each case. They serve as a way to grasp principles and develop skills related to translation practice.
- They develop strategic subcompetence, translation knowledge and certain TC psychophysiological components.

DEFINITION: Apply the methodological principles and strategies necessary to work through the translation process appropriately.

CONTRASTIVE COMPETENCES

- Competences related to knowing the differences between the two languages and monitoring interferences.
- They develop the TC bilingual sub-competence.
- They play an important role in *Introduction to translation B-A* (general translation).

DEFINITION: Differentiate between the two languages, monitoring interferences when it comes to solving translation problems.

EXTRA-LINGUISTIC COMPETENCES

- Competences related to mobilising encyclopaedic, bicultural and thematic knowledge.
- They develop the TC extra-linguistic subcompetence.
- They play an important role in specialised translation (thematic knowledge from specific fields).

DEFINITION: Mobilise encyclopaedic, bicultural and thematic knowledge to solve translation problems.

OCCUPATIONAL COMPETENCES

- Competences related to knowledge of the job market.
- They develop the TC translation knowledge subcompetence (regarding professional aspects).

DEFINITION: Operate in the translation job market appropriately.

INSTRUMENTAL COMPETENCES
<ul style="list-style-type: none"> • Competences related to using documentary sources and a whole array of tools that are useful for the translator (in digital or paper format, etc.). • They develop the TC instrumental sub-competence. <p>DEFINITION: Handle documentary sources and an array of tools to solve translation problems.</p>
TEXTUAL COMPETENCES
<ul style="list-style-type: none"> • Competences related to solving a range of translation problems that arise from different texts. • They imply an integrated development of all the TC sub-competences and of psycho-physiological components. <p>DEFINITION: Solve translation problems in different genres using the appropriate strategies.</p>

Tabla 6. Categorías de competencias específicas en la formación de traductores (Hurtado 2007: 178).

Sin embargo, como la propia Hurtado afirma (2008), las competencias no toman forma hasta que se operacionalizan, es decir, hasta que se definen de tal manera que permitan realizar una planificación didáctica. Para operacionalizar una competencia, según la autora, deben estar presentes, al menos, los siguientes aspectos:

- | |
|---|
| <ul style="list-style-type: none"> - El enunciado de la competencia, que da una idea global de la misma. - Los componentes de la competencia, es decir, los elementos que la forman y que sirven de base para elaborar los indicadores. - El contenido disciplinario asociado a la competencia. - Los indicadores de la competencia, es decir, los elementos observables que sirven de base para la evaluación. |
|---|

Aspectos esenciales a la hora de operacionalizar una competencia (Hurtado 2008).

En el siguiente capítulo se presentan detalladamente los elementos en los que se enmarca la propuesta, en lo que a metodología se refiere. Se exponen las competencias y objetivos que se persiguen, los contenidos y la evaluación, además de la lógica que hay detrás de la elección de todo ello.

6. MARCO METODOLÓGICO Y DE DISEÑO CURRICULAR PARA EL USO DE CORPUS EN LA FORMACIÓN DE TRADUCTORES

Expuestas la evolución de las distintas tendencias que han marcado la didáctica de la traducción y las razones por las que hemos elegido finalmente el enfoque por tareas de traducción integrado con la formación por competencias, base del diseño de objetivos de aprendizaje, dedicaremos este capítulo a la explicación del marco de la propuesta que efectuamos para la utilización de corpus electrónicos en la formación de traductores. Se plantea aquí un marco metodológico y de diseño curricular que incluye el tratamiento en detalle de una competencia específica, el uso de corpus para traducir, así como el diseño de objetivos de aprendizaje y el contenido disciplinario asociado; se realiza una propuesta de evaluación de dicho uso y se exponen los criterios que se han seguido para la elaboración de cada unidad didáctica y una evolución de su elaboración.

6.1. Competencia traductora y uso de corpus para traducir: ubicación y componentes

Creemos necesario explicar la estructura y relación bajo la que hemos concebido nuestra propuesta didáctica. Mostraremos aquí la Tabla 7, en la que se recoge la estructura de competencias y componentes de la competencia específica en que se organiza nuestra propuesta: utilizar corpus para resolver problemas de traducción.

La competencia traductora ha de verse como un todo o una (macro)competencia que comprende a otras como la (sub)competencia instrumental. Ésta a su vez está formada por competencias relacionadas con la documentación y las TIC, entre las que se encuentra la competencia objeto de nuestro estudio. Esta competencia ha de entenderse como una competencia específica que se podría aplicar en todas las materias de traducción.

Los componentes de una competencia son las habilidades, conocimientos y actitudes que la integran. Pensamos que la competencia en el uso de corpus para traducir tiene cuatro componentes:

- (1) identificar los principios básicos relacionados con corpus para resolver problemas de traducción;
- (2) crear corpus para traducir;
- (3) manejar software relacionado con corpus para traducir;
- (4) resolver problemas de traducción utilizando corpus.

De los componentes de una competencia se extraen los indicadores de la misma, que permiten operacionalizar la competencia, sirven de premisas para la evaluación y se redactan en 3ª persona. Los cuatro componentes antes citados sirven de base para el diseño de los cuatro objetivos generales distribuidos en dos fases pedagógicas, iniciación y consolidación en el uso de corpus para traducir. Ésta es nuestra propuesta de objetivos, pero otros docentes podrían hacer cualquier otra. En parte, ése es el fin para el que ofrecemos todo el contenido disciplinario, para que otros docentes lo organicen como deseen o necesiten.

A continuación sigue la estructura completa de lo expuesto:

Competencia traductora ↓	El sistema subyacente de conocimientos declarativos y esencialmente operativos necesarios para traducir.
Competencia instrumental ↓	Manejar documentación y herramientas de todo tipo para resolver problemas de traducción.
Uso de corpus para traducir ↓	Utilizar corpus electrónicos de forma apropiada para resolver problemas de traducción.
Componentes del uso de corpus para traducir	<ol style="list-style-type: none"> 1) Identificar los principios básicos relacionados con corpus. 2) Crear corpus. 3) Manejar software relacionado con corpus. 4) Resolver problemas de traducción utilizando corpus.

Tabla 7. Competencia traductora y uso de corpus electrónicos para traducir.

6.2. Objetivos de aprendizaje

Respecto a los avances que la didáctica de la traducción ha experimentado en los últimos años, Hurtado afirma que “la aportación fundamental son las propuestas de organización de la enseñanza mediante el diseño de objetivos de aprendizaje y la búsqueda de una metodología activa” (1999: 20).

Para el diseño de nuestros objetivos generales partimos de los cuatro componentes de la competencia en el uso de corpus para traducir antes citados (Tabla 7). Dentro de este apartado sobre objetivos de aprendizaje trataremos las fuentes de donde éstos se han extraído o donde nos hemos inspirado para elaborarlos. Los objetivos seleccionados o diseñados están relacionados, principalmente, con el uso de corpus para traducir, pero también con competencias genéricas y con diversas materias de traducción. Dentro de los primeros, distinguimos entre objetivos de la fase de iniciación en el uso de corpus para traducir y objetivos de consolidación.

6.2.1. Fuentes de los objetivos

Detrás del diseño de objetivos para una propuesta didáctica debe haber unas razones que, sobre todo, estarán relacionadas con el contexto que rodea a la propuesta. En este sentido, los objetivos diseñados para la propuesta de unidades didácticas incluida en esta tesis se apoyan en varios de los factores que Kelly (2005: 22) menciona como claves a la hora de definir objetivos.

En primer lugar, el contexto sociocultural e institucional para el que está ideada nuestra propuesta didáctica es el de la enseñanza superior de traducción en España, con un alumnado que se caracteriza principalmente por tener un grado de madurez semi-alto, una experiencia en traducción profesional nula o escasa y, en cualquier caso, heterogénea dentro de cada grupo, una práctica y unos conocimientos de informática a nivel de usuario pero, también, heterogéneos a nivel de grupo, y, eso sí, con una gran necesidad de práctica de las TIC de forma contextualizada; aparte, en cuanto a infraestructuras, se puede contar con que los alumnos suelen tener acceso a ordenadores e Internet, de los que los centros de traducción suelen estar equipados. La observación del alumnado y la experiencia de los últimos años añade también al perfil de alumno el hecho de que éste, por razones laborales o geográficas, pueda seguir los contenidos sólo semipresencialmente, lo que implica que el uso de recursos

electrónicos como los corpus con el software relacionado y unas tareas bien estructuradas sean un punto positivo para aquellos que necesiten hacer la formación parcialmente a distancia. Por otro lado, las necesidades y perspectivas del mundo laboral también se han estudiado a través de dos pequeños sondeos (cfr. supra 2.4. y cfr. infra 6.2.4.1.) y un amplísimo estudio de mercado realizado en 2004 por la Agrupación de Centros Especializados en Traducción (ACT), adquirido por el proyecto TRACE del Grupo Tradumática¹, que se han tenido en consideración a la hora de elaborar nuestra propuesta de objetivos.

Aparte de los factores que menciona Kelly, también hemos tenido en cuenta otros. Así, aunque algunos autores no explicitan los objetivos de aprendizaje del uso de corpus o no hacen referencia a las competencias que el estudiante debe adquirir en este respecto, sí expresan su necesidad en otros términos. Así, Zanettin comenta que “... translators need to be able to see patterns and regularities both within a language and across languages” (Zanettin 1994: 109), y Varantola hace un listado de competencias agrupadas en dos categorías que, de hecho, han servido de base para nuestra propuesta de cuatro componentes de la competencia en el uso de corpus para traducir, que a su vez se relacionan con cuatro objetivos generales.

Corpus compilation

- Corpus design and design criteria
- Search strategies and search word selection
- Source criticism to assess the reliability of corpus texts
- Assessment of corpus adequacy and relevancy
- Software literacy in general
- Selection of Internet search engines
- Integrated use of word processing tools and corpus tools

Use of corpus information

- Deductive corpus analysis skills in general
- Use of preliminary corpus information for more targeted compilation criteria
- Use of corpus evidence for translational decisions
- Corpus evaluation and decision-making skills
- Distinctions between permanent corpus collections and targeted, disposable corpora
- Overall corpus knowledge management skills

Competencias necesarias relacionadas con el uso de corpus (Varantola 2003: 69).

¹ Para más información sobre el grupo y sus actividades, ver <http://tradumatica.net/>.

6.2.2. Objetivos relacionados con el uso de corpus para traducir

Como ya exponíamos en el capítulo anterior, los objetivos de aprendizaje ocupan un lugar primordial en el diseño de cualquier didáctica. Son, se podría decir, el punto de partida y la guía de la enseñanza, a la vez que el punto de llegada y la referencia para la evaluación. Los objetivos están relacionados con las subcompetencias de la competencia traductora y aquí, en concreto, con la subcompetencia instrumental que, como ya se ha definido previamente, es la que se ocupa del uso de las fuentes de documentación y las TIC aplicadas a la traducción. Los objetivos que describiremos recogen los conocimientos y habilidades en diversas áreas que el estudiante de traducción habrá de adquirir para poder efectuar un buen uso de los corpus para traducir. Así, como afirma Aston “to use corpora effectively in increasingly independent research, learners need technical, methodological, and conceptual knowledge and abilities” (2001: 24), que es lo que se ha querido cubrir con el diseño de objetivos incluido en este apartado.

La lingüística de corpus aplicada a la traducción no es una materia en sí de la formación de traductores en el sentido tradicional. Sin embargo, aquí defendemos que sí podría serlo, igual que podría entenderse e integrarse como un contenido transversal de aplicación en materias tanto de iniciación en traducción como de traducción especializada. Y ya que creemos que, hasta cierto punto, deben coincidir en el tiempo la enseñanza de la traducción y la enseñanza del uso de corpus para traducir, en este último caso también hemos establecido dos fases de progresión pedagógica: la iniciación y la consolidación en el uso de corpus para traducir. En la fase de iniciación los alumnos adquieren los principios metodológicos y técnicos básicos del trabajo con corpus, mientras que en la fase de consolidación se profundiza en el aprovechamiento de corpus para traducir. Sin embargo, a pesar de la progresión sugerida en el aprendizaje del uso de corpus para traducir y su relación con las fases de progresión de las materias de traducción, ésta puede verse alterada por factores como los siguientes:

- a) La realidad profesional docente. En muchas ocasiones nos encontramos introduciendo el uso de corpus no en cursos de iniciación a la traducción, sino en cursos de traducción especializada tanto directa como inversa.
- b) La flexibilidad de la propuesta y que algunos objetivos se pueden adelantar o postergar dependiendo de las circunstancias del aprendizaje. Así, por ejemplo, un objetivo relacionado con saber crear un corpus, aunque integrado a priori en la fase de consolidación para poder trabajarlo cuando el alumno ya haya

adquirido conciencia de lo que son los textos paralelos, práctica en la evaluación de fuentes, conocimientos sobre documentación y terminología, un mayor conocimiento de la lengua extranjera, etc., puede simplificarse incluyendo una mayor presencia del profesor como guía, por ejemplo indicando las fuentes de donde recoger los textos.

En cualquier caso, para cada una de las dos fases de progresión indicadas, la de iniciación y la de consolidación, se ha señalado un objetivo global y se han descrito objetivos generales y específicos.

6.2.1.1. Objetivos de la iniciación en el uso de corpus para traducir

El objetivo global de esta fase de iniciación en el uso de corpus para traducir es que el estudiante capte los principios metodológicos básicos del trabajo con corpus para traducir y realice operaciones fundamentales que le permitan enfrentarse después a situaciones de uso más complejas. Este objetivo global se plasma en tres objetivos:

1. Identificar los principios fundamentales del trabajo con corpus para resolver problemas de traducción.
2. Manejar software específico y funciones básicas de los programas para explotar corpus.
3. Resolver problemas de traducción utilizando corpus a nivel básico.

Objetivos de la iniciación al uso de corpus para traducir.

La progresión entre estos objetivos generales se ha de entender de forma flexible ya que, aunque la asimilación de los principios fundamentales de la metodología de corpus venga obligadamente en primer lugar, también es cierto que no es necesario asimilar “todos” los principios para ser capaz de empezar a utilizar corpus para traducir. Por otro lado, el manejo de los programas informáticos utilizados con corpus se puede aprender bajo la excusa de asimilar alguno de aquellos principios, o puede ir de la mano con el uso en sí del corpus para traducir. En cualquier caso, estos tres objetivos generales se desglosan a continuación en objetivos específicos, de lo cual se presenta una tabla sintética.

1. IDENTIFICAR LOS PRINCIPIOS FUNDAMENTALES DEL TRABAJO CON CORPUS PARA RESOLVER PROBLEMAS DE TRADUCCIÓN
1.1. Conocer aspectos básicos sobre corpus
1.2. Identificar los principios metodológicos básicos del trabajo con corpus
2. MANEJAR SOFTWARE ESPECÍFICO Y FUNCIONES BÁSICAS DE LOS PROGRAMAS PARA EXPLOTAR CORPUS PARA TRADUCIR
2.1. Utilizar sistemas de búsqueda en corpus en línea a nivel básico
2.2. Utilizar las funciones básicas de un programa informático de análisis de corpus
2.3. Utilizar las funciones básicas de un programa informático de alineamiento de textos
3. RESOLVER PROBLEMAS DE TRADUCCIÓN UTILIZANDO CORPUS A NIVEL BÁSICO
3.1. Seleccionar el corpus apropiado según las necesidades de traducción
3.2. Integrar conocimientos para resolver problemas de traducción utilizando corpus a nivel básico

Tabla 8. Objetivos generales y específicos de la iniciación en el uso de corpus para traducir.

6.2.1.2. *Objetivos de la consolidación en el uso de corpus para traducir*

El objetivo global de la fase de consolidación en el uso de corpus para traducir es que el estudiante adquiera conocimientos de mayor complejidad y habilidades de uso de corpus más refinadas para profundizar en el aprovechamiento de estos recursos para traducir. Este objetivo global se plasma en tres objetivos:

1. Crear corpus para traducir.
2. Manejar software específico y funciones avanzadas de los programas para explotar corpus para traducir.
3. Resolver problemas de traducción utilizando corpus a nivel avanzado.

Objetivos de la consolidación en el uso de corpus para traducir.

En cuanto a la progresión de objetivos, al igual que ocurre con la fase de iniciación, aquí han de entenderse también de forma flexible y combinada ya que, por ejemplo, el manejo de algunos programas de software es de suma importancia tanto a la hora de crear corpus como de explotarlos para traducir, así como saber cuándo y cómo aumentar el tamaño de un corpus puede redundar en una mejor resolución de

problemas macrotextuales, por poner un ejemplo. A continuación se presentan y detallan los objetivos generales y específicos.

4. CREAR CORPUS
4.1. Conocer fuentes de donde extraer textos
4.2. Evaluar fuentes de donde extraer textos
4.3. Crear corpus de calidad de distintos tipos
5. MANEJAR SOFTWARE ESPECÍFICO Y FUNCIONES AVANZADAS DE LOS PROGRAMAS PARA EXPLOTAR CORPUS
5.1. Utilizar sistemas de búsqueda en corpus en línea a nivel avanzado
5.2. Utilizar software de conversión de textos de papel a formato electrónico
5.3. Utilizar software y recursos de ubicación y descarga de textos electrónicos desde Internet
5.4. Utilizar funciones avanzadas de un programa de análisis de corpus
5.5. Utilizar funciones avanzadas de un programa de alineamiento de corpus
5.6. Utilizar aplicaciones de extracción de terminología
5.7. Calcular la representatividad en un corpus
5.8. Utilizar corpus como memorias de traducción
6. RESOLVER PROBLEMAS DE TRADUCCIÓN UTILIZANDO CORPUS A NIVEL AVANZADO
6.1. Integrar conocimientos para resolver problemas de traducción utilizando corpus a nivel avanzado

Tabla 9. Objetivos generales y específicos de la consolidación en el uso de corpus para traducir.

6.2.3. Objetivos relacionados con competencias genéricas

También se proponen objetivos relacionados con competencias genéricas o transversales extraídas del Proyecto Tuning (González y Wagenaar 2003), y expresadas aquí en infinitivo. En este caso, en nuestras unidades didácticas (cf. infra 6.6.) hay dos competencias transversales, ambas instrumentales, que se trabajan siempre debido a la propia materia. Éstas son:

- Tomar decisiones (en relación con la resolución de problemas de traducción).
- Resolver problemas (de traducción, en sentido amplio).

Aparte de las competencias tratadas por defecto, también se han trabajado otras, si bien en mucho menor grado:

Instrumentales:

- Analizar y sintetizar (p.ej. la información extraída de un corpus).
- Escribir en la propia lengua (cuando se trate de traducción directa o de comunicación con expertos de otras áreas).

Interpersonales:

- Ser crítico y autocrítico (p.ej. con la calidad de la documentación consultada o con el propio conocimiento de la lengua materna o extranjera).
- Trabajar en equipo (como cuando se realiza un trabajo en cadena).

Sistémicas:

- Preocuparse por la calidad (tanto en relación con la documentación consultada como con el resultado final de una traducción).

6.2.4. Objetivos relacionados con materias de traducción**6.2.4.1. Bases profesionales para el diseño de objetivos**

Uno de los factores mencionado por Kelly (2005) a la hora de definir objetivos son las necesidades de la empresa. En lo que respecta a la necesidad de formación del alumnado en traducción inversa en concreto, queremos aportar aquí un pequeño sondeo realizado en relación con el tema de la práctica de la traducción inversa en el mundo profesional. Si bien es cierto que en todas las facultades de traducción de España se imparten asignaturas de traducción inversa “general” y en algunas también especializada, en ocasiones parece necesario justificar la formación en esta materia ante las dudas que existen sobre su práctica real. En este caso, los datos encontrados también han influido en la elaboración de alguna de las unidades didácticas, como es el caso de “Corpus y memoria de traducción”, que se diseñó a partir de la constatación de la traducción técnica como uno de los campos más solicitados. El estudio que se presenta se llevó a cabo en 2004 con 35 participantes, en este caso, 35 traductores autónomos que trabajan con el inglés y el español y no son de origen anglosajón; por otro lado, aunque no se definió la nacionalidad española como condición sine qua non, ésta se deduce en la práctica totalidad de participantes ya que el cuestionario base de este estudio se envió a particulares y a listas de distribución de traductores de Cataluña, en concreto a las listas de ATIC (Associació de Traductors i d’Intèrprets de Catalunya) y TRIAC (Traductors i Intèrprets Associats pro Col·legi). Es también menester subrayar

que, por el tamaño de la muestra, los resultados que aquí se presentan no son generalizables, aunque sí pueden demostrar la tendencia del mercado a pequeña escala.

Las preguntas contenidas en el breve cuestionario distribuido se enumeran a continuación:

- 1) ¿Le han pedido hacer una traducción inversa alguna vez? Sí/No
- 2) ¿Qué tipos de cliente le han pedido traducciones inversas (particular, agencia, empresa...)?
- 3) ¿Qué tipos de textos le han pedido traducir y de qué especialidad (p.ej. artículos científicos, informes financieros, cartas comerciales, páginas web sobre...)?
- 4) ¿Ha hecho traducciones inversas? Sí/No

Documento 3. Cuestionario sobre la práctica profesional de la traducción inversa (español-inglés) en España.

Así pues, según los resultados, al 100% de los traductores que respondieron al cuestionario, es decir, 35, le fue ofrecida en alguna ocasión una traducción inversa del español al inglés. De éstos, sólo 5 rechazaron el encargo o lo derivaron a traductores nativos de la lengua inglesa.

El gráfico que sigue ilustra la cuestión sobre el tipo de clientes que encargan traducciones inversas que, como se puede observar, son mayoritariamente empresas y particulares.

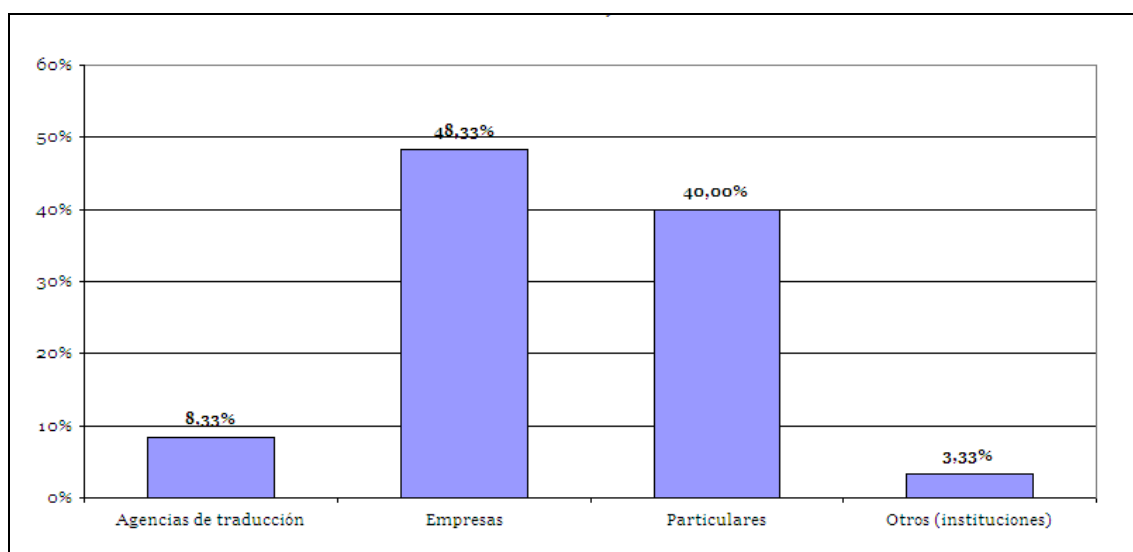


Figura 3. Tipo de clientes que encargan traducciones inversas español-inglés en España.

Por último, recogemos los resultados a la pregunta 3 relacionada con los tipos de texto y campo temático de las traducciones inversas más solicitadas. Aquí hay que decir que, dado que la pregunta era abierta, las respuestas fueron muy variadas. Así, si bien es cierto que los tipos de texto, géneros textuales o soportes mentados van desde el informe (técnico o económico) hasta la página web sobre empresas, pasando por muchos otros que se han explicitado (documentos jurídicos, textos publicitarios), sí hemos podido identificar algunas coincidencias indicativas de la actividad profesional de la traducción inversa. Se ha podido observar, por ejemplo, que los textos más solicitados para la traducción inversa son aquellos de correspondencia comercial, seguidos muy de cerca por textos de especialidad técnica como manuales de usuario, artículos o informes técnicos. Siguiendo en orden descendente, en la traducción científica destacan los artículos sobre tema médico y en la traducción jurídica los contratos y otros tipos de documentos no especificados. Por último, los textos turísticos, publicitarios y periodísticos también aparecen como motivo de encargo de traducciones al inglés, entre, como decimos, muchos otros no concretados o simplemente cubiertos por respuestas del estilo “textos de todo tipo”.

6.2.4.2. *Objetivos relacionados con materias de traducción*

Los objetivos relacionados con las materias de traducción o “colaterales” son de aprendizaje de la traducción y se han extraído de diversas fuentes, entre ellas, Hurtado (1999) (H)² para la iniciación en la traducción directa, Beeby (1996) (B) para la iniciación en la traducción inversa, Gamero y Hurtado (1999) (G-H) para la traducción directa técnica, y Martínez Melis (2001) (MM) y los elaborados por la autora de esta tesis (RI – Rodríguez Inés) para la traducción inversa especializada.

A continuación ofrecemos un listado de los objetivos de aprendizaje relacionados con materias de traducción que hemos seleccionado para tratar en la propuesta didáctica que efectuamos más adelante (cfr. infra *Parte III*):

Iniciación en traducción directa

- Adquirir estrategias de documentación (H)
- Aprender a evaluar las fuentes de documentación (H)
- Conocer y utilizar las herramientas básicas del traductor (H)
- Manejar los útiles informáticos básicos (H)
- Dominar las diferencias de las convenciones de escritura (H)

² Las iniciales que aparecen en este apartado sirven para identificar la autoría de los objetivos.

- Dominar los elementos de discrepancia morfosintáctica (H)
- Dominar las diferencias en cuanto a los mecanismos de coherencia y cohesión (H)

Traducción directa especializada (técnica)

- Dominar las herramientas del traductor técnico y científico (G-H)
- Dominar la terminología (G-H)
- Captar la importancia de la documentación y las estrategias para adquirirla (G-H)
- Saber identificar y caracterizar distintos tipos y géneros textuales en las dos lenguas (G-H)
- Saber traducir documentos industriales de uso externo (descripción técnica) (G-H)

Iniciación en traducción inversa

- Identificar el funcionamiento de la traducción inversa (RI)
- Introducirse a técnicas de documentación (B)
- Escribir textos en la LM que son gramaticalmente apropiados y pragmáticamente adecuados (B)
- Conocer diferencias de léxico entre la LO y la LM dentro de ciertos campos y registros (B)
- Conocer diferentes tipos de textos (B)

Traducción inversa especializada (científico-técnica)

- Aprender y mejorar el léxico (MM)
- Aprender y mejorar la redacción (MM)
- Identificar la importancia de la naturalidad en el discurso especializado en L2 (RI)
- Captar la clase de información que aporta cada recurso de documentación (RI)
- Utilizar recursos de documentación de todo tipo para resolver problemas de traducción especializada inversa (buscadores, diccionarios, glosarios, tesauros, bases de datos, bibliotecas digitales, corpus) (RI)
- Crear recursos de documentación para resolver problemas de traducción especializada inversa (corpus, glosarios, memorias de traducción) (RI)
- Traducir textos científico-técnicos del español al inglés (MM)
 - o Resolver la traducción de elementos específicos del tema (RI)
 - o Resolver problemas conceptuales de especialidad (RI)
 - o Resolver problemas de traducción que van más allá del nivel del término (RI)

6.3. Contenido disciplinario

A continuación siguen los contenidos previstos para cubrir y alcanzar la adquisición de la competencia en el uso de corpus para resolver problemas de traducción en relación con los cuatro componentes de dicha competencia. Lo que aparece entre paréntesis son algunos ejemplos que hemos querido aportar.

I. Principios básicos relacionados con corpus	
1. Corpus de interés para el traductor	
1.1. Tipos de corpus	
1.1.1.	Corpus según el porcentaje y la distribución de los diferentes tipos de texto: corpus grande o extenso; corpus equilibrado; corpus piramidal; corpus monitor; corpus paralelo (<i>COMPARA</i>); corpus comparable
1.1.2.	Corpus según la especificidad de los documentos: corpus general (<i>BNC</i>); corpus especializado; corpus genérico (<i>Business Letter Corpus</i>); corpus canónico; corpus periódico o cronológico (<i>CREA</i>); corpus diacrónico (<i>CORDE</i>)
1.1.3.	Corpus según la cantidad de texto en cada uno de los documentos: corpus textual; corpus de referencia; corpus léxico
1.1.4.	Corpus según la codificación y la anotación: corpus anotado (<i>BNC, IULA</i>); corpus no anotado (<i>CREA</i>)
1.1.5.	Corpus según la documentación que acompaña a los textos: corpus documentado (<i>TEC</i>); corpus no documentado (<i>WordBanksOnline</i>)
1.2. Corpus gratuitos en Internet	
1.2.1.	Corpus de inglés en Internet (<i>BNC, WordBanksOnline, Brown, Business Letter Corpus, ICE, TEC, Edict</i>)
1.2.2.	Corpus de español en Internet (<i>CREA, Corpus del español, SOL, CORDE</i>)
1.2.3.	Corpus de inglés y español en Internet (<i>IULA, CorpusEye, CLUVI, OPUS</i>)
1.3. Repositorios de textos gratuitos en Internet	
1.3.1.	Repositorios de textos en inglés (<i>Guardian Archive, OTA</i>)
1.3.2.	Repositorios de textos en inglés y español (<i>Project Gutenberg, Medline, Scirus, FreeBooks4Doctors</i>)
1.4. La Web como corpus (<i>Internet pública</i>)	
2. Principios básicos del trabajo con corpus	
2.1. Construcción de corpus	
2.1.1.	Ventajas y desventajas en la construcción de un corpus (paralelo o comparable).
2.1.2.	Textos reales e implicaciones

<ul style="list-style-type: none"> 2.1.3. Los derechos de autor de los textos 2.1.4. Textos con o sin etiquetas 2.1.5. El tamaño de un corpus 2.1.6. La representatividad de un corpus 2.2. Uso de corpus <ul style="list-style-type: none"> 2.2.1. Distintas formas de presentación de concordancias (<i>KWIC, KWAL</i>) 2.2.2. Relación entre los resultados arrojados por un corpus y los textos que lo componen 2.2.3. Formas de extraer información relevante de un corpus 2.2.4. Ventajas documentales de un corpus frente a otros recursos electrónicos 2.2.5. Ventajas del software frente al ser humano 2.2.6. Concienciación sobre la observación del contexto 2.2.7. Índices de frecuencia
II. Construcción de corpus
1. Fuentes de donde recoger textos
<ul style="list-style-type: none"> 1.1. Fuentes en papel: libros, etc. 1.2. Fuentes en formato electrónico excepto Internet: CD-ROM (<i>MLCC, Business-in-a-Box</i>) 1.3. Fuentes en Internet <ul style="list-style-type: none"> 1.3.1. Buscadores (<i>Google, Google Books, Amazon</i>) 1.3.2. Buscadores de buscadores especializados (<i>Buscopio</i>) 1.3.3. Buscadores de textos de especialidad (<i>Medline, Scirus, FreeBooks4Doctors</i>) 1.4. Espíritu crítico sobre las fuentes <ul style="list-style-type: none"> 1.4.1. Ayuda en Internet (<i>Intute: Virtual Training Suite, Internet Detective</i>)
2. Gestión
<ul style="list-style-type: none"> 2.1. Importancia de nombrar lógicamente los archivos de un corpus 2.2. Importancia de ordenar lógicamente los archivos de un corpus
3. Construcción
<ul style="list-style-type: none"> 3.1. Creación de corpus <ul style="list-style-type: none"> 3.1.1. Creación de un corpus comparable monolingüe (ad hoc) 3.1.2. Creación de un corpus comparable bilingüe (ad hoc) 3.1.3. Creación de un corpus paralelo 3.2. Cálculo de la representatividad de un corpus <ul style="list-style-type: none"> 3.2.1. Adición de textos a un corpus

III. Software relacionado con corpus y funciones

1. Tipos de software

- 1.1. Software de conversión de textos en papel a formato electrónico: OCR (*HP, Abby...*)
- 1.2. Software de conversión de formato PDF / imagen a texto (*TEXTfromPDF, Solid Converter*)
- 1.3. Software de ubicación o descarga de textos desde Internet
 - 1.3.1. Software de ubicación y descarga de textos (*Copernic*)
 - 1.3.2. Software de descarga de textos (*WebCopier, Free Download Manager, Flashgot, Downthemall*)
- 1.4. Software de etiquetado de textos
 - 1.4.1. Etiquetado morfológico (*CLAWS*)
 - 1.4.2. Etiquetado sintáctico
- 1.5. Software de alineación de textos (*WinAlign* de *Trados*, *WordSmith Tools*, *MultiConcord*)
- 1.6. Software de conversión de un corpus paralelo en memoria de traducción (*Translator's Workbench* de *Trados*)
- 1.7. Software de extracción de terminología (*SystemQuirk*)
- 1.8. Software de cálculo de la representatividad de un corpus (*ReCor*)
- 1.9. Software de extracción de concordancias de Internet (*WebCorp*, *WebCONC*, *Amazon*, *FreeBooks4Doctors*)
- 1.10. Software de análisis de corpus: programas de instalación y acceso local
 - 1.10.1. Software gratuito de análisis de corpus (*AntConc*, *SCP*, *InfoRapid*, *KwicFinder*)
 - 1.10.2. Software comercializado de análisis de corpus (*WordSmith Tools*, *MultiConcord*)

2. Software de análisis de corpus: funciones ofrecidas por el software de acceso local

- 2.1. Extracción de concordancias (simples, con palabra de contexto, bilingües paralelas, etc.)
 - 2.1.1. Función *Concord* (de *WordSmith Tools*)
 - 2.1.2. Función *Concordance* (de *AntConc*)
 - 2.1.3. Función *Búsqueda avanzada* (de *Adobe Reader*)
 - 2.1.4. Función *Concordance* (de *Translator's Workbench* de *Trados*)
 - 2.1.5. Función *Search* (de *MultiConcord*)
 - 2.1.6. Función *Search* (de *InfoRapid Search & Replace*)
- 2.2. Visualización del contexto
 - 2.2.1. Función *Grow/Shrink* (de *WordSmith Tools*)
- 2.3. Reordenación del contexto

<ul style="list-style-type: none"> 2.3.1. Función <i>Resort</i> (de <i>WordSmith Tools</i>) 2.3.2. Función <i>Kwic sort</i> (de <i>AntConc</i>) 2.4. Extracción de colocaciones <ul style="list-style-type: none"> 2.4.1. Función <i>Collocates</i> (de <i>WordSmith Tools/AntConc/SCP</i>) 2.5. Extracción de agrupaciones <ul style="list-style-type: none"> 2.5.1. Función <i>Clusters</i> (de <i>WordSmith Tools/AntConc/SCP</i>) 2.6. Extracción de la progresión de aparición de una expresión <ul style="list-style-type: none"> 2.6.1. Función <i>Plot</i> (de <i>WordSmith Tools/AntConc/SCP</i>) 2.7. Extracción de listas de palabras <ul style="list-style-type: none"> 2.7.1. Listas de palabras por orden alfabético 2.7.2. Listas de palabras por orden de frecuencia (de <i>WordSmith Tools/AntConc/SCP</i>) 2.7.3. Listas con información estadística (de <i>WordSmith Tools/AntConc/SCP</i>) 2.8. Alineación de textos <ul style="list-style-type: none"> 2.8.1. Función <i>Viewer & Aligner</i> (de <i>WordSmith Tools</i>)
IV. Resolución de problemas de traducción utilizando corpus
1. Selección del corpus más apropiado
2. Resolución de problemas de traducción de la lengua general y especializada
<ul style="list-style-type: none"> 2.1. Información conceptual de la lengua general y especializada <ul style="list-style-type: none"> 2.1.1. Extracción de definiciones: hiperónimos, etc. 2.1.2. Extracción de tipos: hipónimos, etc. 2.1.3. Extracción de características 2.2. Elementos microtextuales de la lengua general y especializada <ul style="list-style-type: none"> 2.2.1. Extracción de vocabulario general y términos de especialidad 2.2.2. Extracción de colocaciones 2.2.3. Extracción de fraseología 2.2.4. Extracción de enunciados ilocutivos: agradecer, ordenar (dar instrucciones, indicaciones...) 2.2.5. Inspiración para la traducción creativa 2.3. Elementos macrotextuales <ul style="list-style-type: none"> 2.3.1. Extracción de convenciones de género 2.3.2. Extracción de elementos discursivos 2.4. Elementos de discrepancia entre dos lenguas <ul style="list-style-type: none"> 2.4.1. Discrepancia en convenciones de escritura 2.4.2. Discrepancia en mecanismos de coherencia y cohesión 2.4.3. Discrepancia morfosintáctica

Tabla 10. Contenido disciplinario asociado a la competencia “uso de corpus electrónicos para traducir”.

6.3.1. Desarrollo del contenido

I. Principios básicos relacionados con corpus

Estos contenidos persiguen que el alumno aprenda sobre corpus y comprenda los principios básicos de esta metodología que le permitirán traducir con estos recursos.

1. Corpus de interés para el traductor

1.1. Diversos tipos de corpus

Conocer los distintos tipos de corpus y las diferencias que hay entre ellos a distintos niveles como, por ejemplo, entre un corpus permanente y uno desechable, un corpus general y uno especializado, o un corpus paralelo y uno comparable, y las distintas clasificaciones que se pueden hacer según los criterios que se apliquen.

1.2. Corpus gratuitos en Internet

Conocer la existencia de corpus de diversos tipos que se encuentran disponibles en Internet.

1.3. Repositorios de textos gratuitos en Internet

Conocer la existencia de sitios web donde se puede acceder a una gran cantidad de textos de diverso tipo que no son un corpus pero sí pueden llegar a formarlo si se descargan a una unidad local siguiendo unos criterios.

1.4. La Web como corpus (*Internet pública*)

Conocer la Internet pública, accesible a través de navegadores, que es posiblemente el “corpus” más extenso que existe. Sin embargo, su heterogeneidad en cuanto al tipo y calidad de textos que contiene hace necesario conocerla bien para poder sacarle provecho, así como también conocer herramientas, aparte de los navegadores, que permitan recuperar información útil.

2. Principios básicos del trabajo con corpus

Se trata de que el alumno se introduzca a la metodología de corpus empezando por conocer y comprender algunas cuestiones fundamentales sobre los corpus, su funcionamiento e implicaciones metodológicas para poder así trabajar con estos recursos.

2.1. Construcción de corpus

Identificar lo que implica la construcción de un corpus, sus ventajas y desventajas, y las implicaciones de trabajar con corpus de textos reales, etiquetados o no, grandes o pequeños.

2.2. Uso de corpus

Conocer las distintas formas de presentación de unas concordancias, como el formato KWIC (*Key Word in Context*), el KWAL (*Key Word And Line*) y otros.

Captar las ventajas documentales que pueden proporcionar los corpus frente a otros recursos (electrónicos) comparando la información que proporcionan unos y otros para descubrir que un corpus (p.ej. comparable monolingüe) puede ser un recurso más completo. Darse cuenta de las posibilidades que ofrece un corpus (p.ej. paralelo) a la hora de buscar distintos equivalentes, mantener la consistencia en la traducción de determinados elementos, etc.

Entender lo que un programa informático puede hacer mejor que un humano en cuanto a corpus, sobre todo en lo que respecta a análisis cuantitativo y distributivo.

Captar la relación entre los resultados arrojados por el corpus y los textos que lo componen. Es muy importante que el alumno establezca la relación entre el corpus que utilice y los resultados que le puede ofrecer.

Captar que un corpus más grande no es siempre mejor. Darse cuenta de que un corpus grande será necesario para algunos tipos de búsquedas de información, pero que, en otros casos, un corpus pequeño, equilibrado y específico de un género o tema, puede ser de mayor utilidad.

Captar la forma como se pueden dirigir preguntas a un corpus para poder extraer resultados interpretables. Ser capaz de convertir la pregunta que se desea hacer al corpus en una cadena de caracteres que se pueden introducir en un programa de análisis de corpus para extraer información.

Saber qué parte del cotexto es importante en una concordancia y cómo reordenarlo para facilitar la búsqueda de información. Captar que en una concordancia el co(n)texto de la palabra clave puede dar mucha información si se sabe hacia dónde dirigir la atención y hacia dónde reordenar el contexto.

Interpretar los índices de frecuencia. Entender que el hecho de que una expresión o palabra aparezca más veces que otra en un corpus no significa necesariamente que la primera sea “más correcta”, al igual que si algo no aparece en un corpus tampoco significa necesariamente que no exista o esté mal, aunque en ambos casos estos resultados pueden dar información indicativa de la frecuencia de uso.

II. Construcción de corpus

1. Fuentes de donde recoger textos

Se trata de que el alumno conozca fuentes o sepa encontrarlas (direcciones de Internet, publicaciones electrónicas, etc.) de donde extraer textos con el fin de crear un corpus.

1.1. Fuentes en papel: libros, etc.

No desdeñar la utilidad de las fuentes de documentación en papel aunque el énfasis se pone en los recursos electrónicos. Conocer y utilizar estas fuentes en papel y, en este caso, saber aprovecharlas para la creación de un corpus electrónico.

1.2. Fuentes en formato electrónico excepto Internet: CD-ROM (*MLCC*, *Business-in-a-Box*)

Conocer o saber encontrar fuera de la Web fuentes en formato electrónico que se puedan utilizar como corpus o de donde se puedan extraer textos para formar uno.

1.3. Fuentes en Internet

Conocer o saber buscar fuentes de textos en Internet, no sólo a través de buscadores generales, sino también especializados.

1.4. Espíritu crítico sobre las fuentes

No sólo conocer fuentes de donde recoger textos, sino ser capaz de evaluar su calidad.

2. Gestión

Se trata de que el alumno comprenda la importancia de nombrar y ordenar bien los textos de un corpus. Ser consciente de que el lugar donde se guarden (carpetas y subcarpetas, unidades) y el nombre que se dé a los archivos puede ser relevante a la hora de interpretar una concordancia.

3. Construcción

Se trata de que el alumno aprenda a buscar textos, guardarlos y referenciarlos correctamente para construir corpus de distinto tipo que le ayuden a traducir.

3.1. Creación de corpus

Crear corpus útiles para el traductor, especialmente aquellos desechables compuestos por textos paralelos en lengua meta (comparable monolingüe) o en lengua de partida y lengua meta (comparable bilingüe) o textos originales y su traducción (paralelo).

3.2. Cálculo de la representatividad de un corpus

Detectar cuándo se necesita aumentar el tamaño de un corpus según las necesidades documentales que se tengan y qué otros textos se pueden añadir al corpus y cómo.

III. Software relacionado con corpus y operaciones

Se trata de que el alumno conozca y aprenda a manejar distintos tipos de herramientas informáticas fundamentales para trabajar con corpus, junto con las funciones básicas que éstas ofrecen.

1. Tipos de software

1.1. Software de conversión de textos en papel a formato electrónico: OCR (*HP, Abby...*)

Manejar un programa de reconocimiento óptico de caracteres (OCR) en combinación con un escáner para poder convertir textos en papel a formato digital y procesarlos con programas de análisis de corpus.

1.2. Software de conversión de formato PDF o imagen a texto (*TEXTfromPDF, Solid Converter*)

Manejar programas que convierten textos que se encuentran en formatos que no puede procesar un programa de análisis de corpus a otros formatos que sí son compatibles para poder aprovecharlos.

1.3. Software de ubicación y descarga de textos desde Internet

Manejar programas que permiten encontrar textos en Internet y descargarlos a una unidad local. Ejemplos de este tipo de software son *Copernic*, *WebRecorder* o *WebCopier*, o *AltaVista Discovery* o *KwicFinder*, que son programas que pueden buscar documentos que ya se encuentran en una unidad local.

1.4. Software de etiquetado de textos

Manejar programas para etiquetar textos, tanto morfológica como sintácticamente, para después poder realizar búsquedas por etiquetas en corpus.

1.5. Software de alineación de textos (*WinAlign* de *Trados*, *WordSmith Tools*, *MultiConcord*)

Manejar programas que permiten seleccionar textos, un original y su traducción, y alinearlos bien de forma automática o semi-automática, con la posibilidad de tener que hacer labores de post-edición del alineamiento.

Extraer concordancias bilingües paralelas, es decir, realizar búsquedas para extraer concordancias de textos en lengua original con sus traducciones en lengua meta.

Un ejemplo de este tipo de software específico es *MultiConcord*, pero otros programas como *WordSmith Tools* o memorias de traducción también incluyen esta función.

1.6. Software de conversión de un corpus paralelo en memoria de traducción (*Translator's Workbench* de *Trados*)

Manejar programas y funciones que permiten manipular un corpus paralelo para poder importarlo desde un programa de traducción asistida como una memoria de traducción.

1.7. Software de extracción de terminología (*SystemQuirk*)

Utilizar programas o aplicaciones de extracción de n-gramas para identificar candidatos a términos en un corpus.

1.8. Software de cálculo de la representatividad de un corpus (*ReCor*)

Manejar programas que calculan si un corpus es o no representativo una vez se ha construido.

1.9. Software de extracción de concordancias de Internet (*WebCorp*, *WebCONC*, *Amazon*, *Google Books*)

Manejar aplicaciones que funcionan en línea y extraen concordancias de contenido que también está en línea, como documentos de la Internet pública o libros editados.

1.10. Software de análisis de corpus: programas de instalación y acceso local

Manejar programas fundamentales para la explotación de corpus electrónicos, de los cuales, uno de los más conocidos y utilizados es el *WordSmith Tools*, si bien existen otros gratuitos.

2. Software de análisis de corpus: funciones ofrecidas por el software de acceso local

De mayor o menor complejidad, los programas de análisis de corpus suelen compartir una serie de funciones básicas que el alumno debe aprender a manejar para extraer información del corpus.

2.1. Extracción de concordancias (simples, con palabra de contexto, bilingües paralelas, etc.)

Dominar la función de extracción de concordancias, que es el principal medio de obtener información de un corpus, razón por la que es vital para traducir con este recurso. Extraer concordancias simples o con búsquedas truncadas, ya que es importante que el alumno sea capaz de realizar búsquedas sencillas y más complejas a la hora de extraer concordancias. Extraer concordancias con palabra de contexto, es decir, saber afinar en la búsqueda introducida acotándola con palabras relacionadas.

Utilizar las funciones de una memoria de traducción que se asemejan al trabajo con corpus, como la extracción de segmentos bilingües o de concordancias con un programa como, por ejemplo, el *Translator's Workbench* de Trados.

2.2. Visualización del contexto

Utilizar las funciones disponibles para poder ver más o menos contexto de unas concordancias, según interese por el tipo de búsqueda.

2.3. Reordenación del contexto

Utilizar las funciones disponibles para reorganizar el contexto que rodea a la palabra clave, bien hacia la derecha bien hacia la izquierda, según interese, dependiendo de lo que se necesita saber.

2.4. Extracción de colocaciones

Utilizar la función de extracción de palabras que suelen relacionarse con otras, comprender y aprovechar su utilidad para la traducción.

2.5. Extracción de agrupaciones

Utilizar la función de extracción de n-gramas o agrupaciones de palabras que suelen aparecer juntas, comprender y aprovechar su utilidad para la traducción.

2.6. Extracción de la progresión de aparición de una expresión

Utilizar esta función (llamada *Plot* en *WordSmith Tools*, por ejemplo), comprender y aprovechar su utilidad para la traducción.

2.7. Extracción de listas de palabras

Extraer listas de palabras, tanto aquellas ordenadas alfabéticamente como por frecuencia, ya que son una valiosa fuente de información.

Extraer e interpretar información estadística, es decir, utilizar la función de extracción de estadísticas y aprovechar datos como el ratio formas/palabras, que ofrece información sobre el grado de repetición del léxico en un texto o corpus.

2.8. Alineación de textos

Utilizar la función que permite alinear textos para luego poder realizar búsquedas y extraer concordancias bilingües paralelas.

IV. Resolución de problemas de traducción utilizando corpus

Se trata de que el alumno resuelva problemas de todo tipo utilizando corpus. Tras identificar los problemas de traducción solucionables con un corpus, el alumno debe saber elegir qué tipo de corpus le será útil, hacer búsquedas de calidad ajustadas a los objetivos, interpretar y manejar los datos para obtener respuestas fiables y utilizables; en definitiva, debe ser capaz de buscar de forma eficaz en el corpus, entender los datos que muestra y saber cómo utilizarlos para traducir.

1. Selección del corpus más apropiado

Una vez identificados los problemas, el alumno debe ser capaz de establecer una conexión entre las necesidades de documentación o traducción que tiene y el tipo de corpus que podría darle las soluciones, es decir, saber que un problema X podría resolverse con un corpus Y.

2. Resolución de problemas de traducción de la lengua general y especializada

El alumno debe saber extraer información diversa de un corpus y tomar decisiones en función de los datos arrojados por el corpus analizando y valorando su utilidad de cara a una traducción.

2.1. Información conceptual de la lengua general y especializada

Extraer definiciones, características, explicaciones, tipos, etc.

2.2. Elementos microtextuales de la lengua general y especializada

Extraer vocabulario y términos de especialidad, colocaciones, fraseología tanto de la lengua general como especializada útiles para la traducción.

Encontrar datos que ilustren casos de creatividad en los textos o den información sobre su uso para ayudar con la traducción de elementos lingüísticos creativos.

2.3. Elementos macrotextuales

Solucionar cuestiones de traducción relacionadas con el género textual y el estilo con corpus.

2.4. Elementos de discrepancia entre dos lenguas

Encontrar distintas soluciones para elementos de discrepancia entre dos lenguas utilizando corpus para, por ejemplo, conocer el uso de convenciones de escritura.

6.4. Marco metodológico

Como ya hemos expresado anteriormente, el marco metodológico escogido para llevar a cabo los objetivos presentados ha sido el Enfoque por tareas y, en concreto, el Enfoque por tareas de traducción (Hurtado 1993, 1995, 1996).

El enfoque por tareas de traducción destaca por su flexibilidad a la hora de elaborar las unidades didácticas ya que permite integrar objetivos, actividades, medios y evaluación, todo ello situando al estudiante en el centro del proceso de aprendizaje. Según la misma autora, los aspectos más relevantes que nos proporciona este enfoque son:

- Salva la distancia que se produce en otras propuestas entre objetivos y metodología, proporcionando realmente una metodología activa.
- Arma al estudiante, al introducir tareas posibilitadoras (pedagógicas) que le ayudan a resolver la tarea final (la traducción de determinado género textual, por ejemplo); se consigue así una pedagogía centrada en el recorrido de procesos.
- Encadena constantemente, a través de las tareas, una metodología viva en la que el alumno no sólo aprende haciendo, y capta principios, sino que aprende a resolver problemas y adquiere estrategias de aprendizaje.
- Se logra realmente una pedagogía centrada en el estudiante, que, además, le hace responsable de su propio aprendizaje y, por consiguiente, más autónomo.
- Permite incorporar constantemente tareas de evaluación formativa para el estudiante (que aprende a autoevaluarse y medir sus propias posibilidades) y para el profesor (que puede evaluar su enseñanza).

Aspectos más relevantes del enfoque por tareas de traducción (Hurtado 1999: 56).

El enfoque por tareas de traducción es un marco excelente para organizar el aprendizaje del uso de corpus para traducir. El aprendizaje progresivo del trabajo con corpus se presta perfectamente a una distribución en tareas, y éstas en unidades didácticas. Unas tareas son más sencillas, se centran en la consecución de objetivos concretos, en aspectos determinados del uso de corpus, y se encadenan para ir acumulando conocimientos y habilidades; otras tareas son más complejas (tareas finales) y en ellas se espera que el estudiante ponga en práctica todo lo adquirido en el proceso de aprendizaje que ha tenido lugar a lo largo de la unidad didáctica.

En este sentido, todas las unidades didácticas plantean varios objetivos y contenidos que trabajan la competencia de uso de corpus para traducir y alguna competencia genérica, además de que también muchas tareas cubren más de un objetivo. Este hecho, lejos de sobrecargar las tareas con varios objetivos, les da un mayor realismo ya que, en el caso del trabajo con corpus, aislar los objetivos y tratar cada uno en una tarea independiente hubiera resultado artificial, incompleto o incluso imposible. Por supuesto, huelga decir que la propuesta, tanto de objetivos como de estructuración de los mismos en dos fases pedagógicas, en tareas y unidades didácticas es flexible. Lo que más adelante se presenta como una propuesta concreta está abierta a modificaciones en función de, entre otras cosas, los estudiantes y las circunstancias que les rodeen. En palabras de Hurtado, “la evaluación y la negociación continua con los estudiantes son las mejores guías para la elaboración de las unidades y para la modificación de objetivos” (1999: 57).

6.5. Materias de aplicación

Las materias para las que se proponen las unidades didácticas diseñadas son:

- Iniciación a la traducción directa B/C-A (inglés-español)
- Iniciación a la traducción inversa B-A (español-inglés)
- Traducción directa especializada A-B (inglés-español)
- Traducción inversa especializada B-A (español-inglés)

Si bien todos los contenidos propuestos incluyen el trabajo con corpus electrónicos y tienen la finalidad de que el alumno adquiera conocimientos y habilidades en el uso de corpus para traducir, la consecución de algunos de estos objetivos también proporciona conocimientos sobre el funcionamiento de la traducción y habilidades traductoras. Así,

se han incluido contenidos variados que sirven para materias tanto de iniciación como de traducción especializada en directa e inversa. Nuestra propuesta de unidades didácticas cubre por tanto, aparte de los objetivos principales relacionados con el uso de corpus para traducir, otros objetivos relacionados con las materias de traducción.

6.5.1. La traducción directa

Si entendemos el uso de corpus para traducir en relación con la competencia instrumental o el manejo de recursos documentales e informáticos en traducción, entonces veremos que los objetivos que plantean autores para materias concretas como la iniciación a la traducción o las variedades de traducción especializada (jurídica, científico-técnica, etc.) incluyen siempre objetivos de documentación. En otras palabras, nuestros objetivos de uso de corpus para traducir están en consonancia o se complementan con los objetivos que otros han planteado para el aprendizaje de la traducción en su fase de iniciación o de traducción especializada.

6.5.1.1. Iniciación en traducción directa

Por ejemplo, para Hurtado (1996, 1999, 2007), entre los principios metodológicos básicos que el estudiante debe asimilar en la iniciación a la traducción está la importancia de los conocimientos extralingüísticos y la necesidad de documentación, y entre los aspectos profesionales básicos que éste debe conocer está el uso de las herramientas básicas del traductor, de donde destacamos el conocer y evaluar las fuentes de documentación, así como adquirir estrategias de búsqueda de información.

6.5.1.2. Traducción directa especializada

Ya dentro de la traducción especializada, para Gamero (1996), el objetivo principal en traducción científico-técnica es que los estudiantes sean capaces de enfrentarse a textos de esta especialidad en general, proporcionándoles las herramientas suficientes e indispensables, de modo que puedan especializarse en alguno(s) de los campos de la ciencia y de la técnica en su carrera profesional, a lo que Vega añade que “la principal de tales herramientas es la búsqueda documental y terminológica” (Vega 1999: 242). Para poner un ejemplo, tres de los seis objetivos generales que plantea este último autor en su concepción de la enseñanza de la traducción científica y técnica están directamente relacionados con la documentación. Éstos son:

- Saber documentarse adecuadamente sobre cualquier campo temático y conocer el grado de fiabilidad y las ventajas e inconvenientes de los distintos tipos de fuente documental.
- Saber consultar y utilizar adecuadamente las fuentes de datos terminológicas.
- Saber utilizar los medios técnicos que el mercado de la traducción impone en cada momento y estar en condiciones de aprender a utilizar medios nuevos.

Tres de los seis objetivos generales para la enseñanza de la traducción científica y técnica según Vega (1999: 249).

Gamero y Hurtado (1999), a quienes el propio Vega cita, ofrecen una lista de objetivos para la enseñanza de la traducción técnica y científica más exhaustiva y de la cual merece ser destacada una parte por la relación que guarda con la cuestión de la documentación. Así, Gamero y Hurtado hablan, entre otros, de los siguientes objetivos generales, específicos e intermedios en lo que concierne a la documentación:

- Conocer los aspectos profesionales de la traducción técnica y científica
 - Dominar las herramientas del traductor técnico y científico.
 - Conocer y saber utilizar las fuentes de documentación.
 - Saber manejar los diccionarios especializados en soporte electrónico.
 - Saber acceder y trabajar en Internet.
 - Saber utilizar aplicaciones informáticas útiles para la traducción.
 - Desarrollar un espíritu crítico.
- Asimilar los principios metodológicos de la traducción técnica y científica
 - Dominar la terminología.
 - Saber hallar equivalencias utilizando los recursos terminológicos apropiados.
 - Captar la importancia de la documentación y las estrategias para adquirirla.
 - Valorar la necesidad de documentación.
 - Conocer y saber aplicar estrategias de documentación.

Objetivos generales, específicos e intermedios relacionados con la documentación en la enseñanza de la traducción científico-técnica (Gamero y Hurtado 1999).

De los objetivos de aprendizaje para la traducción científico-técnica aquí seleccionados la mayoría encuentran cobertura en alguna de las unidades didácticas diseñadas en el marco de esta tesis doctoral. Debido a la tardía y escasa repercusión y aplicación de los corpus electrónicos en la enseñanza de la traducción, no es de extrañar que prácticamente nunca aparezcan mencionados entre los objetivos de aprendizaje de ninguna especialidad de traducción. Sin embargo, sí aparecen destacados el manejo de herramientas informáticas, recursos terminológicos, obras de documentación, etc. que, al fin y al cabo, vienen a solaparse con algunas de las utilidades de los corpus de textos.

Una obra dedicada a la traducción directa de especialidad científico-técnica en la que sí se menciona la utilidad de los corpus electrónicos ad hoc a la hora de buscar terminología, fraseología, símbolos, etc. para la traducción es la de Montalt (2005). En cuanto a las necesidades de información del traductor y los tipos de recursos donde encontrarla, reproducimos a continuación una tabla del propio autor a la que sólo queremos añadir que algunos de los recursos que apunta podrían fácilmente formar parte de un corpus (p.ej. entradas de enciclopedias, libros, páginas web, textos paralelos).

<i>Tipus d'informació requerida</i>	<i>Recursos on cercar-la</i>
Temática	Enciclopèdies, llibres de text, tractats, pàgines web, expert en el tema, etc.
Sobre UCE (unitats de coneixement especialitats)	Diccionaris, vocabularis, enciclopèdies, expert en el tema, etc.
Sobre el gènere textual	Textos paral·lels en la llengua d'arribada, expert en el tema, etc.
Lingüística	Gramàtiques, manuals d'estil, diccionaris de dubtes, etc.

Tabla 11. Tipos de información requerida y recursos donde encontrarla (Montalt 2005: 210).

6.5.2. La traducción inversa

Mención aparte requiere la materia de traducción inversa y su definición de objetivos de aprendizaje. En primer lugar, destaca la escasez de estudios de traducción inversa en cuanto al proceso que sigue, las competencias implicadas o los objetivos a perseguir. En relación con la competencia traductora en inversa cabe destacar a los ya citados Beeby, Campbell y Martínez Melis. En muchas ocasiones, las únicas valoraciones que se realizan respecto a la traducción inversa son que su práctica en el mundo real es desaconsejable, los resultados, de mala calidad, y que su enseñanza no responde a una necesidad del mercado. Muchos son también los autores que, sin embargo, defienden la formación en traducción inversa (Kelly et al. 2003) y todos los centros donde ésta se ofrece, al menos en España.

6.5.2.1. *Iniciación en traducción inversa*

En general, en cuanto a la definición de objetivos de aprendizaje para la iniciación o la traducción inversa especializada, no son abundantes las publicaciones donde éstos quedan reflejados. Beeby (1996a, 1996b, 2003) es una de las pocas autoras en España que ha dedicado gran parte de su investigación a la traducción inversa y su docencia. Entre los objetivos específicos que marca para la clase de traducción inversa y más nos interesan está “la introducción a técnicas de documentación, primero a partir de textos paralelos, artículos enciclopédicos y ‘expertos’ invitados [...], y después individualmente” (1996b: 65). La principal preocupación de Beeby en este estadio, el de la iniciación a la traducción inversa, es el de concienciar al alumno de todos los parámetros del proceso de la traducción, un proceso de tres etapas (comprensión del TO, desverbalización y reformulación del TM). Para esta autora el proceso de traducción es el mismo se trate de una traducción directa o inversa o de cualquier combinación lingüística, quedando en manos del traductor decidir qué etapa presenta más dificultades y requiere más atención.

6.5.2.2. *Traducción inversa especializada*

Martínez Melis (2001: 194 y ss) clasifica los objetivos de aprendizaje de la traducción inversa en las cuatro categorías ya definidas por Hurtado (cfr. supra 5.3.4.), es decir, objetivos metodológicos, lingüísticos, profesionales y textuales. Entre éstos, los más interesantes en nuestro contexto de uso de corpus son los profesionales y, en concreto, los que hacen referencia a conocer y saber utilizar las herramientas y recursos propios del traductor, que la autora distingue entre “tradicionales” e “informatizados”, y a saber documentarse. El trabajo con corpus implica manejo de software y recursos y aplicación de estrategias de documentación, lo cual está en sintonía con los objetivos de la autora.

Cámara Aguilera (2003) también habla de los objetivos específicos que persigue en el aula de traducción inversa especializada. Entre ellos, el relacionado con la documentación busca “profundizar en la capacidad para identificar fuentes de documentación apropiadas, para valorar su fiabilidad y para utilizarlas en la traducción especializada” (Cámara Aguilera 2003: 211).

Rodríguez y Schnell (2003), sin hablar expresamente de objetivos, sí hablan de competencias que son parte de la competencia traductora. Estas autoras consideran

que para suplir las insuficiencias que se dan en la competencia lingüística y pragmático-discursiva cuando se traduce a la lengua extranjera, están otras competencias como la textual o la documental y que, de hecho, entre éstas dos últimas también hay una complementariedad ya que “la documentación, como materia dedicada a la recopilación, ordenación, análisis y almacenamiento de documentos [...] pasa a constituir un factor clave en el desarrollo de la competencia textual” (2003: 178).

Por último, queremos mencionar a Neunzig y Grauwinkel (2006), que defienden un modelo de proceso de traducción propio para la inversa. Mientras que para la traducción directa asumen una secuencia de análisis, documentación, transferencia y reformulación, para la traducción inversa, y basados en el experimento sobre competencia traductora de PACTE, los autores proponen la siguiente secuencia: análisis, reformulación, documentación y transferencia. En este modelo hipotético-deductivo del proceso de la traducción inversa los autores observan cinco fases, entre las que destacamos la tercera, la preparación de la traducción, en la que nos congratulamos de la mención que hacen del uso de corpus textuales entre los recursos que ofrece Internet a los traductores. Si hablamos de objetivos, Neunzig lo expresa con clarividencia:

“El objetivo final de la asignatura de traducción especializada inversa es lograr que los alumnos lleguen a convencerse que podrán, parafraseando a Jean Delisle, «hacer con (relativa) facilidad lo que creían sólo poder hacer con (grandísimo) esfuerzo»” (Neunzig 2003: 191-192).

En definitiva, como decíamos, no abundan los listados de objetivos de aprendizaje de la traducción inversa y, aún menos, los estudios empíricos que arrojen luz sobre el proceso que sigue. Sin embargo, de una forma u otra, la insistencia en la necesidad de la documentación está mucho más presente que en la traducción directa. Aquí defendemos que en la enseñanza de la traducción inversa tiene que estar muy presente la adquisición de la competencia instrumental con objetivos específicos relacionados con la documentación incluyendo, desde nuestro punto de vista, el manejo de corpus electrónicos.

6.6. Unidades didácticas

Las unidades didácticas completas que se llegaron a diseñar y probar son 7:

- Introducción al uso de corpus electrónicos para traducir.
- Contrastes de uso en un corpus paralelo.
- Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos.
- Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés.
- El corpus ad hoc como remedio para la traducción médica al inglés.
- Corpus y memoria de traducción.
- Sólo utilizar corpus; utilizar corpus solo.

6.6.1. Criterios para la elaboración de las unidades didácticas

6.6.1.1. Criterios generales

La elaboración de las unidades didácticas que se presentarán en esta tesis ha estado gobernada principalmente por los siguientes criterios:

- Variedad
- Progresión

6.6.1.1.1. Variedad

Con la intención de exponer al alumnado a una amplia gama de situaciones de traducción que se acercaran a la realidad, la variedad ha sido uno de los elementos básicos que ha regido prácticamente todas las elecciones relacionadas con el diseño de las unidades didácticas. Así, se puede observar esta variedad consciente en la composición de las unidades:

- Objetivos de aprendizaje: iniciación y consolidación.
- Materias: iniciación en traducción directa e inversa y traducción especializada.
- Campos de especialidad: científico, técnico y económico.
- Materiales: recursos lexicográficos, tipos de corpus, tipos de textos, géneros textuales, software utilizado.

- Dinámica de trabajo: individual, en parejas y en cadena.
- Evaluación: autoevaluación, puesta en común y comentarios del profesor.

De forma esquemática se presenta a continuación una tabla con todas las características de las unidades didácticas diseñadas con el fin de poner de relieve la variedad de elementos mencionada.

UNIDADES DIDÁCTICAS		Introducción al uso de corpus electrónicos para traducir	Contrastes de uso en un corpus paralelo	Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos	El corpus ad hoc como remedio para la traducción médica al inglés	Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés	Corpus y memoria de traducción	Sólo utilizar corpus; utilizar corpus solo
Dirección	Directa							
	Inversa							
Especialidad	General							
	Científica							
	Técnica							
	Económica							
	Jurídica							
Nivel corpus	Iniciación							
	Consolidación							
Tipo de corpus	General							
	Especializado							
	De fabricación propia							
	Ya construido							
	Comparable monolingüe							
	Comparable bilingüe							
	Paralelo							
Software utilizado		<i>WordSmith Tools</i>	<i>MultiConcord</i>	<i>Free Download Manager, WordSmith Tools, InfoRapid</i>	<i>WordSmith Tools</i>	<i>WordSmith Tools</i>	<i>WebCopier, WordSmith Tools, WinAlign + Translator's Workbench de Trados</i>	<i>WebCopier, Antconc, TagEditor + Translator's Workbench de Trados</i>

Tabla 12. Variedad de algunos elementos incluidos en las unidades didácticas elaboradas.

6.6.1.1.2. Progresión

La progresión de las unidades didácticas se basa en la dificultad en tanto que comienza con unidades que incluyen objetivos de iniciación al uso de corpus para traducir para pasar a la consolidación. Sin embargo, como ya se ha expuesto anteriormente, al estar las unidades concebidas unas para la traducción directa y otras para la inversa, unas más aptas para la iniciación a la traducción (o “traducción general”) y otras para la traducción especializada sólo se puede pensar en ellas como en una serie consecutiva teniendo en cuenta lo siguiente:


Progresión de las unidades didácticas según el nivel de uso de corpus	
Iniciación  Consolidación	❶ Introducción al uso de corpus electrónicos para traducir (directa)
	❷ Contrastes de uso en un corpus paralelo (directa)
	❸ Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés (inversa)
	❹ Utilidad de un corpus comparable bilingüe para traducir textos técnicos (directa)
	❺ El corpus ad hoc como remedio para la traducción médica al inglés (inversa)
	❻ Corpus y memoria de traducción (inversa)
	❼ Sólo utilizar corpus; utilizar corpus solo (inversa)

Tabla 13. Progresión de las unidades didácticas según el nivel de uso de corpus.

Por otra parte, aunque la intención inicial no era realizar toda una serie de unidades que pudieran ser secuenciadas de forma consecutiva ya que, entre otras cosas, les diferencian aspectos de direccionalidad, podríamos proponer una progresión en base al tipo de objetivos relacionados con las materias de traducción o “colaterales” que incluyen. Así, bien podría comenzarse en un curso de traducción directa por la unidad de introducción al uso de corpus para traducir y la unidad sobre contrastividad ya que son las que contienen objetivos generales metodológicos que establecen las bases necesarias para avanzar hacia la traducción especializada, que podría estar representada por la unidad de traducción técnica que incluye básicamente un objetivo general textual. En cuanto a la iniciación en traducción inversa, ésta se daría necesariamente con la unidad dedicada a concienciar al alumno sobre cuestiones

metodológicas sobre el proceso y la práctica de la inversa. Después seguirían la unidad de traducción de textos médicos, bastante dirigida en lo que a construcción y uso de corpus se refiere, la unidad en la que se une el uso de corpus y de memorias de traducción, y finalmente la unidad en la que se propone un proyecto extenso y para realizar de forma bastante autónoma, que incluye tanto objetivos nuevos como de reciclaje.


Progresión de las unidades didácticas según la dirección y el nivel de competencia en traducción		
	Directa	Inversa
Iniciación  Consolidación	❶ Introducción al uso de corpus electrónicos para traducir	❶ Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés
	❷ Contrastes de uso entre inglés y español	❷ El corpus ad hoc como remedio para la traducción médica al inglés
	❸ Utilidad de un corpus comparable bilingüe para traducir textos técnicos	❸ Corpus y memoria de traducción ❹ Sólo utilizar corpus; utilizar corpus solo

Tabla 14. Progresión de las unidades didácticas según la dirección y el nivel de competencia en traducción.

Finalmente, cabe remarcar que de estos dos criterios de progresión se desprende que las unidades didácticas diseñadas podrían:

- a) aislarse en una materia, “Uso de corpus para traducir”, es decir, combinar las unidades independientemente de si están pensadas para la traducción directa o inversa para centrarse en cubrir contenidos de uso de corpus orientado a la resolución de problemas de traducción; o
- b) ubicarse en distintas materias, como apunta la última tabla.

6.6.1.2. Criterios específicos

Apuntadas ya las líneas generales utilizadas para el diseño de las unidades didácticas (variedad y progresión), en este apartado se exponen de una forma un poco más detallada dichos criterios, la lógica que hay detrás de cada unidad, las materias en las que podrían desarrollarse las unidades, los tipos de materiales, etc., si bien es cierto que cada unidad didáctica está explicada en toda su amplitud en la *Parte III* de esta tesis.

De acuerdo con información extraída de nuestra experiencia docente y profesional, así como de los dos sondeos a pequeña escala realizados sobre la práctica profesional de la traducción (cfr. supra 2.4. y 6.2.4.1.) y el amplísimo estudio de mercado realizado por la Agrupación de Centros Especializados en Traducción, se consideraron relevantes para el alumnado de traducción una serie de campos temáticos, tipos de tareas, recursos de documentación, etc.

6.6.1.2.1. Selección de objetivos de aprendizaje

1) Objetivos relacionados con el uso de corpus para traducir

Por un lado, la selección de los 4 objetivos generales que hemos establecido en nuestra propuesta responden a la ya citada observación de Aston (2001) que recoge que el alumno debe adquirir conceptos, metodología y técnica en lo que atañe al uso de corpus en lenguas o traducción.

Por otro lado, la propuesta de objetivos se ha diseñado con flexibilidad. En este sentido, en algunos casos, la consecución de un objetivo específico sería imposible sin la intervención de otro. Además, algunas tareas se han diseñado de tal manera que sean “exportables”, es decir, que se puedan integrar en otras unidades didácticas, como es el caso de las tareas relacionadas con los objetivos de aprendizaje de manejo de software.

A continuación se comentan las razones principales por las que se han seleccionado determinados objetivos dentro de cada unidad.

UD “Introducción al uso de corpus electrónicos para traducir”

Según los planteamientos de Hurtado (1999), es en la fase de iniciación a la traducción donde se adquieren las bases lingüísticas y metodológicas principales. Dentro de las

unidades didácticas diseñadas para la iniciación en el uso de corpus para traducir (inglés-español), ésta se centra en cuestiones de cariz metodológico y encuentra su razón de ser en la necesidad de que los alumnos adquieran unos conocimientos sobre el trabajo con corpus y unas habilidades básicas en su manejo con el fin de empezar a explotar todo ello para la traducción.

UD “Contrastes de uso en un corpus paralelo”

Esta unidad, que también se enmarca dentro de la fase de iniciación en el uso de corpus e iniciación en la traducción directa, incluye contenidos relacionados sobre todo con el trabajo contrastivo con un corpus paralelo y sus implicaciones, ventajas y desventajas, lo cual supone parte de los principios metodológicos básicos del uso de corpus para traducir y de la iniciación en traducción.

UD “Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”

Esta unidad contiene objetivos de aprendizaje de iniciación en el uso de corpus para traducir, en este caso, de la lengua materna a la extranjera. Éstos se ciñen a aspectos muy concretos y básicos del trabajo con corpus y que están relacionados con cómo se puede interpretar la frecuencia de aparición de un elemento en un corpus o dónde fijarse en una concordancia para obtener la información que se quiere, así como conocer corpus diversos. Todo ello es necesario para que el estudiante fije las bases del trabajo con corpus y pueda avanzar hacia usos más complejos.

UD “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”

Moviéndonos ya en la fase de consolidación en el uso de corpus para traducir, esta unidad de traducción directa de especialidad técnica se ha planteado desde la necesidad del estudiante de adquirir una metodología de valoración, recopilación, organización y aprovechamiento de textos paralelos en lengua de partida y de llegada que le sean de ayuda para traducir. El manejo de un corpus comparable bilingüe forma parte de la competencia en el uso de corpus para traducir.

Esta unidad, junto con las tres que siguen, implican la construcción de corpus diversos, una de las habilidades que también forman parte de la competencia objeto de estudio. En este sentido, cabe recordar aquí el sondeo realizado entre 16 agencias de traducción españolas (cfr. supra 2.4.) que revelaba que en todas ellas se realizaban labores de recopilación, almacenamiento o alineación de textos para futuros usos.

UD “El corpus ad hoc como remedio para la traducción médica al inglés”

Esta unidad contempla contenidos más centrados en la compilación de un corpus comparable, en este caso monolingüe, para resolver problemas de traducción a nivel micro y macrotectual dentro del campo médico. La consecución de los objetivos relacionados es crucial para, entre otras cosas, hacer del estudiante un trabajador gradualmente más autónomo.

Esta unidad de traducción inversa especializada y las dos que siguen se centran en la compilación de corpus especializados en lengua meta, en el aprovechamiento de los corpus para resolver problemas y en la asimilación por parte del alumno, especialmente en lo que se refiere a traducción inversa, de que sólo un corpus puede solucionar ciertos problemas de traducción que van más allá del vocabulario especializado.

UD “Corpus y memoria de traducción”

Muy importante también en relación con el aprovechamiento de textos es otro de los objetivos de aprendizaje propuestos, saber convertir un corpus en memoria de traducción o, dicho de otra forma, saber explotar un corpus paralelo con herramientas propias de la traducción asistida por ordenador, objetivos que, junto a otros, se tratan en esta unidad.

De nuevo, cabe citar el sondeo realizado entre 16 agencias de traducción españolas (cfr. supra 2.4.) que, entre otras cosas, reveló que casi todas ellas utilizaban programas de traducción asistida para manejar y aprovechar los textos que acumulan. Además, la observación de la realidad profesional revela que las ofertas de trabajo para traductores por parte de empresas de traducción suelen incluir como requisito el manejo de estos programas.

UD “Sólo utilizar corpus; utilizar corpus solo”

En este caso, la elaboración de esta unidad didáctica viene motivada por la necesidad de los estudiantes de saber evaluar las fuentes que consultan porque se están documentando sobre todo en su lengua extranjera. La selección, evaluación de recursos y documentos, así como la explotación de corpus y recursos similares para producir traducciones de una calidad óptima, aunque sean hacia la lengua extranjera, son habilidades claves que un alumno debe haber adquirido al finalizar la formación en traducción especializada.

Además, según el estudio de mercado publicado en 2005 por la Agrupación de Centros Especializados en Traducción (ACT) apunta que las tecnologías para la traducción “constituyen una de las expectativas de desarrollo del sector de la oferta [y que] la capacidad tecnológica es uno de los puntos fuertes que la demanda valora positivamente en un proveedor” (ACT 2005: 242), lo cual nos sirve para justificar en parte los objetivos y contenidos más tecnológicos de nuestra propuesta.

2) Objetivos relacionados con materias de traducción

UD “Introducción al uso de corpus electrónicos para traducir” y UD “Contrastes de uso en un corpus paralelo”

Con estas dos unidades de iniciación en traducción directa el estudiante empieza a aprender a utilizar corpus en traducción estableciendo al mismo tiempo los principios metodológicos de la traducción y los del trabajo con corpus.

Asimismo, la segunda unidad también incorpora objetivos relacionados con cuestiones de contrastividad entre el inglés y el español, lo cual representa una de las bases lingüísticas que el alumno debe adquirir en su iniciación a la traducción.

UD “Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”

Como ha confirmado el pequeño sondeo realizado entre traductores autónomos (cfr. supra 6.2.4.1.), existe una gran probabilidad de que al traductor se le solicite realizar alguna traducción inversa en algún momento de su carrera profesional. Y si bien al nivel de iniciación en traducción no se realiza un trabajo textual ni temático en profundidad, esta unidad incluye, por ejemplo, un pequeño ejercicio con un corpus de correspondencia comercial y otro con un corpus ad hoc de inteligencia emocional, lo cual viene justificado porque “en segundo orden de volumen [de trabajo] se encuentran los textos de tipo comercial y financiero” (ACT 2005: 108).

UD “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”

Esta unidad permite al alumno dotarse de conocimientos específicos sobre un género técnico y práctica en el trabajo con textos que se adscriben a tal género. Además, la experiencia docente y profesional, junto con el estudio de la ACT (2005: 108, 109) nos revelan que las traducciones técnicas (inglés-español) son las que reciben más demanda en nuestro país.

UD “El corpus ad hoc como remedio para la traducción médica al inglés”

Trabajar el campo temático de la medicina en unidades de traducción inversa puede ser realista dada la función divulgadora de la lengua inglesa. Si bien es indudable que la traducción médica al español tiene su nicho de mercado, también lo es que la cantidad y accesibilidad a los textos médicos en inglés es mucho mayor. Como opina Mayoral,

“... el inglés es la lengua de expresión para cualquier científico que se precie y que quiera que su trabajo se divulgue entre la comunidad científica. Por lo tanto, el traductor no va a encontrar un texto paralelo en español que le facilite la traducción como sí ocurre en el caso de la traducción de textos jurídicos, administrativos y comerciales” (Mayoral 1997: 151).

UD “Corpus y memoria de traducción”

Del sondeo realizado entre traductores autónomos (cfr. supra 6.2.4.1.) se desprende que la especialidad técnica es una de las más solicitadas cuando se da una situación de traducción inversa.

UD “Sólo utilizar corpus; utilizar corpus solo”

Por nuestra experiencia docente y profesional creemos que hay una gran probabilidad de que al traductor se le solicite traducir una página web de contenido especializado. De hecho, el estudio de la Agrupación de Centros Especializados en Traducción, entre sus conclusiones, apunta que “la traducción asistida (...), y la multimedia [en la que incluyen las páginas web] son las tecnologías más extendidas (ACT 2005: 108).

6.6.1.2.2. Selección de materiales

Los materiales utilizados en las unidades didácticas han sido variados:

- Textos: de tipo descriptivo, narrativo, expositivo, argumentativo e instructivo, y de géneros textuales como noticias de prensa, páginas web, extractos de artículos científicos, etc.
- Recursos lexicográficos: diccionarios generales y especializados, monolingües y bilingües, tesauros, glosarios, etc.
- Corpus: de tipo monolingüe, bilingüe, comparable, paralelo, general, especializado, ad hoc, etc.
- Software: de análisis de corpus, de alineación de textos, de descarga de documentos, etc.

Algunos factores determinantes para la elección de los materiales de las unidades han sido la representatividad de la práctica profesional, la utilidad de los recursos para un futuro profesional, la disponibilidad, la facilidad de acceso, la gratuidad, el diseño agradable y, sobre todo, el potencial para el aprendizaje de la traducción. Obviamente, también han influido los motivos pedagógicos y prácticos. El estudio de la ACT nuevamente da soporte a algunas de nuestras elecciones, por ejemplo, en cuanto a las expectativas de la demanda, el informe apunta que “la demanda concede mucho valor al dominio de la temática de la traducción y la adecuación de la terminología al uso en cada sector. [...] Las herramientas tecnológicas son bien valoradas por la demanda” (ACT 2005: 244).

6.6.1.2.2.1. Selección de textos

1) Variedad de textos

Los textos incluidos en las unidades didácticas son de distintas clases: textos originales, paralelos y de apoyo.

a) Por textos originales nos referimos a aquellos que el alumno habrá de traducir total o parcialmente o de donde se han extraído fragmentos para traducir.

b) En cuanto a la definición de texto paralelo, estamos de acuerdo con Sánchez-Gijón (2003) en que la definición de Nord (1991) del término es demasiado restrictiva ya que se refiere sólo a los textos en LM del mismo tema y tipo que el TO. Coincidimos pues con la primera autora en una concepción del texto paralelo más abierta y que incluye “aquell text original o traduït en llengua de partida o d’arribada del qual el traductor extrau informació lingüística o factual que li permet prendre decisions en una nova traducció” (2003: 23).

c) Por textos de apoyo entendemos aquellos que están destinados expresamente a ayudar al aprendizaje como, por ejemplo, un dossier de apoyo o un “tutorial” sobre un programa informático elaborado por el docente.

2) Criterios de selección de los textos

a) Interés

Aclarados los términos principales, junto a los criterios de variedad y progresión que se han seguido para la selección de todos los materiales, en el caso particular de los textos originales que se incluyen en las unidades diseñadas se ha tenido en cuenta también el interés documental. En otras palabras, los textos seleccionados contienen problemas de traducción de diversa índole que requieren estrategias de documentación variadas, razón por la que sirven para ilustrar el manejo de corpus y demuestran su utilidad. En este sentido, son relevantes los diversos tipos de interés de los que habla Hurtado (1995) en relación con la selección de textos:

- 1) Interés lingüístico: textos que contienen problemas de traducción pero no sólo aquellos que pueda solucionar el diccionario.
- 2) Interés extralingüístico: textos que contienen problemas de traducción relacionados con cuestiones conceptuales, conocimientos enciclopédicos o temáticos.
- 3) Interés textual: textos que contienen problemas de traducción relacionados con cuestiones de género textual.

En general, podemos afirmar que hemos intentado cubrir una amplia gama de textos en lo que se refiere a tipos (descriptivos, narrativos, expositivos, argumentativos e instructivos), géneros (noticias de prensa, páginas web personales, artículos científicos, descripciones técnicas, etc.) y campos de especialidad (científico, técnico, económico) con la intención de exponer al alumno a una variedad que refleja, a nuestro parecer, la realidad profesional.

Kelly (2005) habla precisamente de dos tipos de criterios que se complementan a la hora de seleccionar textos adecuados para la formación de traductores, la realidad profesional y la progresión pedagógica.

b) Realidad profesional: autenticidad del material y cuestiones éticas

Respecto a la realidad profesional, Kelly (ídem: 119 y ss) valora positivamente la autenticidad del material, que los textos puedan enmarcarse dentro de una situación de traducción realista y que sirvan para tratar de forma explícita o implícita cuestiones de ética profesional o mercado laboral. En este sentido, debemos decir que, en la medida

de lo posible, los textos utilizados en nuestras unidades didácticas son auténticos ya que algunos se han tomado de proyectos de traducción reales (p.ej. textos en la unidad “El corpus ad hoc como remedio para la traducción médica al inglés”) o pertenecen a un campo con mucha demanda de traducciones. Por ejemplo, en la unidad “Utilidad de un corpus comparable bilingüe para traducir textos técnicos” se trabaja con la traducción técnica, que es la especialidad con más demanda en el mercado de la directa (inglés-español) según el informe de la ACT (2005: 109), lo que añade un punto de motivación entre los estudiantes, que ven su trabajo más cercano a dicha realidad profesional. Además, los criterios que han llevado a seleccionar los textos paralelos tanto en inglés como en español para construir el corpus comparable bilingüe de esta unidad de traducción técnica están condicionados en parte por la disponibilidad. Las descripciones técnicas de microondas, que es el género tratado, pueden encontrarse sin demasiada dificultad en Internet en diversas lenguas; asimismo, la recopilación, análisis y caracterización del género se ven favorecidos por el reducido tamaño de los textos.

Aparte de la autenticidad y el realismo, las cuestiones éticas y de situación profesional se tratan especialmente en la unidad “Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”, donde se muestran ejemplos de malas prácticas en este ámbito.

c) Progresión pedagógica

En relación con la progresión pedagógica a la hora de seleccionar textos, Kelly (ídem: 122 y ss) tiene en consideración diversos factores:

- Tipos de texto: la autora considera que es importante exponer al alumno a los tipos de texto y géneros más relevantes en el mercado laboral, como así hemos hecho y apuntado anteriormente, para ir trabajándolos empezando por los más estandarizados y avanzando hacia los que están sujetos a menos convenciones, como los textos creativos.
- Textos prototípicos: se destaca la importancia de exponer a los alumnos a discursos prototípicos para que afloren las distintas perspectivas culturales de los estudiantes.
- Accesibilidad a los contenidos: se recomienda, entre otras cosas, empezar por textos a los que probablemente los estudiantes hayan estado expuestos.

- Accesibilidad a documentación fiable: el fácil acceso a contenidos relacionados con los textos y a documentación de calidad ha sido un criterio fundamental en la selección del material de nuestras unidades didácticas.
- Interés y motivación del estudiante, y viabilidad de la actividad: la autora recomienda no perder de vista la intención de motivar al alumnado planteándole textos de dificultad asequible que no le hagan perder el interés.

6.6.1.2.2.2. Selección de recursos lexicográficos

La disponibilidad de recursos documentales de tipo lexicográfico como diccionarios generales y especializados, monolingües y bilingües, tesauros, glosarios, etc., junto con algunas enciclopedias, buscadores y portales ha sido un elemento clave para su selección, no sólo por el hecho de estar en la Web y ser gratuitos en su mayoría, sino por el hecho de estar en formato electrónico, lo cual facilita su consulta y, si cabe, la recogida de la información o el propio documento para su inclusión en un corpus.

Así, en la unidad “Introducción al uso de corpus electrónicos para traducir” la selección de los recursos utilizados tales como diccionarios y glosarios en línea, buscadores generales y buscadores especiales como *WebCorp* viene motivada por su disponibilidad en línea, además del hecho de que, en algunos casos, los alumnos ya los reconocen como familiares y, en otros casos, representan recursos que, pensamos, les serán de utilidad no sólo en el marco de nuestras unidades didácticas, sino en su trabajo como alumnos primero y como profesionales más adelante. En el caso de la unidad didáctica “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos” los materiales escogidos fueron direcciones de Internet a las que se llegó a través de buscadores y de donde se podían recoger descripciones técnicas de microondas. En otras unidades, como en “Sólo utilizar corpus; utilizar corpus solo”, se ha hecho uso de recursos que permitían el acceso en línea a libros especializados; en la unidad de “Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés” se han introducido metadicionarios.

6.6.1.2.2.3. Elaboración de fichas

En lo que atañe a las fichas que aparecen en todas las unidades didácticas y que se presentarán en el siguiente capítulo, la elaboración de todas ellas responde a objetivos concretos que se desea que el alumno alcance. En estas fichas se materializa el trabajo propuesto para algunas tareas; se pretende provocar no sólo el trabajo práctico del

alumno, sino sobre todo su reflexión y que ambos aspectos queden plasmados en estas fichas que se guardan para posteriores revisiones o tareas de reciclaje, evaluación, etc. Algunas de las fichas contienen ejercicios inspirados en la obra de otros autores, como es el caso de la unidad “Contrastes de uso en un corpus paralelo”, en la que algunos ejemplos de contrastividad se han extraído de la obra de López Guix y Minett Wilkinson (1997), en la que se llama la atención sobre diversos puntos de divergencia entre el inglés y el español en cuanto a sintaxis, morfología y puntuación.

6.6.1.2.2.4. Selección de tipos de corpus

Aquí también la variedad y la disponibilidad han sido criterios muy importantes, si bien los principales han sido la gratuidad y la relevancia del corpus para cubrir los contenidos propuestos. Aunque se podrían hacer diversas clasificaciones, agrupamos aquí los corpus en dos categorías amplias.

1) Corpus generales

Para la unidad “Introducción al uso de corpus electrónicos para traducir” se seleccionaron corpus generales tanto en inglés (*British National Corpus -BNC-*, *Bank of English -BoE-*) como en español (*Corpus de Referencia del Español Actual -CREA-*, de la *RAE*), cuya disponibilidad en línea, así como la calidad que se les presupone y la cantidad de material que incluyen, fueron algunos de los criterios principales que llevaron a su elección, si bien es cierto que también se tuvieron en cuenta criterios pedagógicos. Por ejemplo, tanto el *BoE* como el *CREA* presentan sus resultados en formato *KWIC*, un formato de visualización de concordancias al que el alumno ha de habituarse y saber explotar, lo que coincide con uno de los objetivos específicos previstos. Por otro lado, el *BNC* y el *CREA* proporcionan información documental a partir de las concordancias, aspecto que también coincide con el objetivo de aprendizaje relacionado con la importancia de saber evaluar las fuentes de los documentos que forman un corpus.

El trabajo de contrastividad en la unidad “Contrastes de uso en un corpus paralelo” está centrado no tanto o no solamente en mejorar el conocimiento de cómo funciona cada una de las lenguas implicadas, sino más bien en ver dicho funcionamiento en contraste, comparando usos. Para tal fin, el corpus seleccionado en este caso fue un corpus paralelo de los mismos discursos de la UE en inglés y en español sobre diversos temas. Aquí el énfasis no está en el tema o el género de los textos que componen el corpus ya

que el trabajo en profundidad con textos pertenecería, como ya se ha indicado, a la traducción especializada. Baste decir que el nivel de lengua y el contexto en el que se utiliza están, en general, al alcance de unos alumnos de iniciación. En este caso, insistimos, el objetivo está en saber utilizar un corpus paralelo para poder analizar contrastivamente, extraer conclusiones y aplicar lo observado, siendo esto un objetivo de aprendizaje de iniciación al uso de corpus para traducir.

Otro ejemplo de corpus de contenido general es uno de los utilizados en “Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”. Éste es el bilingüe *COMPARA* que, aunque sólo trabaja con portugués e inglés, es un perfecto ejemplo de corpus paralelo bidireccional cuidadosamente construido que además es gratis.

2) Corpus especializados

Para la unidad de “Introducción al uso de corpus para traducir” también se seleccionaron corpus especializados, entendiendo que el trabajo con éstos no estaba centrado en análisis complejos de género por pertenecer esto a un estadio más avanzado. Así, se trabajó con una selección de unos 10 millones de palabras del diario *Financial Times* y otros 10 millones del *Expansión* extraídos del CD-ROM *Multilingual Corpora for Co-operation - MLCC*. Las ventajas que ofrece este tipo de corpus son variadas desde un punto de vista práctico ya que se dispone de los textos en formato electrónico, pertenecen a un género concreto, en este caso prensa económica, y cubren un periodo de tiempo breve pero sin interrupciones, con la única desventaja de que en muchas ocasiones no se tiene la certeza de que los textos sean originales (especialmente en otras lenguas que no sean el inglés) ya que las agencias internacionales de noticias suelen escribir originalmente en inglés y pasar su información a otras agencias nacionales que la traducirán. Desde el punto de vista pedagógico, la selección de corpus como el del *Financial Times* o el *Expansión* pueden ofrecer al alumno la oportunidad de enfrentarse a un corpus genérico con unas mínimas garantías de calidad y cantidad, punto éste último difícil de conseguir por sí mismo con la construcción de un corpus ad hoc.

Entre los materiales seleccionados para la unidad “Utilizar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés” está un corpus en línea de correspondencia comercial en inglés original. El corpus se ha seleccionado en base a su disponibilidad en línea, la facilidad de consulta y la variedad de propósitos comerciales

que incluyen las cartas que lo forman lo que, todo junto, permite una gran diversidad de estudios, por ejemplo, sobre un subgénero en particular y las estructuras que ahí se revelan, estudios de tipo cultural o retórico-contrastivo.

Como último ejemplo, en la unidad “Sólo utilizar corpus; utilizar corpus solo” se utilizaron libros en línea a modo de corpus especializados ya que conocer y manejar este tipo de recursos forma parte de los contenidos del uso de corpus para traducir.

6.6.1.2.2.5. Selección de software

En lo que respecta a *WordSmith Tools*, el programa de análisis de corpus utilizado para la práctica totalidad de las unidades didácticas diseñadas, los criterios seguidos para su elección se fundamentan no sólo en su disponibilidad, sino también en cuestiones como la facilidad de uso, un diseño agradable, la inclusión de prácticamente todas las funciones que un alumno de traducción puede necesitar de un programa de análisis de corpus y, finalmente, el hecho práctico de que *WordSmith Tools* es uno de los programas más utilizados en contextos académicos relacionados con el trabajo o enseñanza de metodología de corpus.

Por otra parte, para la consecución de los objetivos de aprendizaje relacionados con el uso de corpus paralelos, también se utilizó el programa informático *MultiConcord*, que permite la alineación de textos, facilita el análisis contrastivo y la generación de ejercicios de contrastividad. *MultiConcord* entra en una categoría de programas cuyo manejo representa un objetivo de aprendizaje de iniciación del alumno de traducción que trabaja con corpus. Objetivo también relacionado con los corpus paralelos, pero ya en la fase de consolidación, es el de manejar éstos como si fueran memorias de traducción, lo cual se trata en otra unidad didáctica, esta vez de inversa, en la que se utilizaron los programas *WinAlign* y *Translator's Workbench* de Trados.

Para la creación de corpus ad hoc, trabajo que pertenece también a la fase de consolidación en el uso de corpus, se utilizaron programas que facilitan la descarga de textos desde Internet (o lo que se conoce como un navegador *offline*). En nuestro caso se utilizaron el *WebCopier* y *Flashgot* por su gratuidad, poco peso en bytes una vez instalado y sencillez a la hora de aprender su manejo. Con el fin de aprovechar al máximo la Web como corpus, también se hizo uso del *KwicFinder*, programa gratuito que permite extraer concordancias de Internet introduciendo varias palabras o frases

clave, observar el contexto y el origen de los documentos recuperados, es decir, contenidos relacionados con el uso de corpus para traducir.

6.6.1.2.2.6. Ejemplo de selección de materiales de una unidad

A continuación sigue un ejemplo completo aunque esquemático de todos los materiales utilizados en la unidad didáctica “El corpus ad hoc como remedio para la traducción médica al inglés” junto con el detalle de los criterios de selección de dichos materiales.

Unidad didáctica: “El corpus ad hoc como remedio para la traducción médica al inglés”		
TIPOS DE MATERIALES	REFERENCIAS DE LOS MATERIALES CONCRETOS	CRITERIOS DE SELECCIÓN O ELABORACIÓN
Textos	TO; (fragmento de) texto especializado de un campo de la medicina en español (1: “Perfil biofísico fetal”)	<ul style="list-style-type: none"> - Representativo de un género típico de traducción al inglés (artículo científico) - Ilustra problemas variados de traducción inversa de un texto especializado: contiene acrónimos y términos específicos del tema y expresiones comunes en el género - Tema ajeno al alumno - Extensión manejable (250-275 palabras) - Facilidad de encontrar TPs de fuentes fiables
	TPs; (fragmentos de) textos especializados pertenecientes a la misma publicación del TO al que sirven de apoyo (2: “Estudio biométrico” y 3: “Resultados clínicos”)	<ul style="list-style-type: none"> - Aportan datos bibliográficos que no contiene el TO
	TP (4: “Exploración ecográfica”)	<ul style="list-style-type: none"> - Reúne varios problemas de traducción del mismo tipo que los del TO
Fichas	Ficha 1: “Análisis del TO”	<ul style="list-style-type: none"> - Permite recoger las características principales del TO - Fomenta el acercamiento al TO antes de traducirlo - Concisión
	Ficha 2: “Identificación de problemas potenciales y análisis de la necesidad de recursos”	<ul style="list-style-type: none"> - Permite recoger los problemas principales de la traducción del TO y reflexionar sobre el tipo de recurso documental que los solucionaría - Fomenta el acercamiento al TO antes de traducirlo

	Ficha 3: “Información sobre el tema extraída del TO y otras fuentes”	<ul style="list-style-type: none"> - Fomenta la extracción de información del TO - Fija la idea de que el TO es la primera fuente de información de cara a una traducción
	Ficha 4: “Lo que sí / no ha conseguido solucionar cada tipo de recurso”	<ul style="list-style-type: none"> - Ayuda a reflexionar sobre el tipo de información que ofrece cada tipo de recurso y fijar la idea de que hay ciertos problemas que sólo un corpus puede solucionar
	Ficha 5: “Ficha de búsqueda y recogida de TPs en inglés”	<ul style="list-style-type: none"> - Permite recoger datos sobre los textos que compondrán el corpus - Ayuda a fijar la importancia de la calidad y la cantidad de los textos que componen un corpus
	Ficha 6: “Ejemplos de cómo utilizar un corpus de TPs para resolver problemas de género, colocaciones y uso”	<ul style="list-style-type: none"> - Ayuda a trabajar la explotación de corpus
	Ficha de autoevaluación de la unidad didáctica	<ul style="list-style-type: none"> - Permite recoger información sobre el proceso y eficacia del aprendizaje desde el punto de vista del alumno - Clave para la evaluación de la UD por parte del docente
	Ficha de valoración de la unidad didáctica	<ul style="list-style-type: none"> - Permite recoger el <i>feedback</i> razonado de los alumnos sobre la unidad didáctica - Clave para la evaluación de la UD por parte del docente
Corpus	Corpus monolingüe general: <i>BNC</i> http://thetis.bl.uk/lookup.html	<ul style="list-style-type: none"> - Representativo del inglés británico general moderno - Accesible en Internet - Fácil de manejar - Ofrece información de las fuentes de las concordancias

	Corpus monolingüe general: <i>CREA</i> http://www.rae.es	<ul style="list-style-type: none"> - Representativo del español general moderno - Accesible en Internet - Fácil de manejar - Ofrece información de las fuentes de las concordancias
Otros recursos documentales	Diccionario bilingüe general: http://wordreference.com	<ul style="list-style-type: none"> - Accesible en Internet - Conocido y utilizado en exceso por el alumnado
	Buscadores generales: http://www.google.com , http://www.alltheweb.com	<ul style="list-style-type: none"> - Accesibles en Internet - Ampliamente utilizados ya por los alumnos aunque no óptimamente
	Enciclopedias en línea: http://www.encyclopedias.com (varios temas), http://www.umm.edu/ency/index.html (<i>Medical Encyclopedia</i>), http://es.encarta.msn.com/ (<i>Encarta</i>)	<ul style="list-style-type: none"> - Accesibles en Internet - Fáciles de utilizar (sintaxis de búsqueda común a otros recursos ya conocidos)
	Buscador de buscadores especializados: http://www.buscopio.net	<ul style="list-style-type: none"> - Accesible en Internet - Fácil de utilizar - Novedoso para los alumnos
	Diccionarios / glosarios multilingües / bilingües especializados: http://allserv.rug.ac.be/~rvdstich/eugloss/welcome.html (<i>European Medical Glossary</i>), http://iate.europa.eu (<i>IATE</i>), http://www.termcat.net (<i>Termcat</i>)	<ul style="list-style-type: none"> - Accesibles en Internet - Fáciles de utilizar - Ofrecen más información aparte del equivalente terminológico (campo temático, registro)

	<p>Diccionarios / glosarios monolingües especializados: http://www.onelook.com</p>	<ul style="list-style-type: none"> - Accesible en Internet - Fácil de utilizar - Ofrece múltiples datos (acepciones, campos de especialidad, etimología) y nuevas formas de encontrar información (términos compuestos, palabras relacionadas)
	<p>Tesauros monolingües especializados: http://www.medical-dictionary.com</p>	<ul style="list-style-type: none"> - Accesible en Internet - Fácil de utilizar - Novedoso por la forma de aprovechar las búsquedas
	<p>Buscador de acrónimos: http://www.acronymfinder.com</p>	<ul style="list-style-type: none"> - Accesible en Internet - Fácil de aprender a utilizar - Novedoso por las posibilidades de búsqueda
	<p>Diccionarios monolingües / bilingües / multilingües generales: http://wordreference.com, http://www.onelook.com, http://www.foreignword.com, http://www.diccionarios.com (Diccionarios <i>Vox</i> en Internet)</p>	<ul style="list-style-type: none"> - Accesibles en Internet - Mecanismo de búsqueda conocido - Ampliamente utilizados (algunos) por los alumnos aunque no de forma óptima
	<p>Diccionarios / tesauros monolingües generales: http://www.wordsmyth.net, http://thesaurus.reference.com/, http://dictionary.reference.com, http://www.hyperdictionary.com</p>	<ul style="list-style-type: none"> - Accesibles en Internet - Mecanismo de búsqueda conocido - Novedoso por la información adicional que ofrecen

	Buscadores de TPs de tema científico: http://www4.ncbi.nlm.nih.gov/entrez/query.fcgi (<i>PubMed</i>), http://www.scirus.com (<i>Scirus</i>)	<ul style="list-style-type: none"> - Accesibles en Internet - Mecanismo de búsqueda conocido o fácil de aprender - Ofrecen fiabilidad en cuanto a la calidad de los textos que contienen, aunque no infalibilidad - Ofrecen una vasta cantidad de textos - Permiten hacer búsquedas restringidas por autor, revista, palabra clave, medio, etc. - <i>PubMed</i> permite guardar cientos o miles de textos en segundos
Software	<i>WordSmith Tools</i>	<ul style="list-style-type: none"> - Disponible en la FTI, UAB - Entre los más utilizados en el trabajo con corpus a nivel de usuario - Ya conocido por los alumnos de este nivel - Diseño “amable”

Tabla 15. Ejemplo de selección de materiales de la unidad didáctica “El corpus ad hoc como remedio para la traducción médica al inglés”.

6.7. Evaluación

Tema espinoso donde los haya dentro de la traducción y, por ende, poco investigado y sistematizado en nuestra disciplina, tratar de la evaluación exige plantearse varias cuestiones principales como son el concepto de evaluación del que se parte (qué es), el objeto (qué se evalúa) y la manera de llevarla a cabo (cómo se evalúa). Estas cuestiones, que se funden con otras como el propósito y el tipo de función que se escogen para la evaluación, constituyen el marco dentro del que también ha de definirse lo que se entiende por problema y error de traducción, o, en nuestro caso, error en el uso de corpus para traducir, han de establecerse criterios claros de evaluación y han de diseñarse o explicitarse los instrumentos utilizados para su realización. Además, el diseño claro de los objetivos de aprendizaje, expresados en términos de competencias, y éstas desgranadas en indicadores, es fundamental para la evaluación.

Respecto a lo que entendemos por problema de traducción, podemos tomar la definición de Nord (1995: 151), según la cual es un problema objetivo que todo traductor [...] debe resolver durante una tarea de traducción determinada. Sin embargo, a tal definición habríamos de añadirle el hecho particular de que, en el marco de este trabajo, todos los problemas de traducción seleccionados tienen solución a través del uso de corpus con el fin de que tal aspecto pueda ser evaluado. Por otro lado, la clasificación de problemas adoptada ha sido la que distingue entre problemas lingüísticos, extralingüísticos, textuales y pragmáticos (Hurtado 1996b). Muy ligado a la noción de problema de traducción está el error de traducción que, en este contexto, puede hallarse tanto en el producto como en el proceso. Los errores en el producto han sido evaluados con la ayuda de los baremos propuestos por Hurtado (1999). En el caso del proceso de uso de corpus para traducir, ha sido necesario desarrollar unos indicadores relacionados directamente con los objetivos de aprendizaje y susceptibles de ser evaluados, una ficha de recogida de datos sobre el uso de corpus para traducir y otra de cruce de datos.

Para poder llevar a cabo tal evaluación del uso de corpus en la traducción, en nuestro caso la observación se limitó por motivos de manejabilidad de datos a cinco problemas de traducción previamente seleccionados por el docente dentro de cada texto original.

6.7.1. La evaluación como fuente de información

La evaluación en traducción, que durante siglos se ha llevado a cabo de forma subjetiva y basándose principalmente en cuestiones estilísticas y de método utilizado (literal o libre), avanza ya desde algún tiempo, aunque a paso lento, hacia una definición de criterios más objetivos y el diseño de instrumentos que ayuden a su sistematización. Martínez Melis y Hurtado (2001) hacen un repaso de las áreas de evaluación que son típicas en traducción, la investigación en evaluación en los Estudios sobre la Traducción, las nociones y modelos de evaluación en traducción, los objetos, tipos, funciones, medios y objetivos de la evaluación, para concluir apuntando la dirección en la que debería caminar la investigación, sobre todo empírica, en evaluación en la enseñanza de la traducción dadas sus carencias.

Dejando atrás la identificación tradicional y casi exclusiva de la evaluación en traducción con la crítica y corrección de traducciones, partimos de un concepto de la evaluación en traducción cuyo objeto de estudio no es sólo el producto final de la traducción, sino también el proceso para llegar a ese producto y el individuo que lo lleva a cabo. En esta misma línea, nos planteamos la evaluación no sólo como un mero proceso de medición, sino como una fuente de información que nos aporte datos para mejorar la enseñanza, es decir, optamos por una evaluación de tipo formativo. Según Abrecht (1991), este tipo de evaluación, que permite extraer datos tanto de los resultados como del proceso de trabajo del estudiante, se caracteriza sobre todo porque:

1. Está dirigida principalmente al estudiante que, al hacerse más consciente, se implica de forma más activa en su propio proceso de aprendizaje.
2. Está totalmente integrada en el proceso de aprendizaje.
3. Es flexible y plural ya que ha de poder adaptarse a situaciones individuales.
4. Se ocupa tanto de procesos como de resultados.
5. Los datos que recoge llevan a la retrospectiva y a la búsqueda de soluciones a problemas a través de su observación y comprensión.
6. Está dirigida también al profesor ya que le permite mejorar y flexibilizar su enseñanza.

Características de la evaluación formativa (adaptado de Abrecht 1991).

De esta lista de características, que a su vez entendemos como ventajas de la evaluación formativa, queremos incidir en el punto 2 ya que, si recordamos lo mencionado en el

capítulo precedente, la posibilidad de integrar la evaluación dentro de la unidad didáctica junto a los objetivos y competencias, los contenidos y los medios es también una de las ventajas de la adopción del enfoque por tareas de traducción como marco metodológico. De hecho, según Hurtado (1999), en la didáctica de la traducción se vuelve indispensable el que los criterios de evaluación y los objetivos de aprendizaje estén directamente relacionados.

Por otra parte, en las publicaciones más recientes de Hurtado (2007, 2008), en las que muestra una evolución hacia la formación por competencias, la autora defiende que son los indicadores de las competencias, es decir, los elementos observables, los que sirven de base para la evaluación, lo cual no está reñido con sus publicaciones anteriores, sino que supone una reconceptualización y un avance a partir de las mismas.

Así, junto a los objetivos de aprendizaje y las competencias se encuentran algunos principios básicos que deben observarse a la hora de establecer los criterios de evaluación, como son el qué se evalúa, en sus vertientes de qué se debe y qué se puede evaluar, por qué se evalúa, para qué y para quién. En nuestro caso particular, la evaluación se ha centrado en la consecución de los objetivos (tanto relacionados con el uso de corpus como con las materias de traducción) seleccionados para cada tarea y unidad didáctica considerando que es necesario tanto para el evaluador como para el evaluado hacer un seguimiento del proceso de enseñanza-aprendizaje con el fin de obtener información que ayude a mejorar tal proceso. Para llevar a cabo tal evaluación se han utilizado diversos instrumentos entre los cuales se encuentran cuestionarios, fichas de recogida de datos y otros recursos encaminados a realizar el seguimiento del alumno.

En primer lugar, es menester distinguir en nuestro caso entre, por un lado, la evaluación de la adquisición de objetivos relacionados con la explotación de corpus para resolver problemas de traducción y, por otro lado, la evaluación de traducciones. En este sentido, podemos hablar de una evaluación tanto del proceso, cuando se evalúa el uso de corpus para una traducción, como del producto, cuando se evalúa la calidad de la traducción resultante de la ejecución de alguna tarea. Siendo el uso de corpus para la enseñanza de la traducción el tema principal del presente trabajo, al hablar de evaluación nos centraremos en dicho aspecto, y no en la calidad final de las traducciones, si bien es cierto que ésta sí se ha tenido en cuenta como variable de control. Aunque no se explique aquí en detalle, baste decir que la evaluación de

traducciones se ha realizado con la ayuda del baremo de corrección y las indicaciones para la notación incluidas en Hurtado (1999: 120, 181).

6.7.2. Indicadores

Los indicadores son comportamientos observables que dan la clave de la evaluación de las competencias, o de los objetivos expresados en términos de competencias. Indican resultados de aprendizaje que el estudiante debe haber asimilado, y nosotros hemos optado por expresarlos en 3ª persona.

La lista que se presenta a continuación es una relación de indicadores susceptibles de ser tenidos en cuenta para la evaluación del uso de corpus para la traducción. En otras palabras, utilizar corpus para traducir implica realizar operaciones relacionadas al menos con varios de estos indicadores; dependiendo del tipo de tarea descrita y los objetivos que se quieran evaluar, se seleccionarán unos u otros indicadores para ser considerados dentro de la Ficha de recogida de datos sobre el uso de corpus y la Ficha de cruce de datos que se presentarán más adelante (cf. infra 6.7.4.1.3. y 6.7.4.1.4.). Es necesario indicar que éstos son los indicadores utilizados para las unidades didácticas propuestas y que el orden de indicadores aquí expuesto es aleatorio. Obviamente, en un diseño curricular más amplio habría que añadir más indicadores.

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación, ...

- 1: selecciona el corpus idóneo
- 2: maneja distintos (tipos de) corpus que están disponibles en línea
- 3: realiza búsquedas en archivos PDF como si fueran un corpus
- 4: realiza una búsqueda de calidad (palabra/cadena de palabras idónea; búsqueda truncada, búsqueda con palabra de contexto)
- 5: reordena los resultados de unas concordancias
- 6: usa las funciones idóneas del programa informático
- 7: selecciona el programa informático idóneo
- 8: compila un corpus ad hoc de calidad
- 9: maneja software de conversión de archivos PDF
- 10: ubica y descarga textos electrónicos desde Internet
- 11: maneja software de alineación de textos
- 12: convierte un corpus paralelo en una memoria de traducción

13: extrae concordancias bilingües en un programa de traducción asistida para utilizar una memoria de traducción a modo de corpus
14: busca equivalentes a elementos de discrepancia entre el inglés y el español
15: interpreta las frecuencias de aparición
16: resuelve la traducción de elementos de la lengua general y especializada
17: encuentra equivalentes terminológicos/colocaciones/fraseología en textos científico-técnicos
18: justifica correctamente su opción de traducción basándose en datos extraídos del corpus
19: emplea un tiempo razonable en resolver un problema
- Aceptabilidad del equivalente propuesto

Tabla 16. Indicadores susceptibles de ser evaluados respecto al uso de corpus para traducir.

6.7.3. Procedimiento de evaluación del uso de corpus cuando se traduce un texto

Aunque, como indica Hurtado (1999), no es obligatorio que cada tarea que compone una unidad didáctica contemple una evaluación, la mayoría de las tareas aquí diseñadas incluyen algún tipo de evaluación o autoevaluación para ser llevada a cabo bien con el profesor, bien entre los alumnos. Como se verá en el siguiente capítulo, donde se explicita cada una de las tareas, algunas de estas actividades de evaluación implican la puesta en común de resultados para propiciar una discusión en voz alta, otras hacen que los alumnos observen los resultados de otros y se evalúen entre sí, otras posibilitan la autoevaluación, etc. En el caso de la evaluación de las tareas finales, en todas las unidades didácticas, excepto en las dedicadas a cuestiones contrastivas, de concienciación de la traducción inversa y de uso de una memoria de traducción, se ha incluido la realización de una traducción con la ayuda de corpus electrónicos con el fin de evaluar su uso en relación con los objetivos de aprendizaje marcados.

Se presenta a continuación el procedimiento estándar del docente para preparar la evaluación del uso que hacen los estudiantes de un corpus a la hora de traducir un texto, por ejemplo, en una tarea final. Este hecho, no obstante, no es siempre así, ya que se dan unidades en las que no hay tarea final, o no incluyen realizar una traducción con corpus como actividad para evaluar.

1. Se seleccionan los objetivos de aprendizaje que se quieren alcanzar.
2. Se seleccionan los indicadores que se quieren evaluar (los que se relacionan con los objetivos de aprendizaje).
3. Se asigna un valor que es un porcentaje de la nota global a cada indicador en función del nivel que se espera del alumno.
4. Se selecciona el TO que el alumno va a traducir con corpus.
5. Se selecciona un número limitado de problemas de traducción (cinco, por ejemplo) del TO que se puedan solucionar utilizando corpus.
6. Se prepara el cuestionario de acuerdo con los problemas de traducción seleccionados.
7. Se facilita a los estudiantes el TO y el encargo, uno o varios corpus que pueden venir ya predeterminados por el docente y el cuestionario con los problemas seleccionados.

6.7.4. Herramientas de evaluación

Como bien indican Martínez Melis y Hurtado (2001), la traducción no sólo puede evaluarse a través de la realización de traducciones, sino que existen otros instrumentos que pueden aportar otro tipo de datos. En el caso del uso de corpus en traducción, también hemos aplicado este criterio e intentado obtener información por diversos medios ya que creemos que tanto el producto como los procesos pueden ser evaluados en la traducción, sólo aplicando diferentes instrumentos. En este caso, las herramientas que hemos desarrollado para evaluar el uso de corpus para resolver problemas de traducción son los siguientes:

- Evaluación del estudiante
 - o Diario de aprendizaje
 - o Ficha de autoevaluación
 - o Ficha de recogida de datos sobre el uso de corpus
 - o Ficha de cruce de datos
- Evaluación de la enseñanza
 - o Ficha de registro de observación
 - o Ficha de valoración de la unidad didáctica

La Ficha de registro de observación y el Diario de aprendizaje son herramientas más generales para el seguimiento de las clases y el alumno durante un periodo que puede extenderse más allá de la unidad didáctica. Las restantes herramientas son más concretas y se modifican para cada unidad didáctica o tarea que se quiera evaluar.

Aunque presentamos todas las herramientas, la Ficha de recogida de datos sobre el uso de corpus al traducir y la Ficha de cruce de datos son los principales dentro de nuestra propuesta de evaluación del uso de corpus para traducir.

6.7.4.1. Herramientas de evaluación del aprendizaje

6.7.4.1.1. Diario de aprendizaje

Con el fin de obtener la apreciación del alumno sobre la marcha del proceso de enseñanza-aprendizaje, se diseña un “diario de aprendizaje” común a todas las unidades didácticas. Este recurso se utiliza con la intención de que el alumno recoja con sus propias palabras, incluso en un tono informal y totalmente subjetivo, reflexiones sobre su propia adquisición de contenidos y habilidades. A continuación se muestran, primero, el modelo de diario de aprendizaje y, segundo, un ejemplo real en el que se ve un formato ligeramente diferente ya que se dio total libertad a los alumnos para que redactaran y organizaran dicho diario; se incluyen también subrayados los comentarios introducidos por la docente:

HOY HE APRENDIDO...	
dd/mm/aa (fecha)	
Aspectos sobre el funcionamiento de la lengua en general	
Cuestiones lingüísticas sobre el inglés	
Cuestiones extralingüísticas sobre el inglés	
Cuestiones textuales sobre el inglés	
Cuestiones pragmáticas sobre el inglés	
Cuestiones lingüísticas sobre el español	
Cuestiones extralingüísticas sobre el español	
Cuestiones textuales sobre el español	
Cuestiones pragmáticas sobre el español	
Funciones del WordSmith Tools u otros programas	
Otros (sobre la lingüística de corpus, etc.)	

Documento 4. Diario de aprendizaje del estudiante.

HOY HE APRENDIDO...

Aspectos sobre el funcionamiento de la lengua en general
<p>-Que los términos especializados de las lenguas se han de estudiar minuciosamente antes de poder dar equivalencias de traducción porque el significado de cada uno puede variar según el campo temático en el que se encuentre.</p> <p>-Que los términos suelen asociarse con otras palabras para formar colocaciones o patrones, de diferente modo según cada lengua, que caracterizan a éstos términos y dan naturalidad a las equivalencias de traducción.</p> <p>-Que la frecuencia de uso determina, en gran medida, la mejor equivalencia de traducción. <i>(Dices bien "en gran medida" porque sabemos que el hecho de que un equivalente sea más frecuente no significa que sea el más adecuado: siempre dependerá del contexto.)</i></p> <p>-Que los términos varían en el tiempo y en el espacio y, por lo tanto, es importante situar el origen de los textos con los que se trabaja.</p> <p>-Que en los lenguajes especializados, los términos suelen utilizarse de un modo concreto y, aunque en la mayoría de los casos otras equivalencias serían correctas, es conveniente utilizar una equivalencia apropiada a la lengua de llegada, para conseguir naturalidad.</p>
Cuestiones lingüísticas sobre el inglés
<p>-Que <i>bourse</i> y <i>stock exchange</i> son sinónimos y equivalen, en castellano, a <i>bolsa</i>. <i>("Stock exchange" y "bourse", aunque significan lo mismo, no son sinónimos al 100%, entre otras cosas, por la economía de la lengua: sería poco lógico tener varias palabras totalmente intercambiables en todos los contextos para designar un mismo concepto. No son 100% sinónimos porque tienen usos diferentes: frecuencia, colocaciones, dan lugar a expresiones distintas ("bourse" es más propenso a funcionar de premodificador), distinta capacidad de generar acrónimos ("Stock exchange" sí los genera, el otro no), aunque en algunos casos se pueden usar en la misma frase con el fin de evitar una repetición).</i></p>
Cuestiones extralingüísticas sobre el inglés
-Que hay bases de datos y corpus de textos en Internet que ayudan mucho en la traducción.
Cuestiones lingüísticas sobre el español
-Que tanto <i>pero</i> como <i>y</i> se pueden utilizar en castellano a principio de frase.
Cuestiones extralingüísticas sobre el español
-Que hay bases de datos y corpus de textos en Internet que ayudan mucho en la traducción.
Funciones del WordSmith Tools
Las de la ficha 2.
Otros (sobre la lingüística de corpus, etc.)
<p>Que el uso de corpus de textos es muy útil para el traductor, tanto en la búsqueda de equivalencias como en la recogida de información sobre un tema especializado, pero sobretodo es útil para confirmar rápidamente equivalencias.</p> <p>-Que dependiendo de cómo y dónde se busca un término, los resultados son más o menos exactos y se encuentran más o menos rápidamente.</p> <p>-Que con el uso de corpus de textos es fácil determinar si un término es actual o si está en desuso.</p> <p>-Que hay grandes diferencias entre el uso de un corpus general, con muchos campos semánticos y varias posibilidades para cada término, y el uso de un corpus especializado que se centra en un campo semántico y recoge sólo información sobre ese campo.</p> <p>-Que el trabajo con corpus de texto y programas informáticos tipo el <i>WordSmith Tools</i> agiliza mucho el trabajo del traductor, ya que el programa hace búsqueda y clasificaciones que una persona tardaría horas en hacer, en pocos segundos.</p> <p>-Que tener información sobre la frecuencia de uso de un término y obtener ejemplos contextualizados permite avanzar mucho en la búsqueda de equivalencias. <i>(Muy buen ejercicio de síntesis)</i></p>

Documento 5. Diario de aprendizaje cumplimentado por un alumno.

6.7.4.1.2. Ficha de autoevaluación

Retomando el tema de la utilización de diversas herramientas para la evaluación y el seguimiento, presentamos una ficha llamada de Autoevaluación que se facilita al alumno al finalizar cada unidad didáctica. Esta ficha contiene preguntas diversas sobre la unidad terminada que pretenden recoger, por un lado, una autoevaluación del alumno sobre lo que cree que ha aprendido y en qué grado ya que se le pide dar una puntuación numérica. Siguiendo este mismo método, en otro bloque se hacen preguntas relacionadas con la percepción del estudiante sobre la utilidad de los corpus. Incluimos aquí un ejemplo real de ficha de autoevaluación aunque, obviamente, las preguntas de cada ficha se adaptan a las características de la unidad didáctica.

FICHA DE AUTOEVALUACIÓN SOBRE LA UNIDAD DE INTRODUCCIÓN A LA METODOLOGÍA DE CORPUS APLICADA A LA TRADUCCIÓN GENERAL (DIRECTA) (Valora de 0 a 10)

Soy capaz de captar...

- que los resultados de una búsqueda en un corpus están directamente relacionados con (el número y la calidad de) los textos que contiene el propio corpus
7
- que hay búsquedas que requieren del uso de un corpus grande, pero que más grande no significa necesariamente mejor
8
- algunas de las ventajas que aporta el uso de un programa informático como el WordSmith Tools a la hora de hacer cálculos sobre grandes cantidades de texto
6
- que un corpus de textos electrónico puede ser un buen recurso documental
7
- qué tipo de preguntas le puedo realizar a un corpus y cómo he de formularlas
6
- sacarle partido a un corpus a la hora de utilizarlo como apoyo en la traducción de un texto
7

Me ha ayudado el corpus...	
- a traducir mejor	7
- a encontrar los equivalentes de traducción	4
- a estar más seguro de la solución	9
- a comprender mejor el texto original	8
- a comprender la importancia de las colocaciones	8
- a ser más consciente de la naturaleza de los textos utilizados en la fase de documentación (origen, fecha, original o traducción, etc.)	6
- otros (por favor, especifica al máximo)	7 a solucionar problemas extralingüísticos de la materia especializada que se trata

Documento 6. Ficha de autoevaluación de la unidad “Introducción al uso de corpus electrónicos para traducir” cumplimentada por una alumna.

6.7.4.1.3. Ficha de recogida de datos sobre el uso de corpus al traducir

Aparte de la decisión de escoger algunos problemas de traducción de un TO solucionables con corpus y llegar a un listado de indicadores para la evaluación del uso de corpus en una traducción concreta, también se diseñó un modelo de ficha de recogida de datos sobre el uso de corpus cuando se traduce como la que se muestra a continuación.

FICHA DE RECOGIDA DE DATOS		
"Título del TO"		
PROBLEMA DE TRADUCCION 1 (lingüístico / extralingüístico / textual / pragmático): "fragmento que representa el problema"		
PREGUNTA	RESPUESTA	
¿Qué palabra o cadena de palabras o caracteres has buscado y en qué corpus? *Para el resto de las preguntas y cuando hayas realizado más de una búsqueda, incluye el número al que haces referencia	Búsqueda 1:	Corpus:
	Búsqueda 2:	Corpus:
	Búsqueda 3:	Corpus:
	Búsqueda 4:	Corpus:
	Búsqueda 5:	Corpus:
¿Has limitado la búsqueda de alguna manera? (p.ej. variante lingüística geográfica como el español peninsular o el inglés de UK, modo oral o escrito, fecha de los documentos, palabra de contexto, etc.)	Búsqueda 1: Búsqueda 2:	
¿Has ordenado alfabéticamente (<i>Resort</i>) las concordancias?	Búsqueda 1: Búsqueda 2:	
¿Has ordenado alfabéticamente (<i>Resort</i>) las concordancias hacia la derecha o hacia la izquierda? Anota el criterio de ordenación (p.ej. L1, R1)	Búsqueda 1: Búsqueda 2:	
Anota si has utilizado otras funciones que te permitía el programa WordSmith Tools (p.ej. <i>Grow, Shrink, Clusters, Collocates</i>).		
Anota las diferentes soluciones de traducción que tenías antes de llegar a la definitiva.		
Justifica tu solución definitiva frente a las otras opciones		
SOLUCIÓN PROPUESTA POR EL ALUMNO: "tu solución definitiva"		

Documento 7. Ficha de recogida de datos sobre el uso de corpus cuando se traduce.

Como se puede ver, las preguntas se corresponden con diversos indicadores que interesa evaluar en la traducción para la que esta ficha en concreto ha sido diseñada. Si, por ejemplo, la traducción implicara la compilación de un corpus ad hoc o el uso de un corpus paralelo alineado con un software específico, habría que adaptar varias de las preguntas para poder evaluar otros objetivos de aprendizaje relacionados con ser capaz de construir un corpus ad hoc, manejar un corpus paralelo, etc.

6.7.4.1.4. Ficha de cruce de datos

Con la información obtenida en la ficha anterior podemos pasar a la ficha de cruce de datos también diseñada en el marco de esta tesis y de la cual mostramos un modelo (sin datos concretos) a continuación.

**FICHA DE CRUCE DE DATOS
UNIDAD DIDÁCTICA "..."**

EVALUACIÓN DEL USO DE CORPUS: Actuación ante los corpus sólo cuando se utilicen
Traducción del texto "..."

Legenda:
Problema 1: ... Problema 2: ... Problema 3: ... Problema 4: ... Problema 5: ...
Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación, ...
Indicador 1: ... Indicador 2: ... Indicador 3: ... Indicador 4: ... Indicador 5: ...
0: Mal 1: Mejorable 2: Bien (entre las diversas soluciones que en ocasiones pueden estar bien)
Aceptabilidad (del equivalente propuesto)
√: Correcto (entre las diversas soluciones que en ocasiones pueden estar bien) X: Incorrecto

Alumno/a X	Problema 1	Problema 2	Problema 3	Problema 4	Problema 5	PUNTUACIÓN PARCIAL	PUNTUACIÓN PARCIAL (después de aplicar los %)	PUNTUACIÓN FINAL (sobre 10)
Indicador 1 (20%)	0/1/2	0/1/2	0/1/2	0/1/2	0/1/2	/10		
Indicador 2 (20%)	0/1/2	0/1/2	0/1/2	0/1/2	0/1/2	/10		
Indicador 3 (20%)	0/1/2	0/1/2	0/1/2	0/1/2	0/1/2	/10		
Indicador 4 (20%)	0/1/2	0/1/2	0/1/2	0/1/2	0/1/2	/10		
Indicador 5 (10%)	0/1/2	0/1/2	0/1/2	0/1/2	0/1/2	/10		
Aceptabilidad	√ / X	√ / X	√ / X	√ / X	√ / X			

Documento 8. Ficha de cruce de datos.

Como se puede apreciar, este modelo de ficha de cruce de datos contiene varias secciones que se explican de forma sintética en la leyenda y que ampliamos ahora con más detalle. Más adelante (cf. infra 6.7.4.1.5.) proporcionaremos un ejemplo concreto en el que se recorrerá todo el proceso de evaluación con el fin de facilitar su comprensión. Por el momento, los aspectos destacables del modelo presentado son:

- Selección de varios problemas de traducción variados (p.ej. cinco) seleccionados por el docente y solucionables con la ayuda de los corpus disponibles, como se ha explicado.
- Uso de varios indicadores (el número puede variar) seleccionados en base a los objetivos de aprendizaje que se supone que el estudiante ha adquirido en esa tarea o unidad didáctica o en otras anteriores. A cada elemento se le asigna un valor porcentual que dependerá de la relevancia de tal elemento y de las competencias que se le suponen al estudiante en el uso de corpus. Al final, el porcentaje total ha de ser sobre 100%, independientemente del número de elementos que se hayan seleccionado.
- Un último indicador (en el modelo aparece con el número 6) que no se tiene en cuenta a la hora de la notación, es decir, la asignación de un valor numérico, pero que, en general, sí se considera a la hora de evaluar el uso de corpus para traducir como variable de control, es decir, que representa a un factor distinto e influye en el funcionamiento del sistema.
- Gradación entre 0, 1 y 2 para valorar si lo que el estudiante ha realizado con el corpus respecto a un indicador está mal, es mejorable o es correcto, entre las diversas opciones que puede haber en ocasiones.
- Distinción entre una solución correcta o incorrecta (\checkmark / X) sólo en lo que atañe al último indicador, es decir, la aceptabilidad del resultado final o traducción.
- Puntuación parcial que ha de interpretarse de forma horizontal, es decir, haciendo los cálculos dentro de cada indicador. Por ejemplo, si el estudiante realiza alguna operación para cada problema dentro de un mismo indicador, la puntuación máxima que podría conseguir sería 10/10 (2 puntos en cada uno de los cinco problemas si es que todas las operaciones que ha realizado son correctas), lo que daría un 1 como puntuación parcial; obviamente, la puntuación baja si en alguno de los problemas el alumno realiza alguna operación mejorable o incorrecta, con lo que obtendríamos una ecuación de aciertos de, por ejemplo, 7/10 (0,7 de puntuación parcial). Caso de no realizar ninguna operación para alguno de los problemas, la puntuación máxima se situaría entonces sobre 8, 6, 4 ó 2 y los aciertos, posibilidades de mejora o fallos se calcularían sobre tales cifras, por ejemplo, 5/6 (0,83 de puntuación parcial), es decir, que no se penaliza la no utilización del corpus si el alumno considera que no lo necesita.
- Puntuación total, que es la suma de la puntuación para cada indicador, y que representa una nota sobre 10, una vez aplicados los porcentajes estipulados por el docente para cada indicador.

Con tal modelo, la gravedad del error está en directa relación con el valor porcentual asignado a cada elemento que, como ya se ha explicado, se relaciona a su vez con los conocimientos y habilidades que se supone que el alumno ha adquirido hasta ese momento en su proceso de aprendizaje. La evaluación por indicadores en este modelo permite también observar claramente la etiología del error, es decir, dónde ha fallado el estudiante en el uso de corpus y por qué, lo que habría de llevarnos a una reflexión, análisis y reestructuración de los objetivos de aprendizaje, quizás moviéndonos a crear nuevas tareas para reforzar algunas cuestiones o preparar para otras, por poner un ejemplo.

6.7.4.1.5. Ejemplo

Para entender mejor cómo se ha llevado a cabo la evaluación del uso de corpus para la realización de una traducción ofrecemos un ejemplo real para poder seguir el proceso completo de evaluación, cuyos elementos y pasos principales son:

1. Competencias y objetivos de aprendizaje: Consideración de los objetivos de aprendizaje y competencias, nivel de la unidad didáctica y de los alumnos respecto al uso de corpus y demás aspectos que han integrado la unidad.
2. Indicadores: Selección de los indicadores que se quieren evaluar respecto al uso de corpus.
3. Porcentaje: Asignación de un valor porcentual a cada indicador.
4. Texto original y problemas: Selección del TO y de varios problemas que se van a observar.
5. Encargo de traducción: Presentación al alumno del TO, el encargo de traducción, los corpus disponibles, el software disponible, el tiempo de duración de la tarea/traducción y la ficha de recogida de datos sobre el uso de corpus al traducir. Respecto a la disponibilidad de materiales por parte del alumno, es importante destacar que sólo se permitió el uso de determinados corpus, los seleccionados por el docente, y no se permite el uso de ningún otro recurso con el fin de poder facilitar la evaluación del uso de corpus para traducir y que ésta no se vea alterada o incluso anulada por el uso de otros materiales.
6. Traducción: Traducción del TO por parte del alumno que, a la vez, rellena la ficha de recogida de datos sobre el uso de corpus al traducir.
7. Cruce de datos (entre la ficha de recogida de datos, los indicadores y sus valores porcentuales – interpretación horizontal de la tabla): Observación por parte del profesor de las respuestas del alumno en la ficha de recogida de datos y puntuación de los pasos realizados para cada indicador. Respecto al uso de

corpus, las operaciones reflejadas en las respuestas de los alumnos se consideran como equivocadas, mejorables o correctas (0, 1, 2) en base a criterios de eficacia en el proceso seguido para llegar a una traducción. Es menester insistir en el hecho de que los alumnos tenían la instrucción de que si requerían del uso de algún recurso, éste sólo podía ser alguno de los corpus disponibles para tal traducción, pero que, por otro lado, el uso de corpus no era obligatorio, obviamente, si no lo necesitaban. El hecho de no utilizar corpus podía provocar que la ficha de cruce de datos contuviera algunos espacios en blanco, bien porque el estudiante no había tenido a bien utilizar ningún corpus o no lo había necesitado, o bien porque no había realizado ninguna operación relacionada con ningún indicador

8. Cruce de datos (entre la ficha de recogida de datos y la aceptabilidad de las soluciones de traducción – interpretación vertical de la tabla): Observación del uso de corpus efectuado para cada problema en relación con la solución o equivalente de traducción proporcionado, es decir, una observación entendida de forma vertical dentro de cada problema. Introducción de un valor “o” en aquellas casillas en blanco cuyo resultado (última casilla) sea “incorrecto” (X). Identificación de las casillas con “o” como puntos de atención para la preparación de tareas de refuerzo. Respecto a la calidad final de la traducción en los problemas seleccionados (cinco) ésta se valoró como correcta o incorrecta (√ / X) atendiendo a criterios funcionales, es decir, de adecuación al encargo, destinatario, y otros como el género, registro, etc.
9. Puntuación: Cálculo de la puntuación sobre 10, entendido de forma horizontal dentro de cada elemento, y de forma vertical dentro de la última columna.

El ejemplo real de evaluación que mostramos a continuación pertenece a dos alumnas que realizaron juntas la unidad didáctica de “Dermovet, 1ª puesta en práctica” para llegar a la tarea final de traducción cuya evaluación se llevó a cabo con algunas de las herramientas comentadas. En este caso, una vez tenidos en cuenta los objetivos de aprendizaje señalados para esta unidad didáctica, que se centraban en la selección y evaluación de textos para su explotación en la traducción inversa, y considerado el hecho de que nos encontramos dentro de una materia de traducción especializada en la que se supone que el estudiante ya ha adquirido los principios metodológicos básicos, entre los que se encuentra el uso de corpus a nivel inicial, pasamos a seleccionar los elementos que queríamos evaluar asignándoles ya un valor porcentual en base a su importancia y a las expectativas del profesor respecto a los conocimientos y habilidades adquiridas por el alumno hasta el momento.

Así, se seleccionaron los siguientes indicadores y valores:

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación, ...
1: selecciona el corpus idóneo (corpus disponibles: dermatología veterinaria ad hoc 01-02; dermatología veterinaria ad hoc 02-03; <i>Wordbanks online</i> ; Internet a través de <i>WebCorp</i>): 20%
2: introduce como búsqueda una palabra/cadena de palabras idónea: 20%
3: limita bien la búsqueda (con asterisco, con palabra de contexto): 20%
4: reordena los resultados: 20%
5: usa las funciones del programa informático idóneas: 20%
----- Aceptabilidad del equivalente propuesto

Indicadores seleccionados para la evaluación de la tarea final de la unidad didáctica “Dermovet, 1ª puesta en práctica”.

Una vez decidido qué se quería evaluar, pasamos a seleccionar un texto original y cinco problemas de diversos tipos, todos ellos susceptibles de ser solucionados con la ayuda de los corpus disponibles. A continuación se ofrece el encargo de traducción y una imagen del TO que es parte de una página web especializada y de donde se extrajeron los cinco problemas que siguen a la imagen (cuatro de los problemas están indicados en ésta):

Translate the following text into English. This text is part of a website (<http://www.dermovet.com>) that offers information and services on veterinary dermatology to a non-expert or semi-expert readership. The translated text will have the same type of target readers and function as the source text.



[Terminología Dermatológica Veterinaria](#) / [Centros Veterinarios](#) / [Pruebas diagnósticas](#) / [Novedades Terapéuticas](#) / [Visita Dermatológica](#)

Mascotas...



¿Qué le sucede a mi perro?



¿Qué le sucede a mi gato?

Terminología dermatológica veterinaria

Lesiones Primarias	Lesiones Secundarias	Lesiones Primarias o Secundarias
--------------------	----------------------	----------------------------------

Mácula: Lesión caracterizada por un cambio de color de la piel, sin elevación ni engrosamiento de la misma. Son focales, bien circunscritas y de un tamaño inferior a 1 cm. Existen varios tipos: eritematosa (atopía), hiperpigmentada (lentigo), hipopigmentada (vitíligo), hemorrágica (intoxicación, reacción a fármaco).

Mancha: Lesión idéntica a la mácula pero de un diámetro mayor a 1 cm, y suelen ser menos bien delimitadas.

Pápula: Área cutánea con relieve, sólida y circunscrita de hasta 1 cm de diámetro. Puede ser folicular o interfolicular. Son indicativas de pioderma y/o parasitosis, en la mayoría de casos.

Eritema: Enrojecimiento de la piel debido a la vasodilatación de los vasos dérmicos superficiales. Indica inflamación cutánea.

Placa: Lesión aplanada mayor de 1 cm de diámetro, la mayoría de veces debido a la unión de varias pápulas. Son indicativas de pioderma y menos frecuentemente de enfermedades autoinmunes.

Pústula: Elevaciones bien delimitadas de los estratos superficiales de la epidermis, normalmente con contenido purulento. Hay de varios tipos: sépticas (piodermas), estériles (Pénfigo Foliáceo), eosinofílicas (alergias, parasitosis), linfocíticas (Linfoma Epiteliotrópico). Pueden ser foliculares o interfoliculares.

Vesícula: Lesión similar a la pústula pero con contenido seroso o con exudado inflamatorio y de diámetro inferior a 1 cm de diámetro. Se origina por un acúmulo de fluido en los espacios intercelulares que aparece y desaparece en minutos u horas.

Bulla: Vesícula mayor de 1 cm de diámetro.

Habón: Lesión con relieve consistente en un edema intercelular de las células de la epidermis. Típica de las reacciones de hipersensibilidad tipo I. Es la lesión principal en reacciones de urticaria y en las reacciones positivas al Skin-Test.

Nódulo: Lesión elevada mayor de 1 cm de diámetro, bien delimitada y sólida. Suelen estar bien infiltradas en la dermis. Típico de neoplasias y Paniculitis Nodular Estéril.

Tumor: Masas neoplásicas tanto benignas como malignas. Se utiliza cuando hay nódulos muy grandes.

Quiste: Cavidades forradas por epitelio localizadas en el interior de la piel y normalmente con contenido glandular.

Documento 9. Texto original para traducir en la tarea final de la unidad “Dermovet, 1ª puesta en práctica”.

- 1: *¿Qué le sucede a mi...?* (problema de tipo pragmático)
- 2: *Son indicativas de pioderma... / Es indicativo de Pénfigo Foliáceo...* (problema de tipo lingüístico)
- 3: *Habón* (problema de tipo extralingüístico)
- 4: *Tumor: masas neoplásicas tanto benignas como malignas. Se utiliza cuando hay nódulos muy grandes.* (problema de tipo textual)
- 5: *estrato córneo / basal* (problema de tipo lingüístico)

Problemas seleccionados para la observación y evaluación del uso de corpus para la traducción en la unidad didáctica “Dermovet, 1ª puesta en práctica”.

Para este texto original, una parte de una página web personal sobre dermatología veterinaria escrita en español, se decidió un encargo de traducción al inglés británico con la misma función expositiva e instructiva que el original y el mismo destinatario no especializado o semi-especializado en el tema. Para la realización de la traducción se pusieron a disponibilidad del alumno varios corpus, dos corpus ad hoc monolingües en inglés sobre dermatología veterinaria para ser explotados con el programa *WordSmith Tools*, un corpus monolingüe general en inglés con acceso en línea y todo Internet, considerado como si fuera un macro-corpus para ser accedido a través del recurso *WebCorp*. El uso de los corpus para llegar a la traducción de los cinco problemas seleccionados se anotó en una ficha de recogida de datos como la anteriormente presentada. Aquí mostramos un ejemplo real de una ficha rellenada por dos alumnas en relación al problema 4.

PROBLEMA DE TRADUCCION 4 (textual, de coherencia y cohesión):		
<p>“Tumor: Masas neoplásicas tanto benignas como malignas. Se utiliza cuando hay nódulos muy grandes.”</p>		
PREGUNTA	RESPUESTA	
<p>¿Qué palabra o cadena de palabras o caracteres has buscado y en qué corpus? *Para el resto de las preguntas y cuando hayas realizado más de una búsqueda, incluye el número al que haces referencia</p>	búsqueda 1: tumor	corpus: nuestro corpus y el del año pasado
	búsqueda 2: malignant	corpus: nuestro corpus y el del año pasado
	búsqueda 3: nodule	corpus: nuestro corpus y el del año pasado
	búsqueda 4:	corpus:
<p>¿Has limitado la búsqueda de alguna manera? (p.ej. variante lingüística geográfica como el español peninsular o el inglés de UK, modo oral o escrito, fecha de los documentos, etc.)</p>		
<p>¿Has ordenado alfabéticamente (<i>Resort</i>) las concordancias?</p>		
<p>¿Has ordenado alfabéticamente (<i>Resort</i>) las concordancias hacia la derecha o hacia la izquierda? Anota el criterio de ordenación (p.ej. L1, R1)</p>	Hemos empleado : L1 y R1	
<p>Anota si has utilizado otras funciones que te permitía el programa WordSmith Tools (p.ej. <i>Grow</i>, <i>Shrink</i>, <i>Clusters</i>, <i>Collocates</i>).</p>	Grow	
<p>Anota las diferentes soluciones de traducción que tenías antes de llegar a la definitiva</p>		
<p>Justifica tu solución definitiva frente a las otras opciones</p>	Nos hemos basado en las definiciones de "tumor" que aparecen en el WordSmith Tools y a partir de allí hemos hecho nuestra traducción.	
<p>SOLUCION PROPUESTA POR EL ALUMNO: Tumor: mass of neoplastic origin, benign or malignant. It is used for large nodules</p>		

Documento 10. Ficha de recogida de datos para el problema 4 de la unidad “Dermovet, 1ª puesta en práctica” cumplimentada por dos estudiantes.

Seguidamente, con la información recogida en esta ficha se pudo rellenar la ficha de cruce de datos diseñada para la evaluación del uso de corpus para traducir de las dos alumnas y que se presenta tal y como se entrega a los alumnos para que ellos mismos vean qué áreas deben reforzar o, en otras palabras, qué objetivos de aprendizaje les quedan por alcanzar.

**FICHA DE CRUCE DE DATOS
UNIDAD DIDÁCTICA "DERMOVET, 1ª PUESTA EN PRÁCTICA"**

EVALUACIÓN DEL USO DE CORPUS: Actuación ante los corpus sólo cuando se utilicen
Traducción del texto "Lesiones primarias; Lesiones secundarias"

Leyenda:
Problema 1: ¿Qué le sucede a mi...? Problema 2: Son indicativas de pioderma... / Es indicativo de Pénfigo Foliáceo... Problema 3: Habón Problema 4: Tumor: masas neoplásicas tanto benignas como malignas. Se utiliza cuando hay nódulos muy grandes. Problema 5: estrato córneo / basal
Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación, ...
Indicador 1: selecciona un corpus adecuado a sus necesidades ("steb 01-02"; "steb 02-03"; <i>wordbanks online</i> ; Internet a través de <i>WebCorp</i>) Indicador 2: busca una palabra/cadena de palabras adecuada a sus necesidades Indicador 3: limita la búsqueda de una forma conveniente (búsqueda truncada, búsqueda con palabra de contexto) Indicador 4: ordena los resultados de una forma conveniente Indicador 5: usa las funciones del programa de una forma conveniente Indicador 6: emplea un tiempo razonable en resolver un problema
0: Mal 1: Mejorable 2: Bien (entre las diversas soluciones que en ocasiones pueden estar bien)
Acceptabilidad (del equivalente propuesto)
√: Correcto (entre las diversas soluciones que en ocasiones pueden estar bien) X: Incorrecto

Alumnas P.R.+V.V.	Problema 1	Problema 2	Problema 3	Problema 4	Problema 5	PUNTUACIÓN PARCIAL	PUNTUACIÓN PARCIAL (después de aplicar los %)	PUNTUACIÓN FINAL (sobre 10)
Indicador 1 (20%)		2	2	2	2	8/8 = 1	2	9,2
Indicador 2 (20%)		2	2	2	2	8/8 = 1	2	
Indicador 3 (20%)		2		0	2	4/6 = 0,6	1,2	
Indicador 4 (20%)		2		2	2	6/6 = 1	2	
Indicador 5 (10%)		2	2	2	2	6/6 = 1	1	
Indicador 6 (10%)		2	2	2	2	6/6 = 1	1	
Acceptabilidad	√	√	√	X	√			

Documento 11. Ficha de cruce de datos cumplimentada por la docente para la evaluación del uso de corpus de dos alumnas que realizaron la UD "Dermovet, 1ª puesta en práctica".

Este ejemplo nos sirve para ilustrar varios de los casos que pueden darse al rellenar la ficha de cruce de datos diseñada. La primera columna, es decir, la correspondiente al problema 1, es una muestra de que, sin necesidad de utilizar ningún corpus, las alumnas llegaron a una solución correcta. No usar los recursos no es penalizable en

tanto que el alumno tiene la libertad de utilizarlos o no según sus necesidades; obviamente, si conoce o cree que conoce la solución, se le permite ofrecerla sin más. Los problemas 2 y 5 ilustran un uso de corpus adecuado en cada elemento con un resultado correcto, al igual que ocurre en el problema 3, con la diferencia de que aquí las alumnas no han necesitado hacer determinadas operaciones relacionadas con varios de los elementos para llegar a un resultado satisfactorio. En el problema 4 nos detendremos para poder seguirlo de cerca con la ayuda de la ficha de recogida de datos incluida más arriba. Las respuestas aportadas por las alumnas nos permiten observar que para el problema concreto, un problema de tipo textual por el error de cohesión ya introducido en el TO pero también de variante geográfica por las diferencias en la ortografía del equivalente en inglés británico (*tumour*) e inglés americano (*tumor*), tanto la selección de los corpus (los dos corpus ad hoc especializados en dermatología veterinaria en inglés) como la selección de las palabras de búsqueda fue correcta. Incluso la ordenación de los resultados (1L, 1R) y el uso de las funciones del programa *WordSmith Tools (Grow)* estaban bien encaminados. Sin embargo, y a pesar de la aparente eficacia del proceso seguido, el resultado final fue incorrecto. El error, el hecho de incluir una ortografía de inglés americano y no británico, como así lo exigía el encargo de traducción, puede tener su origen en dos aspectos relacionados entre sí: la no observación del origen de los documentos consultados, ya que ambos corpus ad hoc contenían las dos variantes del inglés y, por otra parte, si nos fijamos en la ficha de recogida de datos, la no limitación de la búsqueda en cuanto al uso del asterisco, por ejemplo, ya que la búsqueda de "tum*r" hubiera proporcionado a las estudiantes concordancias con las dos variantes ortográficas, lo que, seguramente, les hubiera llevado a una observación más detallada y a un resultado óptimo. Así, como se puede observar en la ficha de cruce de datos, la no realización de operaciones en alguno de los elementos pero que acaba en un resultado correcto, no se premia ni se penaliza; mientras tanto, un resultado incorrecto se interpreta como un fallo en el uso de corpus, razonamiento que lleva a penalizar (con un "0") la no utilización de tal recurso.

Una vez introducida la valoración (0, 1 ó 2) de las operaciones realizadas para la resolución de cada problema, podemos pasar a realizar los cálculos para extraer la puntuación obtenida en cada elemento, es decir, si antes trabajamos con la ficha de cruce de datos de forma vertical, ahora lo hacemos de forma horizontal. Como se puede observar en la columna reservada para la puntuación parcial, el divisor es la cifra máxima que se puede conseguir sumando un 2 por cada elemento utilizado y el dividendo es la suma de los puntos que de hecho ha conseguido el alumno con sus operaciones. Así, por ejemplo, en el caso del indicador 3 más arriba, la máxima

puntuación que se podía haber conseguido al realizar operaciones relacionadas con 3 indicadores sería un 6 pero como una de ellas se ha puntuado con un “0”, la relación queda en un 4/6. La última columna, como se puede ver, refleja la puntuación sobre 10 una vez se han aplicado los tantos por ciento estipulados para cada elemento, lo que, todo sumado, nos da una nota de 9,2 en el uso de corpus para traducir en este caso.

En esta forma de evaluación también pueden darse algunos casos especiales que ahora explicamos. Se podría dar el caso de que, sistemáticamente, un estudiante nunca realizase ninguna operación en uno o más elementos, lo que provocaría que al menos una fila entera estuviera vacía, sin puntuación. En esa situación sería necesario alterar el cálculo del porcentaje y, si el elemento que ha sido consistentemente ignorado valiera, por ejemplo, un 20%, entonces el total ya no podría calcularse sobre un 100% sino que tendría que calcularse sobre 80. Este caso, por otra parte, sólo se podría dar si todas las soluciones son correctas ya que si no, como hemos visto antes, ha de introducirse un “0” en cada casilla vacía.

Otro caso más improbable sería aquél en el que todas las operaciones realizadas en todos los elementos fueran correctas o mejorables pero se llegara a un resultado incorrecto. Es decir, tendríamos la columna de un problema rellena con valores 1 y 2 pero con un resultado incorrecto (X) en la última casilla, la variable de control. En este caso, pensamos, sería necesario revisar muy detenidamente el proceso seguido en el uso de corpus para llegar a la traducción e intentar explicar la etiología del error. Éste sería, sin embargo, un caso de una gran rareza al implicar que un proceso correcto puede llevar a un producto incorrecto. Por otro lado, implicaría también que la adquisición de los objetivos de aprendizaje ha fallado en algún punto.

6.7.4.2. Herramientas de evaluación de la enseñanza

6.7.4.2.1. Ficha de registro de observación

Se han utilizado registros de observación con la finalidad de que el profesor, sobre todo, recoja información sobre la adecuación de las unidades y tareas diseñadas, impresiones sobre la asimilación de los objetivos, la progresión de los contenidos, el tiempo empleado, etc. A continuación se muestra un ejemplo real de un registro de observación:

COMENTARIOS TRAS LAS CLASES DE SEMINARIO DE TRADUCCIÓN C	
25/02/03	
COMENTARIOS GENERALES	
-	Les cuesta un poco asimilar el concepto de corpus. Se lo definí brevemente el día anterior y hoy ya no lo tenían muy claro.
-	Les ayuda ver el recurso de donde han salido los textos (ej. ver las páginas web del <i>Financial Times</i> y el <i>Expansión</i>) para hacerse una idea de los textos.
-	También les ayuda el ver los textos tal y como están dentro del corpus (archivos de texto).
-	Así todo, la alumna M.M. entrega su ficha sobre "fund-s" basándose en los textos que hay en la página web del <i>Financial Times</i> y no en el corpus.
COMENTARIOS SOBRE TAREAS	
Tarea 1	
-	Tardamos 2 horas en explicar claramente lo que significa un corpus y enseñarles lo que hay dentro de los corpus <i>Financial Times</i> y <i>Expansión</i> + enseñar las 4 primeras funciones del WST (con ejercicio Stock Exchange/bourse).
COMENTARIOS SOBRE FICHAS	
Ficha 1	
-	Aclarar que los apartados "Qué tipo de documento es ...?" y "En qué tipo de documentos aparece la palabra" no es necesariamente lo mismo. Sí lo es en el caso de la 1ª columna.

Documento 12. Ficha de registro de observación cumplimentada por la docente tras poner en práctica una tarea de la UD "Introducción al uso de corpus electrónicos para traducir".

6.7.4.2.2. Ficha de valoración de la unidad didáctica

En esta última ficha se invita al alumno a expresar libremente y de forma redactada sus opiniones, comentarios y sugerencias sobre la unidad didáctica, de la que mostramos un ejemplo real.

FICHA DE VALORACIÓN DE LA UNIDAD UNIDAD DIDÁCTICA "Dermovet, 1ª puesta en práctica"	
-	¿Crees que utilizando corpus has aprendido también una nueva forma de documentarme para resolver los problemas de traducción del texto? Por favor, explicita al máximo. Sí, sobre todo con los corpus especializados, que ayudan a contextualizar términos y a descubrir su significado.....
-	¿Hay algo de esta unidad didáctica que te haya gustado especialmente? Por favor, explicita al máximo. El mero hecho de aprender nuevas maneras de documentarme aporta mucho y, particularmente, el uso de <i>WordSmith Tools</i> con los corpus me parece un modo muy rápido y eficaz de obtener o comprobar información.....
-	¿Hay algo que no te haya gustado? Por favor, explicita al máximo. Sinceramente, la traducción financiera no es la que más me atrae pero aún así, y gracias a todo lo que he aprendido sobre corpus y <i>WordSmith Tools</i> , me ha parecido bastante amena...
-	¿Cómo lo cambiarías? Trabajando otro tipo de textos especializados aunque, evidentemente, se tendrían que utilizar también otros corpus especializados.....
-	Otros comentarios ...

Documento 13. Ficha de valoración cumplimentada por un estudiante tras realizar la UD "Dermovet, 1ª puesta en práctica".

La información tanto en forma numérica como redactada extraída de las fichas de autoevaluación y de valoración como las anteriores se comentará con más detalle en el marco de cada unidad didáctica. En el apartado de “Evolución de la unidad” se incluirán datos relevantes para la reestructuración de las unidades a partir de las respuestas de los estudiantes.

En resumen, hemos intentado poner en práctica una evaluación del uso de corpus para traducir que nos aporte información sobre la enseñanza para poder mejorarla. Para poder llevar a cabo tal evaluación ha sido necesario diseñar herramientas variadas entre las que se encuentran diversos tipos de fichas, un listado de indicadores y otros. Todas ellas han sido probadas en el aula y han demostrado su utilidad a la hora de evaluar diversos aspectos del proceso de aprendizaje de los estudiantes en varios niveles, materias, fases pedagógicas, etc., así como en todas las unidades didácticas que se describen en el siguiente capítulo.

6.8. Evolución de la propuesta

6.8.1. Fase previa al diseño

La propuesta didáctica que aquí se realiza comenzó con el planteamiento de la necesidad de incorporar el uso de corpus electrónicos a las clases de traducción por todo lo que éstos pueden aportar tal como textos reales, cantidad y calidad en el manejo de textos paralelos o de apoyo, reducción del peso del docente y del componente intuitivo en la enseñanza, mayor protagonismo del estudiante y el recurso, mayor autonomía del alumno, incorporación de las nuevas tecnologías, etc. Seguidamente, la puesta en práctica de ejercicios o actividades sueltas a modo de prueba, no específicamente concatenadas para conformar una unidad didáctica completa y sobre todo en la materia de traducción inversa especializada, nos aportó un *feedback* inicial que, entre otras cosas, nos permitió hacer una observación del alumnado y tener una idea más realista del tiempo necesario para que los estudiantes absorban los conocimientos y adquieran algunas habilidades de manejo de corpus, así como también nos permite, por ejemplo, identificar usos no previstos para el corpus como ayudar al alumno a evitar errores producidos por la cercanía de dos lenguas. Asimismo, estas primeras pruebas nos hicieron constatar que la enseñanza del uso de corpus aplicado a la traducción se debía sistematizar para no quedarse en lo anecdótico e impresionista.

Por otro lado, para no sólo conocer la situación dentro del aula, sino también fuera, en el mundo real, se realizaron dos sondeos a pequeña escala, uno sobre la práctica profesional de la traducción inversa que, entre otras cosas, nos reveló los campos temáticos y géneros textuales más solicitados (cfr. supra 6.2.4.1.), y otro sondeo sobre el uso de corpus electrónicos en agencias de traducción que constató que, si bien las agencias no utilizan corpus como tal, sí utilizan textos paralelos y de apoyo en forma de memorias de traducción, que en realidad viene a ser un corpus paralelo (cfr. supra 2.4.). Por último, la obtención del informe sobre la situación del mercado de la traducción en España realizado por la ACT y adquirido por el grupo Tradumática, también aportó datos clave para apoyar la selección de objetivos, materias y materiales para el diseño de algunas de las unidades didácticas.

6.8.2. Diseño y experimentación

La lectura de bibliografía sobre la evolución y estado de la cuestión sobre la enseñanza de la traducción nos llevó a la adopción del enfoque por tareas de traducción de Hurtado como marco metodológico que permite la sistematización e integración de todos los elementos que forman parte de la enseñanza.

Como bien indica Kelly (2005: 21-22), antes de pasar a plantear los objetivos de aprendizaje, en este caso, del uso de corpus para traducir, cabe preguntarse sobre el origen de los mismos y los factores que los determinan. En el caso que nos ocupa, la observación de las características del alumnado implicado nos ofrece un perfil variado pero con elementos comunes como por ejemplo los conocimientos y práctica limitados de las nuevas tecnologías, con acceso a ordenadores e Internet, con necesidad de una mayor práctica de sus habilidades informáticas en contexto, un grado de madurez medio-alto, una experiencia en traducción heterogénea en cada grupo y una buena infraestructura en el centro de traducción, un perfil que, con seguridad, se podría hacer extensible al resto de los alumnos de traducción de España. La encuesta a varias agencias de traducción sobre su uso de corpus y textos paralelos es un dato sobre las necesidades del mercado y otro de los factores que marcan el diseño de los objetivos de aprendizaje en esta propuesta, una propuesta que también cuenta con las limitaciones del contexto como son la duración de las materias o contenidos formativos.

Teniendo todos estos elementos en cuenta, se plantearon los objetivos de aprendizaje para dos fases pedagógicas del uso de corpus para traducir, una de iniciación y otra de

consolidación, se diseñaron y pusieron en práctica todas las unidades didácticas, que resultaron siendo 7 en total.

6.8.3. Propuesta definitiva

Una vez realizada la puesta en práctica de todas las unidades didácticas, identificados los puntos susceptibles de mejora y analizados los resultados de la evaluación de las mismas se procedió a introducir cambios en las unidades. El *feedback* recogido de la puesta en práctica y, en gran medida, de las fichas de autoevaluación provocó cambios en las fichas para, por ejemplo, aclarar alguna pregunta, cambios en los materiales aportados para ofrecer recursos de mayor ayuda, cambios en el desarrollo para permitir más o menos tiempo en la realización de una tarea, etc. Con esto, las unidades se dan por probadas y revisadas y están preparadas para ser utilizadas de nuevo por otros estudiantes o docentes.

PARTE III.

PROPUESTA DE UNIDADES DIDÁCTICAS

“Para que el compromiso entre calidad y rentabilidad sea posible
no queda más salida para el traductor
que el uso de procedimientos de documentación muy eficaces”.
(Mayoral 1997: 138)

“Corpus data has to be interpreted and evaluated comparatively to reach conclusions,
and this requires not only technical skill [...], but above all critical thought.
Training would-be translators to use corpora goes hand in hand with educating them
to think about the translation process and the learning process,
developing their sensitivity as to how they can use corpora in these processes”.
(Aston 2008)

INTRODUCCIÓN

En este capítulo nos centramos en la exposición de las unidades didácticas, cómo se han organizado y la planificación detallada de cada una de ellas.

1. Procedimiento de presentación de las unidades didácticas

A continuación se muestra el esquema concreto que se ha seguido para la presentación de cada unidad didáctica (UD), que consta de varias secciones. Estas secciones se han distribuido en dos volúmenes con el fin de facilitar la lectura y la consulta del material.

VOLUMEN 1

- ❶ PRESENTACIÓN DE LA UNIDAD
- ❷ DESARROLLO DE LA UNIDAD
- ❸ EVALUACIÓN DE LA UNIDAD
- ❹ EVOLUCIÓN DE LA UNIDAD

VOLUMEN 2

- 📖 MATERIALES
- ↔ SOLUCIONES
- ☑ EVALUACIÓN DE LA UNIDAD

Lo que sigue es una breve explicación de lo que incluye cada sección de cada volumen:

VOLUMEN 1

❶ PRESENTACIÓN DE LA UNIDAD

En la presentación de cada unidad se indican el título, el nivel de la unidad respecto al uso de corpus (iniciación o consolidación), así como la o las materias en las que se puede poner en práctica. Además, los objetivos se desglosan en varios tipos: relacionados con el uso de corpus, con la materia de traducción, con competencias genéricas, siendo los primeros los más importantes dentro de esta propuesta. Por otro lado, se listan los contenidos tratados, las tareas que componen la UD, los indicadores para la evaluación y todos los materiales utilizados; y se muestra una estimación de la duración total de la UD completa. La estructura general de una unidad didáctica es la siguiente:

<ul style="list-style-type: none"> • TÍTULO • NIVEL • MATERIA(S) • OBJETIVOS <ul style="list-style-type: none"> - Objetivos relacionados con el uso de corpus para traducir - Objetivos relacionados con la materia - Objetivos relacionados con competencias genéricas • CONTENIDOS • ESTRUCTURA <ul style="list-style-type: none"> - TAREA 1 - TAREA 2 - TAREA ... - (TAREA FINAL) • EVALUACIÓN <ul style="list-style-type: none"> - Indicadores - Fichas de evaluación • MATERIALES <ul style="list-style-type: none"> - Fichas - Textos (originales y paralelos) - Textos de apoyo - Recursos electrónicos • DURACIÓN ESTIMADA
--

Tabla 17. Elementos incluidos en la presentación de cada unidad didáctica.

❷ DESARROLLO DE LAS TAREAS DE CADA UNIDAD

Ésta es una sección dirigida al profesor y que se presenta en formato de ficha en la que se incluyen todos los apartados que contempla cada tarea.

- **Introducción:** apunta la razón de ser de la tarea y su finalidad.
- **Desarrollo:** actividades, dinámica de grupos, duración estimada, etc.
- **Materiales:** textos originales para traducir o contextualizar elementos lingüísticos sobre las que se trabaja, textos paralelos, textos de apoyo, fichas de ejercicios, recursos electrónicos, etc.
- **Evaluación:** el tipo y los medios de evaluación, si cabe, ya que no es necesario que todas las tareas se evalúen.

- Comentarios: sobre el desarrollo de la tarea, antes o después de su puesta en práctica (anotaciones, recordatorios, conveniencia de la realización de pre- o postareas).

ESTRUCTURACIÓN DE CADA TAREA
<ul style="list-style-type: none"> • INTRODUCCIÓN • DESARROLLO DE LA TAREA • MATERIALES • (EVALUACIÓN) • COMENTARIOS

Tabla 5. Estructuración de la tarea de traducción (Hurtado 1999).

③ EVALUACIÓN DE LA UNIDAD

En el Volumen I se ofrecen los resultados de la evaluación desde dos perspectivas:

1. Valoración realizada por los estudiantes.
 - a. Valoración de su aprendizaje (recogida en la ficha de autoevaluación).
 - b. Valoración de la unidad didáctica (recogida en la ficha de valoración).
2. Valoración realizada por el docente.
 - a. Valoración del aprendizaje de los estudiantes (extraída de la observación de sus tareas, de su diario de aprendizaje, sus fichas de recogida de datos sobre el uso de corpus y las fichas de cruce de datos).
 - b. Valoración de la unidad didáctica (valoración final o comentarios por parte del docente sobre la experiencia).

④ EVOLUCIÓN DE LA UNIDAD

Se comenta aquí el contexto real en el que la unidad se pilotó y se incluyen los cambios efectuados a raíz de su puesta en práctica.

VOLUMEN 2

📖 MATERIALES

Ésta es la sección dirigida al estudiante y que incluye todos los materiales que utilizará para la realización de la unidad didáctica.

- Fichas de trabajo
- Fichas de evaluación
 - Autoevaluación
 - Valoración de la unidad
 - Cruce de datos

- Textos
 - o Textos originales y paralelos
 - o Textos de apoyo
- Descripciones: de corpus*
- Listas: de palabras ya extraídas de corpus*
- Concordancias: ya extraídas de corpus*
- Recursos electrónicos
 - o Recursos lexicográficos
 - o Corpus
 - o Software

↔ **SOLUCIONES**

Es un “solucionario” de las fichas propuestas o, al menos, un ejemplo o una previsión de las posibilidades de solución.

EVALUACIÓN DE LA UNIDAD

En el Volumen II se ofrecen las pruebas o evidencias de la evaluación, es decir, las fichas de cruces de datos rellenas para cada alumno en su tarea final, si es que la UD la contempla, y las respuestas redactadas (no numéricas) de todos los estudiantes a fichas como la de autoevaluación o valoración de la unidad.



Es necesario precisar que existen elementos que se dan en todas las unidades didácticas diseñadas y que por ello no se especificarán en la presentación de cada una. Estos elementos comunes son:

- Todos los materiales y recursos facilitados a los alumnos están en formato electrónico, a no ser que se indique lo contrario, y disponibles a través de una página web institucional, una unidad compartida o a través de una plataforma virtual de aprendizaje para estudios no presenciales o semi-presenciales.
- Todas las unidades didácticas diseñadas siempre se han llevado a cabo en aulas de informática con acceso a Internet y todo el software específico instalado.
- Todos los corpus electrónicos utilizados son de textos actuales (p.ej. siempre que hablemos del corpus de la RAE, nos referiremos al *CREA* y no al *CORDE*), a no ser que se indique lo contrario.

* Este tipo de materiales no se utiliza siempre.

- Todos los encargos de traducción al y del inglés se refieren a la variante británica; por esta razón, hemos intentado fijarnos especialmente en textos y recursos de esta variante, si bien es cierto que no se desdeñan otros de origen norteamericano total o parcial (p.ej. una parte de los textos del corpus *Bank of English – WordBanksOnline*- que está accesible en Internet es de inglés norteamericano, aunque la mayoría es británico). En el caso del español siempre nos referimos a la variante peninsular.
- En la duración estimada de cada tarea se ha tomado en cuenta la suma de lo siguiente:
 - o Trabajo presencial
 - o Trabajo supervisado
 - o Trabajo autónomo

Además, también es necesario puntualizar que cada unidad didáctica se ha probado con un grupo de estudiantes diferente, excepto en las unidades de introducción y contrastividad, que se trabajaron en el mismo grupo y materia. La información sobre el grupo o grupos de estudiantes donde se probaron las unidades se facilita al comienzo de cada unidad. Ahí se especifica su ubicación, es decir, la materia o asignatura en la que de hecho se puso en práctica, y en las secciones dedicadas a la evaluación se podrá comprobar que el número e iniciales de los alumnos evaluados cambian de una unidad a otra.

2. Unidades didácticas

A continuación se presentan las unidades didácticas diseñadas y puestas en práctica en el marco del presente trabajo y centradas en el uso de corpus para traducir en diversos contextos. Las unidades didácticas están elaboradas en torno a varios objetivos para captar los contenidos y persiguen en casi todos los casos una tarea final. Siendo la flexibilidad una de las ventajas que permite el marco metodológico adoptado, las unidades aquí presentadas son susceptibles de ser adaptadas a otras situaciones de aprendizaje y los objetivos perseguidos también podrían alcanzarse con otras unidades.

Las unidades que siguen son siete:



“Introducción al uso de corpus electrónicos para traducir”



“Contrastes de uso en un corpus paralelo”



“Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”



“Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”



“El corpus ad hoc como remedio para la traducción médica al inglés”



“Corpus y memoria de traducción”



“Sólo utilizar corpus; utilizar corpus solo”



UNIDAD DIDÁCTICA

Introducción al uso de corpus electrónicos para traducir



Con esta unidad didáctica los estudiantes empiezan a conocer algunas de las ventajas que supone trabajar con corpus electrónicos a la hora de documentarse para una traducción. Aquí se sientan algunas de las bases fundamentales del uso de corpus, como por ejemplo la importancia de la composición y tamaño de un corpus y su influencia en los resultados que se puedan obtener de él. Se identifican diversas formas de extraer información de un corpus a través de algunas funciones del programa *WordSmith Tools*, un tipo de software fundamental para la explotación de corpus, que probará ser de gran ayuda en cuanto a la cantidad de textos que permite manejar en poco tiempo, el tipo de datos que permite obtener, y la objetividad de los resultados. Se empieza aquí a guiar al estudiante hacia el desarrollo de un espíritu crítico, autónomo e investigador ya que estará en sus manos la formulación de las preguntas y las búsquedas en el corpus, así como la interpretación de los resultados.



Grupos de prueba: La unidad original se probó en un grupo de 8 alumnos de 4º curso de la licenciatura de Traducción e Interpretación de la UAB, en una asignatura de traducción (inglés C-español), durante el año académico 2002-03.

Los materiales de esta unidad se encuentran en el *Volumen 2* de esta tesis (págs. 411-456).



❶ PRESENTACIÓN DE LA UNIDAD DIDÁCTICA

“Introducción al uso de corpus electrónicos para traducir”

NIVEL DE USO DE CORPUS

- Iniciación

MATERIA

- Iniciación a la traducción directa (inglés B/C – español A)

OBJETIVOS

Objetivos relacionados con el uso de corpus

- Conocer aspectos básicos sobre corpus
- Identificar los principios metodológicos básicos del trabajo con corpus
- Utilizar sistemas de búsqueda en corpus en línea a nivel básico
- Utilizar las funciones básicas de un programa informático de análisis de corpus
- Integrar conocimientos para resolver problemas de traducción utilizando corpus a nivel básico

Objetivos relacionados con la materia

- Conocer y utilizar las herramientas básicas del traductor
- Manejar los útiles informáticos básicos
- Aprender a evaluar las fuentes de documentación
- Adquirir estrategias de documentación

CONTENIDOS

- Diversos tipos de corpus (I. 1.1.)
- Corpus de inglés en Internet (*BNC*) (I. 1.2.1.)
- Corpus de español en Internet (*CREA*) (I. 1.2.2.)
- La Web como corpus (*Internet pública*) (I. 1.4.)
- El tamaño de un corpus (I. 2.1.5.)
- Relación entre los resultados arrojados por un corpus y los textos que lo componen (I. 2.2.2.)
- Formas de extraer información relevante de un corpus (I. 2.2.3.)
- Ventajas documentales de un corpus frente a otros recursos electrónicos (I. 2.2.4.)
- Ventajas del software frente al ser humano (I. 2.2.5.)
- Software comercializado de análisis de corpus (*WordSmith Tools*) (III. 1.10.2.)
- Función *Concord* (de *WordSmith Tools*) (III. 2.1.1.)
- Función *Grow/Shrink* (de *WordSmith Tools*) (III. 2.2.1.)
- Función *Resort* (de *WordSmith Tools*) (III. 2.3.1.)
- Extracción de características (IV. 2.1.3.)
- Extracción de colocaciones (IV. 2.2.2.)
- Extracción de fraseología (IV. 2.2.3.)



ESTRUCTURA

- Tarea 1: "Los resultados que ofrece el corpus dependen de los ingredientes que contenga"
- Tarea 2: "Funciones básicas de un programa de análisis de corpus para empezar a ver la utilidad de este recurso"
- Tarea 3: "Un corpus más grande no es siempre un corpus mejor"
- Tarea 4: "El ojo humano puede fallar"
- Tarea 5: "El corpus como recurso documental más completo que otros"
- Tarea 6: "Cómo hacer preguntas a un corpus"
- Tarea final: "Explotación real de un corpus en la traducción de un texto"

EVALUACIÓN

Indicadores

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación,...

- Indicador 1: selecciona un corpus adecuado
- Indicador 2: busca una palabra/cadena de palabras adecuada
- Indicador 3: limita la búsqueda
- Indicador 4: ordena los resultados
- Indicador 5: usa las funciones del programa de análisis de corpus

Fichas de evaluación

- Ficha de recogida de datos sobre el uso de corpus al traducir
- Ficha de autoevaluación
- Ficha de valoración de la unidad
- Diario de aprendizaje

MATERIALES

Fichas

- Ficha 1: "Qué contiene cada recurso"
- Ficha 2: "Funciones de un programa de análisis de corpus"
- Ficha 3: "Información extraída de una forma diferente"
- Ficha 4: "Mini-estudio del uso de 'Pero'"
- Ficha 5: "Explicación de las cifras"
- Ficha 6: "Observación humana frente a objetividad informática"
- Ficha 7: "Recursos documentales"
- Ficha 8: "Preguntas al corpus"
- Ficha de recogida de datos sobre el uso de corpus al traducir
- Ficha de autoevaluación
- Ficha de valoración de la unidad

Textos originales y paralelos

- Texto en inglés que contenga un mismo término con varias acepciones y, entre ellas, una del campo de la economía (Texto 1: "Women's Growth Capital Fund")



- Un texto en español en el que aparezcan las palabras "fondo" y "fondos", así como la conjunción adversativa "pero" a principio de oración (Texto 2: "Invertir los ahorros")
- Un texto en inglés o español que hable sobre sinergias (Texto 3: "Synergies: A business guide")
- Dos textos en español que contengan "fondo" o "fondos" con distintas acepciones (Texto 4: "Fondo Global para Mujeres"; Texto 5: "Perspectivas esperanzadoras para los Fondos de Pensiones")
- Un texto en inglés sobre economía y de género periodístico (Texto 6: "JSE expects gold futures to excite strong appetites" o Texto 7: "Britain must avoid Germany's mistake, says Feldstein")

Textos de apoyo

- 1: Material de apoyo al programa *WordSmith Tools*

Concordancias

- Concordancias de "fondo" y "fondos" extraídas del corpus del *Expansión*

Descripciones

- Descripción 1: Corpus de textos periodísticos financieros en inglés (*Financial Times*)
- Descripción 2: Corpus de textos periodísticos financieros en español (*Expansión*)

Recursos electrónicos

- URL del diccionario normativo del español [DRAE: <http://www.rae.es>]
- URL de un diccionario de español general [Vox: <http://www.diccionarios.com>]
- URL de un diccionario bilingüe general [WordReference: <http://wordreference.com>]
- URL de glosarios multilingües o bilingües de economía [FMI: http://www.imf.org/external/np/term/index.asp?index=esl&index_langid=3 ; Andy Miles: <http://www.andymiles.com>]
- URL de una página web divulgativa sobre inversiones [*El Inversor*: <http://www.elinversor.com>]
- URL de un corpus de inglés general [*British National Corpus -BNC*: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- URL de un corpus de español general [CREA: <http://www.rae.es>]
- Corpus de textos periodísticos financieros en español (*Expansión*)
- Corpus de textos periodísticos financieros en inglés (*Financial Times*)
- Programa informático *WordSmith Tools*

DURACIÓN ESTIMADA

16 horas

2 DESARROLLO DE LAS TAREAS

TAREA 1: LOS RESULTADOS QUE OFRECE EL CORPUS DEPENDEN DE LOS INGREDIENTES QUE CONTENGA

INTRODUCCIÓN:

En esta primera tarea se expone al alumno a diferentes fuentes de información o recursos de documentación, entre ellos, un corpus, para que constate que los resultados obtenidos siempre deben interpretarse a la luz de los textos que contiene dicho recurso. Dada la novedad de la metodología para los estudiantes, especialmente cuando empiecen a trabajar con corpus propiamente dichos, las primeras tareas requerirán de una monitorización y un control intensos por parte del profesor.

DESARROLLO:

1. Se facilita a los alumnos un modelo de Diario de aprendizaje para que recojan todo aquello que consideren que van aprendiendo o adquiriendo a medida que realizan la unidad.
2. Tras organizar a los alumnos en parejas, se les da el Texto 1 ("Women's Growth Capital Fund"), que contiene las palabras "fund" y "funds" con diferentes acepciones; cada miembro de la pareja debe leerlo fijándose en una de dichas palabras.
3. Tras leer el Texto 1 se les facilita la Ficha 1 ("Qué contiene cada recurso") para "fund" y "funds" para que cada uno vaya cubriéndola a lo largo de la tarea con la información que aportan los distintos recursos empleados.
4. Para que puedan rellenar la última columna del ejercicio, se les facilita un corpus de textos periodísticos en inglés del *Financial Times* extraído de un CD-ROM comercializado por ELRA y una breve descripción del mismo (Descripción 1).
5. A la vez, se va explicando el programa de análisis de corpus *WordSmith Tools* y 4 funciones básicas: *Concord*, *Grow/Shrink* y *Resort*.
6. Se facilita el material de apoyo a este programa (Texto de apoyo 1: "Material de apoyo al programa *WordSmith Tools*").
7. Se facilita la Ficha 2 ("Funciones de un programa de análisis de corpus") para que recojan las funciones que van aprendiendo del *WordSmith Tools*.

Dinámica de grupos: en parejas e individual

Duración estimada: 2 horas

MATERIALES:

Fichas

- Ficha 1: "Qué contiene cada recurso"
- Ficha 2: "Funciones de un programa de análisis de corpus"
- Diario de aprendizaje

Textos originales y paralelos

- Texto en inglés que contenga un mismo término con varias acepciones y, entre ellas, una del campo de la economía (Texto 1: "Women's Growth Capital Fund")

Textos de apoyo

- 1: Material de apoyo al programa *WordSmith Tools*

Descripciones de corpus

- Descripción 1: Corpus de textos periodísticos financieros en inglés (*Financial Times*)

Recursos electrónicos

- URL de un corpus de inglés general [*British National Corpus -BNC*: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- Corpus de textos periodísticos financieros en inglés (*Financial Times*)
- Programa informático *WordSmith Tools*

EVALUACIÓN:

Considerando que, en cada pareja, un miembro ha cubierto una Ficha 1 para "fund" y otro una Ficha 1 para "funds", al terminar repasan las fichas de ambos, comparan y comentan sobre las diferencias respecto a la información encontrada, primero dentro de cada ficha y después entre las dos fichas.

COMENTARIOS:

- Cuando los alumnos busquen su término en el BNC, se debe insistir en que miren las fuentes de los documentos de donde ha salido cada concordancia.
- Al ir rellenando la Ficha 1 se darán cuenta de que cada recurso les da un tipo, cantidad y variedad de información más o menos diferente, lo que se debe a la composición de cada "corpus", conclusión que debe quedar clara.

TAREA 2: FUNCIONES BÁSICAS DE UN PROGRAMA DE ANÁLISIS DE CORPUS PARA EMPEZAR A VER LA UTILIDAD DE ESTE RECURSO**INTRODUCCIÓN:**

Esta tarea busca un objetivo doble; por un lado, que el estudiante asimile cómo se puede llegar a extraer de un corpus información variada y de distintas maneras, y por otro lado, que con ese pretexto, aprenda y practique algunas funciones básicas del programa *WordSmith Tools*. Dado que el énfasis está en aprender a traducir, es importante concienciar a los alumnos sobre los patrones que se dan en la lengua e introducirles en su búsqueda, aspectos éstos que son, en cierto modo, los cimientos del trabajo con corpus, y que se empiezan a trabajar en esta tarea.

DESARROLLO:

1. Se facilita la Ficha 3 ("Información extraída de una forma diferente") en la que los estudiantes deben utilizar el corpus del *Financial Times* y diversas funciones del programa *WordSmith Tools* para poder responder.
2. Para la pregunta sobre siglas se ilustra la posibilidad de realizar búsquedas truncadas desde la función *Concord*. En este caso una buena búsqueda sería (*SE) para extraer siglas compuestas con "Stock Exchange".
3. Para la pregunta sobre si los dos términos en cuestión aparecen en la misma frase, los alumnos utilizan la función *Concord* pero ahora ayudándose del apartado para introducir una palabra de contexto (*context word*). Deben buscar "Stock Exchange" + "bourse" con un horizonte de 10 palabras a la derecha y a la izquierda. Se analizan en voz alta los 4 casos que aparecen (en 1 caso "bourse" está en francés; en 2 casos parece que la alternancia se usa para evitar una repetición y en otro caso "bourse" está en una construcción como premodificador – "off-bourse trade").
4. Para poder ver bien los contextos anteriores es necesario aumentar el contexto y luego reducirlo para conseguir una visualización más rápida de las concordancias, lo que es posible con las funciones *Grow/Shrink*.
5. Los alumnos incorporan a la Ficha 2 ("Funciones de un programa de análisis de corpus") las funciones nuevas que han aprendido del programa *WordSmith Tools*.

Dinámica de grupos: individual

Duración estimada: 2 horas

MATERIALES:**Fichas**

- Ficha 2: "Funciones de un programa de análisis de corpus"
- Ficha 3: "Información extraída de una forma diferente"

Recursos electrónicos

- URL de un glosario bilingüe de economía y finanzas [AndyMiles: <http://www.andymiles.com>]
- Corpus de textos periodísticos financieros en inglés (*Financial Times*)
- Programa informático *WordSmith Tools*

EVALUACIÓN:

Comentando los resultados todos juntos se llega a la conclusión de que parece que ambos términos pueden significar lo mismo en algunos contextos pero tienen usos diferentes, es decir, que no son 100% sinónimos porque no son 100% intercambiables.

COMENTARIOS:

- Cuando se den pocos casos en una búsqueda frente a un patrón más frecuente (p.ej. nombre de ciudad+Stock Exchange/bourse), se debe insistir en que no significa que el primer caso esté mal.

- La Ficha 2 se irá rellenando con las funciones del programa *WordSmith Tools* que vayan aprendiendo en ésta y otras tareas que vendrán más adelante.

TAREA 3: UN CORPUS MÁS GRANDE NO ES SIEMPRE MEJOR

INTRODUCCIÓN:

Esta tarea enlaza con la Tarea 1 en tanto que recuerda al alumno que los resultados de un corpus dependen de lo que éste contiene, pero también tiene que ver con cuestiones de tamaño. Con los ejercicios que se plantean aquí se darán cuenta de que un corpus grande puede ser de gran utilidad y, de hecho, imprescindible cuando se quiere hacer un estudio general de un aspecto de la lengua. Sin embargo, también verán que cuanto más pequeño el corpus y más especializado, más útil para un encargo de traducción también especializado.

DESARROLLO:

1. Se facilita el Texto 2 ("Invertir los ahorros"), que trata sobre fondos de inversión, para que lo lean rápidamente (10 min.) e intenten localizar algo que crean anómalo en el uso del español (para ahorrar tiempo, se les puede dirigir la atención hacia las conjunciones).
2. Se llama la atención sobre el uso de "Pero" a principio de frase en el texto. Los alumnos comentan este uso en voz alta y hacen la Ficha 4 ("Mini-estudio del uso de 'Pero'"), que no sólo les hará reflexionar sobre la intuición del hablante nativo, sino que les hará ver que para estudiar un fenómeno de la lengua general, un corpus como el *CREA* de la RAE es apropiado por sus características (representa el español general, permite seleccionar documentos sólo en español de España, documentos relativamente actuales, distintos medios y campos temáticos, etc.).
3. Para acabar de asimilar la utilidad de un corpus grande y general o pequeño y especializado, los alumnos buscan "fondo" y "fondos" en el corpus del *Expansión* y el *CREA*, es decir, en un corpus especializado y en un corpus general.
4. Rellenan la Ficha 5 ("Explicación de las cifras") en la que tienen que explicar las cifras obtenidas en las búsquedas de "fondo/s" y dar ejemplos extraídos de ambos corpus.
5. Cada pareja se junta con otra pareja para comentar sus resultados, tanto de la Ficha 4 como de la Ficha 5 para llegar a respuestas comunes y más completas.

Dinámica de grupos: en parejas

Duración estimada: 2 horas

MATERIALES:

Fichas

- Ficha 4: "Mini-estudio del uso de 'Pero'"
- Ficha 5: "Explicación de las cifras"

Textos

- Un texto en español en el que aparezcan las palabras “fondo” y “fondos”, así como la conjunción adversativa “pero” tras un punto (Texto 2: “Invertir los ahorros”)

Recursos electrónicos

- URL de un corpus de español general [CREA: <http://www.rae.es>]
- Corpus de textos periodísticos financieros en español (*Expansión*) y su descripción (Descripción 2)

TAREA 4: EL OJO HUMANO PUEDE FALLAR**INTRODUCCIÓN:**

Esta tarea conecta con la anterior en tanto que hace reflexionar al estudiante sobre su capacidad intuitiva y observadora, no para minar su confianza pero sí para fomentar la autocrítica de las propias competencias y confiar en las herramientas informáticas para algunas operaciones, especialmente de tipo cuantitativo.

DESARROLLO:

1. Se facilita un texto en inglés o español (Texto 3: “Synergies: A business guide”) en el que aparecen palabras de significado o uso desconocido para los alumnos. En este caso se trata del término “synergy” / “synergies”, y el texto es para que lo contextualicen.
2. Se facilita a los alumnos la Ficha 6 (“Observación humana frente a objetividad informática”) para que recojan datos encontrados acerca del término tras efectuar un análisis humano y uno automático de unas concordancias.
3. Se facilita la URL de un corpus online de español general (CREA) y un corpus de textos especializados en español (*Expansión*).
4. De los resultados del CREA (sólo variante de español de España) tienen que intentar extraer patrones “a ojo” de la palabra “sinergia/s” (repeticiones, características de la palabra) y hacer lo mismo con los resultados del corpus del *Expansión* pero con la ayuda del programa *WordSmith Tools*.
5. Se practica la función de *Resort* y se explican e ilustran a la vez las nuevas funciones *Collocates* y *Clusters*.
6. Se recupera la Ficha 2 (“Funciones de un programa de análisis de corpus”) para introducir las nuevas utilidades del *WordSmith Tools*, que aún estaba sin completar.
7. Acaban de rellenar la Ficha 2 y la Ficha 6.

Dinámica de grupos: en parejas

Duración estimada: 3 horas

MATERIALES:**Fichas**

- Ficha 2: “Funciones de un programa de análisis de corpus”
- Ficha 6: “Observación humana frente a objetividad informática”

Textos

- Un texto en inglés o español que hable sobre sinergias (Texto 3: "Synergies: A business guide")

Recursos electrónicos

- URL de un corpus de español general [CREA: <http://www.rae.es>]
- Corpus de textos periodísticos financieros en español (*Expansión*)

EVALUACIÓN:

Cada pareja realiza los mismos pasos para el análisis del término "fluctuación" (CREA=73 casos; *Expansión*= 145) y cubre otra Ficha 6 para este término.

COMENTARIOS:

- Se prevé que el conocimiento previo sobre el concepto de "sinergia" sea muy escaso o nulo. Se darán cuenta de que sus observaciones "a ojo" son menos exactas, más intuitivas y más lentas, mientras que las hechas con ayuda del corpus y el *WordSmith Tools* son exactas, objetivables, repetibles y más rápidas.
- Las funciones del *WordSmith Tools* de extracción de colocaciones y agrupaciones, *Collocates* y *Clusters*, respectivamente, se suelen mostrar en la fase de consolidación en el uso de corpus para traducir. Sin embargo, su comprensión y ejecución se pueden adelantar si el profesor realiza una actividad más guiada, como es el caso.

TAREA 5: EL CORPUS COMO RECURSO DOCUMENTAL MÁS COMPLETO QUE OTROS**INTRODUCCIÓN:**

Esta tarea busca mostrar al estudiante la amplitud y variedad de información que puede mostrar un corpus especializado en combinación con un programa de análisis de corpus frente a la menor cantidad de información que ofrecen otros recursos.

DESARROLLO:

1. Se facilitan a los alumnos el Texto 4 ("Fondo Global para Mujeres") y el Texto 5 ("Perspectivas esperanzadoras para los Fondos de Pensiones"), donde aparecen las palabras "fondo" y "fondos" para que hagan una lectura rápida del texto (15 min.) y extraigan sus varias acepciones.
2. Se les facilita la Ficha 7 ("Recursos documentales") para que, a medida que realizan la tarea buscando las palabras "fondo" o "fondos" en los distintos recursos que se proponen, vayan rellenándola con la información que cada uno ofrece.
3. Guardan las concordancias de "fondo" / "fondos" extraídas del corpus del *Expansión* para futuros usos.

Dinámica de grupos: en parejas

Duración estimada: 2 horas

MATERIALES:

Fichas

- Ficha 7: "Recursos documentales"

Textos

- Dos textos en español que contengan "fondo" o "fondos" con distintas acepciones (Texto 4: "Fondo Global para Mujeres"; Texto 5: "Perspectivas esperanzadoras para los Fondos de Pensiones")

Recursos electrónicos

- URL del diccionario normativo del español [DRAE: <http://www.rae.es>]
- URL de un diccionario de español general [Vox: <http://www.diccionarios.com>]
- URL de un diccionario bilingüe general [WordReference: <http://wordreference.com>]
- URL de glosarios multilingües o bilingües de economía [FMI: http://www.imf.org/external/np/term/index.asp?index=esl&index_langid=3 ; Andy Miles: <http://www.andymiles.com>]
- URL de una página web divulgativa sobre inversiones [El Inversor: <http://www.elinversor.com>]
- URL de un corpus de español general [CREA: <http://www.rae.es>]
- Corpus de textos periodísticos financieros en español (*Expansión*)
- Programa informático *WordSmith Tools*

EVALUACIÓN:

Una vez rellenada la Ficha 7 por parejas, los alumnos se juntan en grupos de 4 y comentan sus resultados brevemente. Se comentan en voz alta las conclusiones y se incide en las ventajas que aporta el uso de corpus que no hayan sido captadas.

TAREA 6: CÓMO EXTRAER INFORMACIÓN DE UN CORPUS

INTRODUCCIÓN:

Esta tarea es clave en tanto que busca que el alumno capte cómo puede obtener la información que necesita de un corpus utilizando diversas funciones de un programa de análisis de corpus.

DESARROLLO:

1. Los estudiantes recuperan las concordancias que extrajeron en la tarea anterior de "fondo" y "fondos" del corpus del *Expansión* con ayuda del *WordSmith Tools*.

2. Se les facilita la Ficha 8 ("Preguntas al corpus"), donde recogerán las estrategias de búsqueda y funciones del *WordSmith Tools* con las que extraigan información del corpus del *Expansión*.
3. Para poder observar mejor los patrones que aparezcan en las concordancias y rellenar la ficha, se practican las funciones *Collocates* y *Clusters*.
4. Un miembro de la pareja rellena una Ficha 8 para "fondo" y el otro miembro otra ficha para "fondos".
5. Se intercambian las fichas dentro de cada pareja, comentan sus soluciones y completan, si es necesario, sus propias fichas.

Dinámica de grupos: en parejas

Duración estimada: 2 horas

MATERIALES:

Fichas

- Ficha 8: "Preguntas al corpus"

Recursos electrónicos

- Concordancias de "fondo" y "fondos" extraídas del corpus del *Expansión* y guardadas en la tarea anterior
- Programa informático *WordSmith Tools*

COMENTARIOS:

- Si se aprecia que no se han recogido todas las posibilidades de interrogar un corpus, también se pueden comentar las fichas en voz alta para que todos tengan fichas lo más completas posibles.

TAREA FINAL: EXPLOTACIÓN DE CORPUS EN UNA TRADUCCIÓN

INTRODUCCIÓN:

Esta tarea se plantea para ofrecer a los estudiantes la posibilidad de empezar a aplicar lo aprendido y adquirido en la traducción de una parte de un texto, para lo que tendrán que recorrer todo el camino desde la selección del o de los corpus hasta su análisis y explotación como recurso documental en la traducción de un texto.

Materiales

Textos

- Un texto en inglés género periodístico y tema financiero (Texto 6: "JSE expects gold futures to excite strong appetites" o Texto 7: "Britain must avoid Germany's mistake, says Feldstein")

Fichas

- Ficha de recogida de datos sobre el uso de corpus al traducir
- Ficha de autoevaluación
- Ficha de valoración de la unidad

Recursos electrónicos

- URL de un corpus de inglés general [*British National Corpus* –BNC: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- URL de un corpus de español general [CREA: <http://www.rae.es>]
- Corpus de textos periodísticos en inglés del *Financial Times*
- Corpus del de textos periodísticos financieros del *Expansión*
- Programa informático *WordSmith Tools*

DESARROLLO:

1. Se proporciona el TO (Texto 6 o Texto 7) en inglés a cada estudiante y se explica el encargo de traducción. El TO consiste en un texto de género periodístico y tema financiero como los que se han tratado a lo largo de la unidad; los alumnos tendrán que traducir algunos fragmentos seleccionados del TO.
2. Como único recurso de documentación para traducir se ofrece una serie de corpus tanto generales como especializados, tanto en lengua original como en lengua meta.
3. Al mismo tiempo que realizan la traducción, los alumnos han de rellenar la ficha de recogida de datos que se centra en el uso de corpus para la resolución de cinco problemas de traducción contenidos en el TO.
4. Al terminar la tarea, que supone también el fin de la unidad, se pasa a los alumnos las fichas de autoevaluación y de valoración de la unidad para que puntúen y comenten diversos aspectos de la misma.

Dinámica de grupos: individual

Duración estimada: 3 horas

EVALUACIÓN:

- Cada solución de cada uno de los 5 problemas seleccionados se valorará como correcta o incorrecta dependiendo de su adecuación al encargo y la función de la traducción. Para valorar el uso de corpus para traducir se utilizarán algunos de los indicadores diseñados en esta tesis.
- Los alumnos entregan su Diario de aprendizaje rellenado a lo largo de la unidad.

③ EVALUACIÓN DE LA UNIDAD DIDÁCTICA

1. Valoración realizada por los estudiantes

a) Valoración de su aprendizaje: ficha de autoevaluación

En este apartado se presentan los resultados obtenidos a partir de la Ficha de autoevaluación rellena por cada alumno al finalizar la unidad “Introducción al uso de corpus electrónicos para traducir”. Como puede observarse en la tabla que sigue (Tabla 18), las respuestas de puntuación personal de los alumnos sobre su aprendizaje y aprovechamiento de la unidad tratada revelan algunas tendencias que merecen ser comentadas.

- Puntuación media de las respuestas proporcionadas por cada alumno. Se observa un *continuum* que va desde el 6,8 hasta el 9,6 en la parte referente a las capacidades adquiridas por el alumno según su propia apreciación y notas que oscilan entre un 5,6 y un 9,6 en la parte referente a la utilidad de los corpus juzgada por los mismos sujetos. Aquí es necesario destacar los diferentes estilos de aprendizaje de los alumnos, junto a otras cuestiones que pueden afectar a la puntuación, como son la regularidad en la asistencia a las clases o la autoestima, punto éste último comprobado en la ficha de autoevaluación perteneciente a la alumna M.P., que otorgó las puntuaciones más bajas en la segunda parte de la ficha, añadiendo el comentario al margen de que era capaz de ver la utilidad de los corpus pero aún encontraba dificultades a la hora de explotarla.
- Puntuación media de las respuestas en cada pregunta. El interés de observar esta puntuación radica, para el primer bloque (“Soy capaz de captar...”), en la posibilidad de ver gráficamente en qué partes de la unidad didáctica los alumnos han ganado más o menos confianza. En este sentido, si nos fijamos en las puntuaciones extremas, se aprecia que los estudiantes se ven menos hábiles a la hora de, por ejemplo, interrogar un corpus (puntuación: 7), pero sí han entendido perfectamente la relación entre el tamaño y la utilidad de un corpus a la hora de extraer información (puntuación: 8,8). En cuanto al segundo bloque (“Me ha ayudado el corpus a...”), las puntuaciones medias nos permiten ver que los estudiantes no han asimilado totalmente cómo utilizar corpus a la hora de extraer equivalentes de traducción (puntuación: 5,7), pero sí se han dado perfecta y destacada

cuenta del valor de las colocaciones en la traducción (puntuación: 8,5). Como es lógico, aquellas cuestiones en las que los alumnos han dado una puntuación baja han de llevar necesariamente a una revisión de las tareas centradas en la consecución de los objetivos asociados con la finalidad de que éstos sean adquiridos, como se verá más adelante.

El segundo bloque de la ficha de autoevaluación incluye, aparte de un espacio para las puntuaciones numéricas, un espacio (“Otros comentarios”) para que los estudiantes añadan otros aspectos que consideran dignos de destacar en cuanto al uso de corpus electrónicos. En este sentido, los comentarios aportados por los alumnos han sido sobre todo con respecto a la utilidad de la información sobre frecuencia de uso, siempre interpretado en su contexto, así como información sobre colocaciones, usos cambiantes de la lengua (p.ej. el uso de “pero” a principio de frase), equivalentes de difícil acceso en otros recursos (p.ej. siglas como “JSE”), etc. Las respuestas completas pueden leerse en el *Volumen 2* de esta tesis.

FICHA DE AUTOEVALUACIÓN: RESULTADOS UNIDAD DIDÁCTICA "INTRODUCCIÓN AL USO DE CORPUS ELECTRÓNICOS PARA TRADUCIR"											
Estudiantes	A.G.	L.C.	M.M.	M.P.	P.A.	R.G.	S.L.	Media de cada pregunta			
Soy capaz de captar...	6	8	10	10	10	8	7	8,4			
	8	8	10	10	10	8	8	8,8			
	7	5	10	8	10	9	6	7,8			
	8	8	10	5	10	8	7	8			
	7	7	8	7	8	6	6	7			
	6	6	8	5	10	7	7	7			
	7	7	9,3	7,5	9,6	7,6	6,8				
	6	6	8	3	10	6	7	6,5			
	5	6	8	3	8	6	4	5,7			
	5	8	7	9	10	7	9	7,8			
Me ha ayudado el corpus a...	6	7	7	4	10	6	8	6,8			
	8	8	9	10	10	7	8	8,5			
	8	6	7	5	10	9	6	7,2			
	6,3	6,8	7,6	5,6	9,6	6,8	7				
	Media de cada alumno										
	traducir mejor										
	encontrar los equivalentes de traducción										
	estar más seguro de la solución										
	comprender mejor el texto original										
	comprender la importancia de las colocaciones (palabras que suelen acompañar a otras)										
a ser más consciente de la naturaleza de los textos utilizados en la fase de documentación (origen, fecha, original o traducción, etc.)											
Media de cada alumno											

Tabla 18. Resultados de la Ficha de autoevaluación de la UD "Introducción al uso de corpus electrónicos para traducir".

b) Valoración de la unidad didáctica: ficha de valoración

En la ficha de valoración de la unidad los estudiantes debían responder extensamente a preguntas sobre la unidad didáctica en general, los puntos que encontraban a favor y en contra, etc. A continuación se muestran los temas más destacados, si bien las respuestas completas se encuentran en el *Volumen 2*.

(i) Nueva forma de documentarse

En lo que respecta a esta ficha de valoración de la unidad, destaca el hecho de que la mayoría de los alumnos, si no todos, desconocían el uso de corpus electrónicos y las herramientas informáticas relacionadas. Es grato comprobar que varios estudiantes (p.ej. A.G., M.M., P.A.) consideran que han descubierto en los corpus una rica fuente de información que les puede ayudar a contextualizar términos, descubrir su significado y observar su uso con la ayuda de los datos de frecuencia.

(ii) Destacado positivamente por los alumnos

Lo innovador del enfoque, la orientación hacia el alumno, el repaso de los conceptos más importantes y el clima de cooperación entre profesora y alumnos son aspectos que los alumnos han destacado como positivos, así como la claridad en el proceso de llegar a soluciones. De especial interés para los alumnos parece haber sido el hecho de que las clases fueran eminentemente prácticas y que siempre se llevaran a cabo en un aula multimedia ya que esto les permitía desarrollar o practicar sus habilidades informáticas, así como aprender a utilizar de forma contextualizada el programa *WordSmith Tools*.

(iii) Sugerencias de mejora y comentarios

Entre las sugerencias, algunos alumnos manifiestan su deseo de realizar más traducciones, en otros campos temáticos, y de compilar ellos mismos su propio corpus e incluso trabajar con corpus bilingües que les faciliten la traducción de una forma más inmediata que a través de corpus comparables. En este sentido, lo único que se puede comentar es que la presente unidad didáctica está dedicada a la introducción al uso de corpus para traducir y está pensada para un nivel de iniciación a la traducción, razones por las cuales, por ejemplo, no se trataron más campos temáticos, ya que los intereses de esta unidad no son ni el trabajo con textos ni el estudio de un dominio, y las cuestiones temporales y prácticas no permiten trabajar todos los tipos de corpus existentes en una sola unidad, con lo que la compilación de corpus ad hoc y el análisis de corpus bilingües se practica en unidades posteriores.

2. Valoración realizada por el docente

a) Valoración del aprendizaje de los estudiantes

A continuación presentamos una tabla con las puntuaciones obtenidas por cada alumno tras haber analizado su traducción y su ficha de recogida de datos sobre el uso de corpus al traducir, así como las puntuaciones medias obtenidas para cada uno de los indicadores seleccionados.

FICHA DE CRUCE DE DATOS: RESUMEN DE RESULTADOS										
UNIDAD DIDÁCTICA: "INTRODUCCIÓN AL USO DE CORPUS ELECTRÓNICOS PARA TRADUCIR"										
Estudiantes	A.G.	C.P.	J.A.	M.M.	M.P.	P.A.	R.G.	S.L.	Media de cada indicador (tras aplicar el %)	Media de cada indicador (sobre 10 puntos)
Indicador 1: Selecciona el corpus idóneo (30%)	2,4	1,8	2,4	2,4	2,7	3	2,7	2,7	2,5/3	8,3
Indicador 2: Busca la palabra / cadena de palabras idónea (20%)	1,2	1,2	1,4	1,8	1,6	1,6	1,6	1,6	1,5/2	7,5
Indicador 3: Limita la búsqueda (búsqueda truncada, búsqueda con palabra de contexto) de forma idónea (20%)	2	1,4	0	1	1	2	0	2	1,2/2	6
Indicador 4: Ordena los resultados (20%)	1	1,6	0,6	1,2	1	2	0	1,6	1,1/2	5,5
Indicador 5: Usa las funciones del programa (10%)	0,8	1	0	1	1	1	0,5	1	0,8/1	8
Nota de cada alumno	7,4	7	4,4	7,4	7,3	9,6	4,8	8,9		

Tabla 19. Resumen de resultados de la Ficha de cruce de datos de la UD "Introducción al uso de corpus electrónicos para traducir".

- Puntuaciones medias de los alumnos. Destacan sobremanera en los extremos un 4,4 y un 9,6, es decir, la peor y mejor puntuaciones. Para comprender estos resultados se puede observar con detalle cada columna que, en el caso del resultado más óptimo (P.A.), refleja aciertos en todas las categorías en las que el alumno ha considerado que debía utilizar corpus para ayudarle con la traducción. En el otro extremo, la columna del resultado más negativo, se observa que el alumno (J.A.), por ejemplo, ha ignorado consistentemente las categorías 3 y 5, es decir, no ha refinado sus búsquedas ni utilizado el programa *WordSmith Tools* en ningún momento, lo que le ha conducido en ocasiones a resultados de traducción incorrectos, razón por la cual cuenta con varios “o” en su columna.
- Puntuaciones medias obtenidas por el conjunto de alumnos en cada indicador. Se observa claramente que éstos han captado bastante bien qué tipo de corpus necesitan consultar dependiendo del tipo de problema de traducción que se les plantea. Por otro lado, destaca la baja media obtenida en la categoría de reordenación de resultados ya que, como se aprecia, o ha sido ignorada o no se ha aplicado con la adecuación necesaria como para servir de ayuda a la hora de traducir. La observación de este aspecto ha de llevar necesariamente a una revisión de la tarea dedicada a tal objetivo de aprendizaje.

b) Valoración de la unidad didáctica

La presente unidad está entre las primeras que se diseñaron y probaron ya no como ejercicios sueltos, sino como una unidad didáctica completa. Si bien la valoración general es positiva, el análisis de los resultados extraídos tanto de las fichas de autoevaluación y valoración, como la de cruce de datos, ha llevado a realizar una revisión de algunas de las tareas que conforman la unidad. Así, los resultados más bajos observados en dichas fichas se han tomado como índice de los elementos susceptibles de modificación.

④ EVOLUCIÓN

Modificaciones efectuadas

Éstos son los principales cambios que se realizaron en Fichas o Tareas tras la puesta en práctica de la unidad :

- Tarea 1, Ficha 1 (“Qué contiene cada recurso”): Se observó la necesidad de redactar de forma más clara los enunciados.
- Tarea 3, Ficha 3 (“Información extraída de una forma diferente”): Esta tarea, en la que se evaluaba el indicador relacionado con la reordenación del contexto en las concordancias obtuvo la puntuación más baja en la Ficha de cruce de datos (Tabla 19). La Ficha 3, cuya versión mejorada es la que se ha incluido ya en esta tesis, se modificó para incluir preguntas explícitas sobre la ordenación de los resultados con el fin de fijar tal concepto.
- Tarea 6, Ficha 6 (“Preguntas al corpus”): La observación de las puntuaciones más bajas en los resultados obtenidos de la Ficha de autoevaluación apunta hacia tres aspectos relacionados entre sí directamente. En conjunto, la capacidad de interrogar un corpus para encontrar los equivalentes de traducción o, en general, sacarle partido para traducir viene a representar un punto débil del uso de corpus desde la propia perspectiva de los alumnos. En este caso, la Ficha 6, cuyo objetivo es captar la forma como se pueden dirigir las preguntas a un corpus para poder extraer resultados interpretables, ha sido el objeto de cambios. La propuesta de mejora se basa en una mayor práctica con el programa *WordSmith Tools* en este caso, con otros corpus y otras palabras de búsqueda para que el alumno trabaje la extracción de información.



UNIDAD DIDÁCTICA

Contrastes de uso en un corpus paralelo



En esta unidad se introduce al estudiante en el uso de un corpus paralelo ya creado con el fin de que, por un lado, aprenda a manejar software que le permita aprovechar un corpus de este tipo, bilingüe y alineado, y que se dé cuenta de las ventajas que ofrece su utilización. Por otro lado, gracias a este recurso, se trabajan cuestiones contrastivas entre el inglés y el español, como son elementos de discrepancia en cuestiones de puntuación, deixis o morfosintaxis.



Grupos de prueba: La unidad original se probó en un grupo de 8 alumnos de 4º curso de la licenciatura de Traducción e Interpretación de la UAB, en una asignatura de traducción (inglés C-español), durante el año académico 2002-03.

Los materiales de esta unidad se encuentran en el *Volumen 2* de esta tesis (págs. 457-484).



❶ PRESENTACIÓN DE LA UNIDAD DIDÁCTICA

“Contrastes de uso en un corpus paralelo”

NIVEL DE USO DE CORPUS

- Iniciación

MATERIA

- Iniciación a la traducción directa (inglés-español)

OBJETIVOS DE LA UNIDAD

Objetivos relacionados con el uso de corpus

- Identificar los principios metodológicos básicos del trabajo con corpus
- Utilizar las funciones básicas de un programa informático de alineamiento de textos
- Integrar conocimientos para resolver problemas de traducción utilizando corpus a nivel básico

Objetivos relacionados con la materia

- Dominar las diferencias de las convenciones de escritura (uso del guión)
- Dominar las diferencias en cuanto a los mecanismos de coherencia y cohesión (uso del demostrativo)
- Dominar los elementos de discrepancia morfosintáctica (uso de la voz pasiva)

Objetivos relacionados con competencias genéricas

- Trabajar en equipo

CONTENIDOS

- Ventajas y desventajas en la construcción de un corpus (paralelo) (I. 2.1.1.)
- Software comercializado de análisis de corpus (*MultiConcord*) (III. 1.10.2.)
- Función *Search* (de *MultiConcord*) (III. 2.1.5.)
- Discrepancia en convenciones de escritura (IV. 2.4.1.)
- Discrepancia en mecanismos de coherencia y cohesión (IV. 2.4.2.)
- Discrepancia morfosintáctica (IV. 2.4.3.)

ESTRUCTURA

- Tarea 1: “De qué y cómo está hecho un corpus paralelo”
- Tarea 2: “Uso de la raya o guión largo en inglés y español”
- Tarea 3: “Búsqueda de equivalentes para ‘this’”
- Tarea 4: “Búsqueda de equivalentes para la voz pasiva”
- Tarea final: “Encontrar equivalentes para diversos elementos de discrepancia”



EVALUACIÓN

Indicadores

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación,...

- Indicador 1: alinea textos automáticamente
- Indicador 2: extrae concordancias bilingües
- Indicador 3: busca equivalentes a elementos de discrepancia entre el inglés y el español

Fichas de evaluación

- Ficha de autoevaluación
- Ficha de valoración de la unidad
- Diario de aprendizaje

MATERIALES

Fichas

- Ficha 1: "Puntos a favor y en contra de un corpus paralelo"
- Ficha 2: "Uso de la raya o guión largo"
- Ficha 3: "Equivalentes de 'this'"
- Ficha 4: "Equivalentes de la voz pasiva"
- Ficha 5: "Ejercicio de rellenar huecos"
- Ficha de autoevaluación
- Ficha de valoración de la unidad didáctica

Texto original y traducido

- Textos 1 y 2 (debates en el Parlamento Europeo el 24/10/96 en inglés y español, respectivamente, marcados para ser procesados por el programa *MultiConcord*)

Recursos electrónicos

- URL de corpus de inglés general [BNC: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- URL de corpus de español general [*Corpus del español*: <http://www.corpusdelespanol.org>]
- URL de un conjugador de verbos [*Verbix*: <http://www.verbix.com/webverbix/index.asp>]
- Programa informático *MultiConcord*: alineador de textos, extractor de concordancias bilingües y generador de ejercicios de rellenar huecos

DURACIÓN ESTIMADA

8 horas

2 DESARROLLO DE LAS TAREAS

TAREA 1: DE QUÉ Y CÓMO ESTÁ HECHO UN CORPUS PARALELO

INTRODUCCIÓN:

Esta primera tarea tiene la finalidad de que el alumno capte la idea de lo que es un corpus paralelo y lo que implica su construcción, sus ventajas y desventajas. También busca que el alumno empiece a manejar las funciones básicas del programa de alineación de corpus *MultiConcord*.

DESARROLLO:

1. Se facilita a los alumnos un modelo de Diario de aprendizaje para que recojan todo aquello que consideren que van aprendiendo o adquiriendo a medida que realizan la unidad.
2. Se explica qué es un corpus paralelo, llamando la atención sobre las diferencias entre un corpus comparable y un corpus paralelo y se aclaran las dudas terminológicas respecto a la diferencia entre "texto paralelo" y "corpus paralelo".
3. Se muestran los textos que se van a alinear (Textos 1 y 2, que son debates llevados a cabo en el Parlamento Europeo el 24/10/96 en inglés y español, respectivamente, marcados para ser procesados por el programa *MultiConcord*), primero en formato TXT, y después una vez procesados y alineados con el programa *MultiConcord* para que aprecien qué es lo que compone un corpus paralelo y lo que implica la alineación de textos (p.ej. criterios de alineación según puntuación).
4. Para poder mostrar los textos alineados, se explican brevemente las funciones básicas del programa:
 - Selección de archivos.
 - Búsqueda de una palabra.
 - Visualización de las concordancias bilingües.
5. Se propone alguna palabra de búsqueda relevante y, observando las concordancias alineadas, los alumnos comentan en voz alta las características generales del texto.
6. Rellenan la Ficha 1 ("Puntos a favor y en contra de un corpus paralelo") con las dificultades aparentes que implica la construcción de un corpus paralelo y las ventajas que aporta.
7. El resultado se comenta en voz alta entre todos.

Dinámica de grupos: en grupos de 3 ó 4 personas

Duración estimada: 1 hora

MATERIALES:

Fichas

- Ficha 1: "Puntos a favor y en contra de un corpus paralelo"
- Diario de aprendizaje

Texto original y traducido

- Textos 1 y 2 (debates en el Parlamento Europeo el 24/10/96 en inglés y español, respectivamente, marcados para ser procesados por el programa *MultiConcord*)

Recursos electrónicos

- Programa informático *MultiConcord*: alineador de textos, extractor de concordancias bilingües y generador de ejercicios de rellenar huecos

COMENTARIOS:

- A lo largo de toda la unidad se utilizarán los mismos textos (1 y 2) provenientes de discursos llevados a cabo en el seno del Parlamento Europeo. Es recomendable escoger textos adecuados para un nivel de iniciación a la traducción, es decir, de tema o género no "especializado", entendido en el sentido tradicional, si bien es cierto que siempre habrá de comentarse con qué tipo y género de textos se va a trabajar y en qué contexto fueron producidos (en este caso, discursos en la UE).
- Esta tarea se realiza entre todos los estudiantes, en voz alta y guiados de cerca por el docente porque es probablemente la primera vez que ven un corpus paralelo y es previsible que no sean capaces de extraer conclusiones por sí solos.

TAREA 2: USO DE LA RAYA O GUIÓN LARGO EN INGLÉS Y ESPAÑOL**INTRODUCCIÓN:**

Esta tarea busca practicar la observación del uso de un elemento de discrepancia entre el inglés y el español, la raya o guión largo. Para ello, se observa este signo de puntuación en corpus monolingües de inglés y español general y después se ve cómo se ha traducido del inglés al español en el corpus paralelo de textos de la UE.

DESARROLLO:

1. Se facilita la Ficha 2 ("Uso de la raya o guión largo") que, para rellenarla, los alumnos tendrán que buscar rayas o guiones largos en distintos corpus.
2. Los alumnos buscan en el corpus *BNC* y en el *Corpus del Español* y se atrae la atención hacia el hecho de cómo tienen que formular la búsqueda: "-", es decir, la "sintaxis de búsqueda", aspecto en el que se incidirá para cada herramienta y corpus. Rellenan la parte de la ficha correspondiente al uso del guión largo en el *BNC* y en el *Corpus del Español*.
3. Buscan en el corpus paralelo de la UE (Textos 1 y 2 alineados) con la ayuda del *MultiConcord*: "-". La búsqueda se realiza en la parte de los textos en inglés, que se ha seleccionado como TO. Debido a la numerosa cuantía de casos que aparecen, éstos se reparten entre los alumnos para que cada uno se encargue de observar sólo una cantidad determinada y así agilizar el trabajo. Rellenan la parte de la ficha correspondiente al uso

del gui3n largo en los textos en ingl3s del corpus paralelo para lo que tendr3n que observar los casos asignados y explicitar la funci3n que cumple el gui3n en cada uno de ellos.

4. Observan los p3rrafos equivalentes en espa3ol y se fijan en c3mo se ha transferido la funci3n de la raya del ingl3s al espa3ol (en este caso ser3n siempre comas) y cu3l es esa funci3n (inciso, comentario al margen, etc.). Rellenan la parte de la ficha restante.
5. Tras comentar los resultados de la ficha entre todos, los alumnos fijan por escrito sus conclusiones al final de la misma.

Din3mica de grupos: individual

Duraci3n estimada: 1 hora y media

MATERIALES:

Fichas

- Ficha 2: "Uso de la raya o gui3n largo"

Texto original y traducido

- Textos 1 y 2 (debates en el Parlamento Europeo el 24/10/96 en ingl3s y espa3ol)

Recursos electr3nicos

- URL de corpus de ingl3s general [BNC: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- URL de corpus de espa3ol general [*Corpus del espa3ol*: <http://www.corpusdelespanol.org>]
- Programa inform3tico *MultiConcord*

EVALUACI3N:

Dado que, a pesar de que los casos sean numerosos, tambi3n son repetitivos en sus funciones, los resultados de la ficha se puede comentar y discutir en voz alta. As3, aunque cada alumno se haya centrado en un n3mero de casos, todos podr3n tener una percepci3n general de los resultados.

COMENTARIOS:

- Es necesario insistir en la importancia de conocer la sintaxis de b3squeda de cada corpus para sacarle el m3ximo provecho.

TAREA 3: BÚSQUEDA DE EQUIVALENTES PARA "THIS"

INTRODUCCIÓN:

Esta tarea busca practicar la observación del uso de un elemento de discrepancia entre el inglés y el español, el pronombre/adjetivo demostrativo "this".

DESARROLLO:

1. Se pregunta en voz alta de qué maneras se puede traducir "this" al español. Se comentan brevemente las soluciones propuestas (p.ej. este/a/o) y las diferencias, por ejemplo, de categoría gramatical (adjetivo, pronombre) y sintáctica (modificador, sujeto) que puede representar.
2. Los alumnos buscan la palabra "this" en el corpus bilingüe de la UE (Textos 1 y 2 alineados) con la ayuda del *MultiConcord*.
3. A pesar de encontrar una gran cantidad de casos de "this", el programa sólo ofrece 250 como máximo. Se reparten todos los casos encontrados entre los alumnos.
4. Apuntan en la Ficha 3 ("Equivalentes de 'this'") los datos encontrados en su "paquete" de casos.
5. Cada alumno escribe en la Ficha 3 los equivalentes que ha encontrado para "this" en español en su "paquete" de casos, apuntando los números de los ejemplos.
6. Dado que cada alumno sólo recoge los datos correspondientes a un número determinado de casos, pero no de todos, se establece una cadena de trabajo entre los estudiantes para que entre todos lleguen a una ficha única que contenga toda la información. La cadena consiste en que cada alumno facilita su ficha a otro alumno, que fundirá sus datos con los de la ficha que reciba produciendo una ficha más completa que a su vez pasará a otro alumno, que seguirá el mismo proceso para llegar a una ficha que contenga más datos cada vez. El último alumno en recibir la ficha deberá completarla con sus propios datos y entonces pasarla a todos los estudiantes para que posean la misma información.

Dinámica de grupos: individual y en cadena

Duración estimada: 2 horas

MATERIALES:

Fichas

- Ficha 3: "Equivalentes de 'this'"

Texto original y traducido

- Textos 1 y 2 (debates en el Parlamento Europeo el 24/10/96 en inglés y español)

Recursos electrónicos

- Programa informático *MultiConcord*

EVALUACIÓN:

Tras observar la última ficha con los datos de todos recogidos, cada alumno redactará unas conclusiones en la misma ficha sobre los equivalentes del "this" en español. También pueden comentarse algunos ejemplos en voz alta si es necesario.

COMENTARIOS:

- Se espera que, en un principio, a los alumnos sólo se les ocurran dos o tres equivalentes para "this"; con este ejercicio verán que este pronombre/adjetivo puede traducirse de muchas formas, o incluso no traducirse en muchos casos, siempre dependiendo del contexto.
- El ejercicio en cadena hace hincapié en la responsabilidad compartida sobre un trabajo final, lo que refleja una parte de la realidad profesional. Cada alumno es responsable de su parte de los datos, de integrarlos con los de los demás, y de pasarlos a tiempo a otro compañero para que éste pueda seguir trabajando.
- Si se desea ampliar o reforzar las habilidades que practica esta tarea se puede sugerir, por ejemplo, tratar la riqueza de formas posibles en las que se puede expresar el agradecimiento en español, haciendo búsquedas del tipo "thank*" en el corpus paralelo.

TAREA 4: BÚSQUEDA DE EQUIVALENTES PARA LA VOZ PASIVA

INTRODUCCIÓN:

Esta tarea busca practicar la observación del uso de un elemento de discrepancia entre el inglés y el español, la voz pasiva. Incide también en cómo plantear las búsquedas en un corpus, es decir, pensar en la cadena de caracteres que, al introducirla como búsqueda en un corpus, ofrezca el máximo número de resultados relevantes.

DESARROLLO:

1. Por parejas, los estudiantes van a observar las distintas formas de equivalencia que puede tener en español la voz pasiva, mucho más frecuente en inglés.
2. Ahora que los estudiantes ya están familiarizados con los textos, el software y han realizado varias búsquedas diferentes, se les plantea la cuestión de cómo habría de ser la sintaxis de búsqueda para poder extraer construcciones pasivas del corpus bilingüe que tienen a su disposición (Textos 1 y 2 alineados).
3. Buscan uno de los patrones propuestos, por ejemplo, "was *ed", en el corpus bilingüe con *MultiConcord*.
4. Cada pareja rellena la Ficha 4 ("Equivalentes de la voz pasiva") con los datos extraídos.
5. Para facilitar la identificación de los tiempos verbales con sus nombres correctos se les da la URL de un conjugador verbal en línea.

Dinámica de grupos: en parejas

Duración estimada: 1 hora y media

MATERIALES:

Fichas

- Ficha 4: "Equivalentes de la voz pasiva"

Textos original y traducido

- Textos 1 y 2 (debates en el Parlamento Europeo el 24/10/96 en inglés y español)

Recursos electrónicos

- URL de un conjugador de verbos [*Verbix*: <http://www.verbix.com/webverbix/index.asp>]
- Programa informático *MultiConcord*

EVALUACIÓN:

Se pide que cada pareja comente sus resultados con otra pareja y que después, ya individualmente, cada alumno rellene una ficha similar introduciendo otro patrón de búsqueda, por ejemplo, "were *ed". Se pueden comentar algunos casos en voz alta.

COMENTARIOS:

- Se prevé que, en un principio, los alumnos no piensen en muchas maneras diferentes de reexpresar la voz pasiva inglesa en español y que al final se den cuenta de que hay formas muy diversas.

TAREA FINAL: ENCONTRAR EQUIVALENTES PARA DIVERSOS ELEMENTOS DE DISCREPANCIA

INTRODUCCIÓN:

En esta tarea se practican los conocimientos y habilidades contrastivos adquiridos a lo largo de la unidad para encontrar equivalentes a elementos de discrepancia de diverso tipo ya tratados.

DESARROLLO:

1. Se facilita la Ficha 5 ("Ejercicio de rellenar huecos"), que es un ejercicio generado a partir de los textos trabajados en la unidad (Textos 1 y 2) y el programa *MultiConcord*. En el ejercicio se observan oraciones en inglés donde se ha utilizado el guión largo, el "this" o la voz pasiva, y el estudiante debe rellenar con un equivalente los espacios en blanco que aparecen en las oraciones en español.
2. Al finalizar, los alumnos rellenan la ficha de autoevaluación y la ficha de valoración de la unidad didáctica.

MATERIALES:

Fichas

- Ficha 5: "Ejercicio de rellenar huecos"
- Ficha de autoevaluación
- Ficha de valoración de la unidad didáctica

Texto original y traducido

- Textos 1 y 2 (debates en el Parlamento Europeo el 24/10/96 en inglés y español)

Recursos electrónicos

- Programa informático *MultiConcord*

EVALUACIÓN:

- El docente valora cada hueco rellenado de la Ficha 5 según si el equivalente propuesto por el alumno es correcto (2 puntos), mejorable (1 punto) o incorrecto (0 puntos) y realiza la media aritmética.
- Los alumnos entregan su Diario de aprendizaje rellenado a lo largo de la unidad.

Dinámica de grupo: individual

Duración estimada: 2 horas

COMENTARIOS:

- Se puede resaltar la relevancia de comprobar en corpus más grandes y generales los datos obtenidos en el corpus de debates de la UE.

③ EVALUACIÓN DE LA UNIDAD DIDÁCTICA

1. Valoración realizada por los estudiantes

a) Valoración de su aprendizaje: ficha de autoevaluación

Se presentan aquí los resultados obtenidos a partir de la Ficha de autoevaluación rellena por cada alumno al finalizar la unidad. En la tabla que sigue (Tabla 20) pueden observarse las pruebas de algunas tendencias observadas.

- Puntuación media de las respuestas proporcionadas por cada alumno. La observación de las medias de las puntuaciones otorgadas por los alumnos revela patrones consistentes como por ejemplo que los resultados más bajos y más altos provienen siempre de los mismos alumnos. Este hecho, que se repite en la evaluación llevada a cabo en la unidad anterior, “Introducción al uso de corpus electrónicos para traducir”, apunta una tendencia que podría estar relacionada con gustos personales, como ya señalamos en su momento.
- Puntuación media de las respuestas en cada pregunta. En cuanto a las medias máximas y mínimas que se pueden observar respecto a cada pregunta destacan, tanto para el primer bloque como para el segundo, una puntuación más baja en la pregunta más genérica y puntuaciones más altas en preguntas concretas, si bien es cierto que la diferencia en este caso no es muy significativa.

Según parece a razón de las respuestas al apartado “Otros comentarios” de la Ficha de autoevaluación, el trabajo de contrastividad ha tenido un efecto muy positivo en la observación por parte de los alumnos del contexto lingüístico y extralingüístico a la hora de pensar en posibles equivalentes. Se constata que los alumnos han apreciado la utilidad que tiene un corpus paralelo, especialmente en el sentido de que este tipo de corpus permite encontrar equivalentes de traducción de una forma mucho más rápida que un corpus comparable, sea éste bilingüe o monolingüe en la lengua de llegada.

FICHA DE AUTOEVALUACIÓN: RESULTADOS												
UNIDAD DIDÁCTICA "CONTRASTES DE USO EN UN CORPUS PARALELO"												
Estudiantes												
	A.G.	C.P.	J.A.	L.C.	M.M.	M.P.	P.A.	R.G.	S.L.	Media de cada pregunta		
Soy capaz de captar...	las ventajas y desventajas que ofrece un corpus paralelo	8	8	6	7	8	10	8	6	7,6		
	el uso del guión largo o raya en inglés y en español	7	9	7	9	9	10	7	7	8,1		
	la variedad de equivalentes que puede tener la raya en español	7	9	6	10	10	10	7	7	8,1		
	el uso del "this" en inglés y sus equivalentes en español	8	8	7	9	8	10	8	6	8		
	la variedad de equivalentes que puede tener "this" en español	8	9	8	7	10	8	10	8	8,3		
	el uso de la voz pasiva en inglés y español	8	8	8	7	9	9	10	6	7	8	
	la variedad de equivalentes que puede tener la voz pasiva en español	8	8	7	7	10	10	6	6	7,7		
	Media de cada alumno	7,7	8,4	7,7	6,7	9,1	8,6	10	7,1	6,5		
	Te ha ayudado el corpus a ...	traducir mejor	7	8	6	6	7	6	10	9	7,4	
		encontrar equivalentes o soluciones para elementos de discrepancia en convenciones de escritura entre el inglés y el español (guión largo o raya)	7	9	8	7	8	8	8	6	7,6	
encontrar equivalentes o soluciones para elementos de discrepancia morfosintáctica entre el inglés y el español (<i>this</i>)		8	8	10	7	8	9	8	7	8,2		
encontrar equivalentes o soluciones para elementos de discrepancia morfosintáctica entre el inglés y el español (voz pasiva)		8	9	8	7	8	9	8	6	7,6		
observar diferencias entre el inglés y el español		9	8	8	7	8	8	10	9	8,3		
pensar en más de una posible equivalencia para un elemento del TO		8	7	8	7	9	8	10	9	8		
comprender la importancia del contexto (lingüístico, extralingüístico) del TO a la hora de pensar en equivalencias		8	9	8	7	10	8	10	10	6	8,4	
Media de cada alumno		7,8	8,2	8	6,8	8,2	8	9,1	8,5	6,7		

Tabla 20. Resultados de la Ficha de autoevaluación de la UD "Contrastes de uso en un corpus paralelo".

b) Valoración de la unidad didáctica: ficha de valoración

Recordamos que las respuestas completas a esta ficha se encuentran en el *Volumen 2*.

(i) Nueva forma de documentarse

Según las respuestas a la Ficha de valoración, los estudiantes insisten en los aspectos positivos del uso de corpus paralelos en cuanto a que les permite ver distintos equivalentes de traducción para un mismo término o estructura, cotejar estos equivalentes con los suyos propios, extraer conclusiones, etc. aprendiendo de una forma práctica.

(ii) Destacado positivamente por los alumnos

Entre las tareas llevadas a cabo destacan especialmente las dedicadas a los equivalentes del “this” y de la voz pasiva por lo reveladoras que han resultado. Los estudiantes subrayan también lo conveniente que es conocer recursos y herramientas que desconocían hasta el momento y lo interesante que es trabajar en grupo. En este sentido, el trabajo en cadena llevado a cabo en una tarea parece haber añadido un interés especial no sólo en cuanto a la agilización del trabajo, sino también porque ha permitido una observación y evaluación del trabajo de otros compañeros.

(iii) Sugerencias de mejora y comentarios

Las sugerencias de cambio o mejora aportadas por los estudiantes para la presente unidad van en la línea de un cambio de género textual y en la idea de construcción de un corpus propio. Podemos comentar al respecto que, en primer lugar, por tratarse de una unidad diseñada para un nivel de iniciación a la traducción, ésta no está centrada en el trabajo con textos o el estudio de géneros determinados y que, en segundo lugar, la construcción de un corpus paralelo no se consideró como objetivo de esta unidad por estar ésta centrada exclusivamente en cuestiones contrastivas de la traducción. La construcción de un corpus sí se contempla, sin embargo, en otras unidades didácticas que pertenecen a la fase pedagógica de consolidación en el uso de corpus para traducir.

2. Valoración realizada por el docente

a) Valoración del aprendizaje de los estudiantes

Todas las herramientas utilizadas para hacer el seguimiento y evaluación de la unidad revelan que ésta se ha seguido con normalidad por los alumnos e, incluso, con mayor

entusiasmo que la unidad anterior, la de introducción al uso de corpus para traducir, especialmente por ver en el corpus paralelo una forma más “directa” de encontrar equivalentes de traducción que en un corpus comparable.

En esta unidad dedicada a cuestiones contrastivas no se planteó una traducción como tarea final. En lugar de una traducción, se decidió que la forma de evaluar lo adquirido a lo largo de la unidad sería con un ejercicio de rellenar huecos generado con el programa *MultiConcord* a partir del corpus paralelo con el que los alumnos habían estado trabajando. Así, en este caso la notación se limitó a una jerarquía de 0, 1 ó 2 puntos según si la respuesta del alumno era incorrecta, mejorable o correcta. Las puntuaciones medias, calculadas sobre 32 casos de contrastividad, muestran unos muy buenos resultados en cuanto a las soluciones de los alumnos en el ejercicio de rellenar huecos, como muestra la siguiente tabla.

FICHA 5: EJERCICIO DE RELLENAR HUECOS: RESULTADOS									
UNIDAD DIDÁCTICA “CONTRASTES DE USO EN UN CORPUS PARALELO”									
Estudiantes	A.G.	C.P.	J.A.	L.C.	M.M.	M.P.	P.A.	R.G.	S.L.
Notas sobre 10	8,2	8,4	7,3	8,4	8,1	8,7	10	8,2	8,5

Tabla 21. Notas medias resultantes de puntuar la Ficha 5 de la UD “Contrastes de uso en un corpus paralelo”.

b) Valoración de la unidad didáctica

Si relacionamos los resultados mostrados en la última tabla con los comentarios de los alumnos en la Ficha de autoevaluación y la observación por parte de la profesora, podemos concluir que la unidad “Contrastes de uso en un corpus paralelo” es del todo adecuada para el nivel y objetivos propuestos.

④ EVOLUCIÓN DE LA UNIDAD DIDÁCTICA

Modificaciones efectuadas

Tras la puesta en práctica y evaluación de la unidad se realizaron modificaciones mínimas en los enunciados de alguna ficha.



UNIDAD DIDÁCTICA

Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos



Esta unidad didáctica tiene como elemento original el que el alumno comienza su proceso de documentación sin partir siquiera de un texto original concreto, sino que primero localiza las fuentes de información que le podrán servir de ayuda e identifica las características del género textual que se trata en la unidad, la descripción técnica de pequeños electrodomésticos (microondas) en Internet, para lanzarse a construir un corpus comparable bilingüe en el que basará su búsqueda de equivalencias.



Grupos de prueba: Los materiales y datos mostrados en esta unidad didáctica corresponden a la que se probó en su forma completa en el curso 2006-07 y 2007-08 con alumnos de la asignatura Traducción directa especializada (inglés B-español) de 3º de licenciatura en Traducción e Interpretación de la UAB. Antes, esta unidad se había probado por primera vez en forma de ejercicios independientes en el curso 2000-01 y, más adelante, se probó una versión intermedia en el curso 2005-06.

Los materiales de esta unidad se encuentran en el *Volumen 2* de esta tesis (págs. 485-554).



❶ PRESENTACIÓN DE LA UNIDAD DIDÁCTICA

“Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”

NIVEL DE USO DE CORPUS

- Consolidación

MATERIA

- Traducción técnica directa (inglés B - español A)

OBJETIVOS

Objetivos relacionados con el uso de corpus

- Conocer fuentes de donde extraer textos
- Crear un corpus comparable bilingüe
- Utilizar software y recursos de ubicación y descarga de textos electrónicos desde Internet
- Utilizar funciones avanzadas de un programa de análisis de corpus
- Integrar conocimientos para resolver problemas de traducción utilizando corpus a nivel avanzado

Objetivos relacionados con la materia

- Dominar las herramientas del traductor técnico y científico
- Captar la importancia de la documentación y las estrategias para adquirirla
- Saber identificar y caracterizar distintos tipos y géneros textuales en las dos lenguas
- Saber traducir documentos industriales de uso externo (descripción técnica)

Objetivos relacionados con competencias genéricas

- Comunicarse con expertos de otras áreas

CONTENIDOS

- Textos reales e implicaciones (I. 2.1.2.)
- Importancia de nombrar lógicamente los archivos de un corpus (II. 2.1.)
- Importancia de ordenar lógicamente los archivos de un corpus (II. 2.2.)
- Creación de un corpus comparable bilingüe (ad hoc) (II. 3.1.2.)
- Software de descarga de textos (*Free Download Manager*) (III. 1.3.2.)
- Software gratuito de análisis de corpus (*InfoRapid*) (III. 1.10.1.)
- Software comercializado de análisis de corpus (*WordSmith Tools*) (III. 1.10.2.)
- Función *Concord* (de *WordSmith Tools*) (III. 2.1.1.)
- Función *Search* (de *InfoRapid Search & Replace*) (III. 2.1.6.)
- Función *Clusters* (de *WordSmith Tools*) (III. 2.5.1.)
- Listas de palabras por orden de frecuencia (III. 2.7.2.)
- Extracción de vocabulario general y términos de especialidad (IV. 2.2.1.)
- Inspiración para la traducción creativa (IV. 2.2.5.)



ESTRUCTURA

- Tarea 1: "Encontrar fuentes y recursos"
- Tarea 2: "Identificar el género textual"
- Tarea 3: "Compilación de un corpus comparable bilingüe"
- Tarea 4: "Terminología diferente y terminología compartida"
- Tarea 5: "Identificación de la fiabilidad de los textos o la importancia de ser ordenado"
- Tarea 6: "Creatividad en un corpus técnico"
- Tarea 7: "Aprovechar las pistas que da el corpus en lengua original (LO)"
- Tarea final: "Encontrar equivalencias utilizando el recurso creado"

EVALUACIÓN

Indicadores

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación,...

- Indicador 1: hace una búsqueda de calidad y económica en el corpus
- Indicador 2: justifica correctamente su opción de traducción basándose en datos extraídos del corpus

Fichas de evaluación

- Ficha de evaluación del uso de corpus al traducir
- Ficha de autoevaluación
- Ficha de valoración de la unidad
- Diario de aprendizaje

MATERIALES

Fichas

- Ficha 0: "*Brainstorming*"
- Ficha 1: "Información sobre empresas que fabriquen o comercialicen microondas"
- Ficha 2: "Características del género de la descripción técnica de microondas en Internet, en inglés y en español"
- Ficha 3: "Términos diferentes y términos en común"
- Ficha 4: "Errores en el corpus de español"
- Ficha 5: "Creatividad en los textos técnicos"
- Ficha 6a: "Cuando el diccionario no ayuda con la terminología..."
- Ficha 6b: "...La clave puede estar en el cotexto de unas concordancias"
- Ficha 6c: "...La clave puede estar en el cotexto de unas concordancias"
- Ficha de evaluación del uso de corpus al traducir
- Ficha de autoevaluación
- Ficha de valoración de la unidad

Textos originales y paralelos

- Texto 1: "Panasonic NN-T888S"
- Texto 2: "Fagor MW3-176G"
- Texto 3 (expositivo): "Descripción técnica de una lavadora, en inglés"
- Texto 4 (expositivo): "Descripción técnica de una impresora, en inglés"



- Texto 5 (expositivo): "Catálogo de materiales según las normas que se emplean en la industria petroquímica, en inglés"
- Texto 6 (expositivo): "Descripción técnica de una mesa de mezclas, en español"
- Texto 7 (expositivo): "Lista de materiales, en español"
- Texto 8 (expositivo con función secundaria instructiva): "Prospecto de un medicamento, en español"
- Texto 9 (instructivo): "Manual de instrucciones de un teléfono móvil, en español"
- Texto 10 (instructivo con función secundaria expositiva): "Publirreportaje sobre tuberías y accesorios, en español"
- Textos 11, 12: descripciones técnicas de microondas, en español
- Textos 13, 14: descripciones técnicas de microondas, en inglés
- Texto 15: "Snapple-crapple"
- Texto 16: "Siemens HF17065GB"
- Texto 17: "Roper MHE14XMD"

Textos de apoyo

- 1: Material de apoyo al programa *WordSmith Tools*

Descripciones de corpus

- Descripción 1: corpus ad hoc de descripciones técnicas de microondas en inglés y español extraídas de Internet

Concordancias

- Concordancia 1: "cocción"

Listas

- Lista 1: "Lista de frecuencias en el corpus de español"

Recursos electrónicos

- URL de conversor de cantidades en línea [*Ask Numbers*: <http://www.asknumbers.com/>]
- URL de buscador de imágenes [*Google Imágenes*: <http://images.google.es/>]
- Corpus ad hoc de descripciones técnicas microondas compilado
- Programa informático *Free Download Manager*
- Programa informático *InfoRapid*
- Programa informático *WordSmith Tools*

DURACIÓN ESTIMADA

16 horas

2 DESARROLLO DE LA UNIDAD DIDÁCTICA

TAREA 1: ENCONTRAR FUENTES Y RECURSOS

INTRODUCCIÓN:

Esta tarea representa una primera aproximación de los alumnos al sector comercial al que pertenecen los textos que tratarán en la unidad. La documentación para la traducción comienza por la búsqueda de fuentes de donde recoger textos paralelos y no por el uso de recursos lexicográficos, aspecto que, más adelante, a la hora de buscar equivalencias, descubrirán como muy beneficioso.

DESARROLLO:

1. Se facilita a los alumnos un modelo de Diario de aprendizaje para que recojan todo aquello que consideren que van aprendiendo o adquiriendo a medida que realizan la unidad.
2. Se facilitan a los alumnos dos textos originales extraídos de Internet que contengan una descripción técnica de un microondas, uno en inglés (Texto 1: "Panasonic NN-T888S"), y otro en español (Texto 2: "Fagor MW3-176G"), para que hagan una lectura rápida (15 minutos para los dos).
3. Se facilita la Ficha 0 ("*Brainstorming*") para que, con la colaboración de todos, al final se consiga tener un listado de empresas del sector.
4. Se asigna una empresa a cada estudiante y se les entrega la Ficha 1 ("Información sobre empresas que fabriquen o comercialicen microondas"); cada uno tendrá que investigar si la empresa que le ha sido asignada es española o es extranjera con filiales o sedes en España y si los documentos que aparecen en su web son originales o traducidos. Para ello deberá contactar con las empresas por correo electrónico.

Dinámica de grupos: individual

Duración estimada: 1 hora y media

MATERIALES:

Fichas

- Ficha 0: "*Brainstorming*"
- Ficha 1: "Información sobre empresas que fabriquen o comercialicen microondas"
- Diario de aprendizaje

Textos originales y paralelos

- Texto 1: "Panasonic NN-T888S"
- Texto 2: "Fagor MW3-176G"

EVALUACIÓN:

Una vez hayan conseguido la información para completar la Ficha 1, ésta se pondrá en común para que todos cuenten con la misma información. El docente se encargará de distribuir a todos la ficha completa con los datos de todas las empresas investigadas.

COMENTARIOS:

- Dependiendo del número de alumnos que haya y las empresas encontradas, en esta tarea se puede trabajar por parejas.
- Quizás se puede dedicar un poco de tiempo a comentar en voz alta el tipo de fuentes que se han encontrado (páginas web de las propias empresas, de minoristas, de distribuidores, etc.).

TAREA 2: IDENTIFICAR EL GÉNERO TEXTUAL**INTRODUCCIÓN:**

Esta tarea se centra en la identificación del género de la descripción técnica de microondas en Internet en inglés y español mediante el contraste con otros géneros del campo técnico. Tras su reconocimiento se llegan a extraer las características principales de este género en ambos idiomas.

DESARROLLO:

1. Los estudiantes observan varios textos expositivos e instructivos de distintos géneros del campo científico-técnico en inglés y español (Texto 3: "Descripción técnica de una lavadora, en inglés"; Texto 4: "Descripción técnica de una impresora, en inglés"; Texto 5: "Catálogo de materiales según las normas que se emplean en la industria petroquímica, en inglés"; Texto 6: "Descripción técnica de una mesa de mezclas, en español"; Texto 7: "Lista de materiales, en español"; Texto 8: "Prospecto de un medicamento, en español"; Texto 9: "Manual de instrucciones de un teléfono móvil, en español"; Texto 10: "Publirreportaje sobre tuberías y accesorios, en español"). Junto a éstos, también se incluyen textos que son descripciones técnicas de microondas (Textos 11-14: descripciones técnicas de microondas, en inglés y español).
2. Los alumnos deben discriminar el género de la descripción técnica del resto e identificar su foco dominante.
3. Se organizan en parejas y rellenan la Ficha 2 ("Características del género de la descripción técnica de microondas en Internet en inglés y en español") con las características propias que ven en el género de la descripción técnica. Unas parejas se centran en los textos en inglés y otras en los de español.
4. Al finalizar, las parejas se intercambian las fichas para que todos vean las características del género en ambas lenguas.

Dinámica de grupos: en parejas

Duración estimada: 2 horas

MATERIALES:

Fichas

- Ficha 2: "Características del género de la descripción técnica de microondas en Internet, en inglés y en español"

Textos

- Texto 3 (expositivo): "Descripción técnica de una lavadora, en inglés"
- Texto 4 (expositivo): "Descripción técnica de una impresora, en inglés"
- Texto 5 (expositivo): "Catálogo de materiales según las normas que se emplean en la industria petroquímica, en inglés"
- Texto 6 (expositivo): "Descripción técnica de una mesa de mezclas, en español"
- Texto 7 (expositivo): "Lista de materiales, en español"
- Texto 8 (expositivo con función secundaria instructiva): "Prospecto de un medicamento, en español"
- Texto 9 (instructivo): "Manual de instrucciones de un teléfono móvil, en español"
- Texto 10 (instructivo con función secundaria expositiva): "Publirreportaje sobre tuberías y accesorios, en español"
- Textos 11, 12: descripciones técnicas de microondas, en español
- Textos 13, 14: descripciones técnicas de microondas, en inglés

TAREA 3: COMPILACIÓN DE UN CORPUS BILINGÜE COMPARABLE

INTRODUCCIÓN:

Esta tarea tiene como fin que los estudiantes compilen un corpus bilingüe comparable (textos paralelos en LO y LM) con textos originales de descripciones técnicas de microondas en inglés y español extraídos de los sitios web de las empresas de este sector.

DESARROLLO:

1. Con la información que han recogido en la Ficha 1 ("Información sobre empresas que fabriquen o comercialicen microondas") y sabiendo las características del género textual que les ocupa, los estudiantes buscan descripciones técnicas de microondas en Internet en inglés y en español para formar entre todos un corpus bilingüe de textos paralelos; cada estudiante debe ocuparse de recoger todas las descripciones que encuentre de la empresa o marca que tenía asignada en un principio.
2. Guardan las descripciones en 2 carpetas (una para el corpus en español; otra para el de inglés – con subcarpetas para cada empresa). Se practican dos formas de recoger y guardar dichas descripciones desde Internet:

- Cuando sólo hay una descripción técnica en cada página web o la descripción está insertada dentro de un texto: Se puede recoger cada descripción una a una seleccionando el texto que corresponde a la descripción, copiándolo y pegándolo en un archivo de texto (.txt) o *MSWord* (.doc), por ejemplo.

- Cuando hay varios enlaces a descripciones técnicas en una página web: Se pueden recoger varias descripciones a la vez seleccionando varios enlaces con el puntero del ratón y descargando su contenido con un gestor de descarga como *Free Download Manager*, *Flashgot* o *Downthemall*, por ejemplo. Estos gestores deben estar instalados previamente en el ordenador para que estén integrados con el navegador (*Internet Explorer*, *Mozilla*).

3. Los nombres que se den a los documentos deben ser representativos de la marca y, si cabe, el origen (p.ej. míele1, míele2... o teka1mex, teka2mex...).
4. Una vez se ha acabado de guardar todos los textos, se procede a comentar con los alumnos en voz alta la naturaleza de los mismos, es decir, si son originales o traducciones según su observación y la información que han obtenido por correo electrónico de las empresas. Se guía la discusión hacia la necesidad de clasificar los textos no sólo por empresa e idioma, sino también por grado de certeza sobre si son originales o no. Se ordenan los textos de las empresas según estos criterios.
5. El docente puede encargarse de asegurar que al final todos los textos estén reunidos en un solo corpus y ordenados según la jerarquía acordada.

Dinámica de grupos: individual

Duración estimada: 3 horas

MATERIALES:

Fichas

- Ficha 1: "Información sobre empresas que fabriquen o comercialicen microondas"

Recursos electrónicos

- Programa informático de descarga de contenido Web *Free Download Manager*

COMENTARIOS:

- Es recomendable incidir en la importancia de conocer el origen de los textos, recordando, por ejemplo, que en el corpus se pueden incluir consciente o inconscientemente tanto textos originales como traducidos en distintas variantes geográficas del español, como el de Sudamérica, o del inglés.

TAREA 4: TERMINOLOGÍA DIFERENTE Y TERMINOLOGÍA COMPARTIDA**INTRODUCCIÓN:**

Esta tarea se ha introducido con el fin de concienciar a los estudiantes sobre una realidad que se da en los textos técnicos. Si bien en la mayoría de ocasiones distintos agentes pueden utilizar la misma terminología, en otras, pueden existir varios términos para referirse a un concepto igual o similar.

DESARROLLO:

1. Se facilita a los estudiantes la Ficha 3 ("Términos diferentes y términos en común"), junto con el corpus técnico que han creado y el programa *WordSmith Tools*.
2. Para rellenar la Ficha 3 los estudiantes tendrán que volver a los sitios web de cada empresa de microondas y fijarse en el título que tiene la sección donde se ofrece la información técnica de cada aparato, tanto en inglés como en español, para que se den cuenta de que cada marca se refiere a esta información de una forma diferente. Para la segunda parte de la ficha, tendrán que extraer las concordancias y agrupaciones de los términos propuestos y fijarse en aquellos compartidos por varias empresas.

Dinámica de grupos: individual

Duración estimada: 2 horas

MATERIALES:**Fichas**

- Ficha 3: "Términos diferentes y términos en común"

Descripciones de corpus

- Descripción 1: Corpus ad hoc de descripciones técnicas de microondas en inglés y español extraídas de Internet

Recursos electrónicos

- Corpus comparable bilingüe compilado entre todos ("Corpus descripciones técnicas microondas")
- Programa informático *WordSmith Tools*

TAREA 5: IDENTIFICACIÓN DE LA FIABILIDAD DE LOS TEXTOS Y LA IMPORTANCIA DE SER ORDENADO

INTRODUCCIÓN:

En esta tarea se pone de relieve la importancia de nombrar y ordenar bien los textos de un corpus para facilitar la posterior interpretación de las concordancias extraídas del mismo. A su vez, se incide en la necesidad de desarrollar un espíritu crítico sobre las fuentes que han aportado los textos para el corpus.

DESARROLLO:

1. Con la ayuda del programa *WordSmith Tools* se muestran a los estudiantes unas concordancias extraídas del corpus creado por ellos (Concordancia 1: "cocción") donde se ve un ejemplo de error de traducción en español.
2. Por parejas, los alumnos deben observar las concordancias, localizar el error y rellenar la Ficha 4 ("Errores en el corpus de español") con más errores o palabras que no pertenecen al español que aparecen en el corpus.
3. Para poder completar la Ficha 4 los alumnos tendrán que extraer del corpus de español una lista de palabras por orden de frecuencia y guardarla para futuros usos (Lista 1: "Lista de frecuencias en el corpus de español"). En este caso deberán fijarse en las palabras con frecuencias más bajas, ya que en un corpus de este tipo se espera que no sean abundantes.
4. Se facilitan unos materiales de apoyo al programa de análisis de corpus *WordSmith Tools*, que contiene indicaciones sobre cómo extraer listas de palabras.

Dinámica de grupos: en parejas

Duración estimada: 1 hora y media

MATERIALES:

Fichas

- Ficha 4 ("Errores en el corpus de español")

Concordancias

- Concordancia 1: "cocción"

Textos de apoyo

- 1: Material de apoyo al programa *WordSmith Tools*

Recursos electrónicos

- Corpus comparable bilingüe compilado entre todos ("Corpus descripciones técnicas microondas")
- Concordancias que muestran un error de traducción (Concordancia 1: "cocción")
- Programa informático *WordSmith Tools*

TAREA 6: CREATIVIDAD EN UN CORPUS DE TEXTOS TÉCNICOS

INTRODUCCIÓN:

Esta tarea persigue eliminar la idea de que en los textos técnicos originales no hay espacio para la creatividad y se centra en aprovechar un corpus comparable bilingüe precisamente como inspiración para la traducción creativa.

DESARROLLO:

1. Para introducir a los alumnos en el tema de la creatividad en traducción se les facilita un texto no necesariamente especializado donde aparece un término inventado en inglés para plantearles el ejercicio de traducirlo de forma creativa (Texto 15: "Snapple-crapple").
2. En parejas, los alumnos contestan a las primeras preguntas de la Ficha 5 ("Creatividad en los textos técnicos") para reflexionar sobre el tema de la creatividad en los textos técnicos.
3. Un miembro de la pareja abre la lista de palabras guardada en la tarea anterior (Lista 1: "Lista de frecuencias en el corpus de español") y el otro extrae otra lista de palabras por orden de frecuencia del corpus de inglés (Lista 2: "Lista de frecuencias en el corpus de inglés"). Para ambas acciones se debe utilizar la función *Wordlist* del *WordSmith Tools*.
4. Cada uno observa las palabras con frecuencias más bajas en la lista de su corpus, extrae aquellas que cree que son inventadas y las apunta en la Ficha 5.
5. A modo de conclusión acerca de la creatividad en los textos técnicos se comentan en voz alta las respuestas a la última pregunta de la Ficha 5.

Dinámica de grupos: en parejas

Duración estimada: 1 hora y media

MATERIALES:

Fichas

- Ficha 5: "Creatividad en los textos técnicos"

Listas

- Lista 1: "Lista de frecuencias en el corpus de español"

Textos

- Texto 15: "Snapple-crapple"

Recursos electrónicos

- Corpus comparable bilingüe compilado entre todos ("Corpus descripciones técnicas microondas")
- Lista 1: "Lista de frecuencias en el corpus de español"
- Programa informático *WordSmith Tools*

TAREA 7: SABER APROVECHAR LAS PISTAS QUE DA LA LENGUA ORIGINAL (LO)**INTRODUCCIÓN:**

Esta tarea es clave en tanto que muestra al alumno cómo aprovechar la información cotextual de unas concordancias en lengua original para buscar equivalentes de traducción en la lengua de llegada. Partiendo de un término especializado desconocido y que no aparece en los diccionarios, el alumno debe acudir al corpus comparable bilingüe que ha creado para encontrar vías de solución a sus problemas de traducción.

DESARROLLO:

1. Se facilita una descripción técnica de microondas en inglés que incluya términos técnicos que no resuelva un diccionario (Texto 13: "Siemens HF17065GB"), se dirige la atención de los alumnos hacia uno de esos términos ("trim kit") en principio desconocido para ellos y se busca en el corpus de inglés.
2. Se les facilita la Ficha 6a ("Cuando el diccionario no ayuda con la terminología...") con preguntas con las que se pretende que cuestionen la utilidad del diccionario para este tipo de traducción.
3. Se comentan en voz alta las respuestas hasta el momento.
4. Se les facilita la Ficha 6b ("...La clave puede estar en el cotexto de unas concordancias") para que comprueben la utilidad de observar el cotexto en unas concordancias extraídas del corpus comparable bilingüe creado a la hora de buscar equivalentes de traducción.
5. Se comentan en voz alta las respuestas encontradas hasta el momento.
6. Se les facilita la Ficha 6c ("...La clave puede estar en el cotexto de unas concordancias") para que practiquen esta observación del cotexto y ratifiquen la importancia de contar con un corpus en lengua de partida y de llegada.

Dinámica de grupos: en parejas

Duración estimada: 2 horas

MATERIALES:**Fichas**

- Ficha 6a: "Cuando el diccionario no ayuda con la terminología..."
- Ficha 6b: "...La clave puede estar en el cotexto de unas concordancias"
- Ficha 6c: "...La clave puede estar en el cotexto de unas concordancias"

Textos

- Texto 16: "Siemens HF17065GB"

Recursos electrónicos

- Corpus comparable bilingüe compilado entre todos ("Corpus descripciones técnicas microondas")
- Programa informático *WordSmith Tools*

EVALUACIÓN:

La puesta en común de las primeras respuestas de la Ficha 6a y la evolución hasta llegar a contestar la última pregunta de la Ficha 6b constituyen la evaluación de esta tarea. La Ficha 6c incide en la práctica de lo aprendido, con una ligera variación.

COMENTARIOS:

- Para atraer la atención de los alumnos hacia el término "trim kit" es conveniente que comprueben que es un término de este campo; para este fin se puede empezar por recuperar la lista de palabras por orden de frecuencia y hacerles ver que aparece varias veces en el corpus de inglés.

TAREA FINAL: ENCONTRAR EQUIVALENCIAS UTILIZANDO EL RECURSO CREADO

INTRODUCCIÓN:

En esta tarea los alumnos integran lo aprendido en la unidad, especialmente en lo que se refiere a aprovechar el corpus comparable bilingüe especializado que han creado entre todos para resolver la traducción de una descripción técnica de un microondas.

DESARROLLO:

1. Se da a los alumnos un texto original en inglés ilustrativo del género de la descripción técnica y que contenga problemas diversos de traducción (puede ser un texto utilizado durante el trascurso de la unidad didáctica o, como en este caso, un texto manipulado) para que lo traduzcan (Texto 17: "Roper MHE14XMD").
2. Se aclara que para traducir este texto sólo podrán utilizar como fuente de documentación el corpus comparable bilingüe creado entre todos, en combinación con el programa *WordSmith Tools*. También podrán utilizar un conversor de cantidades y un buscador de imágenes en Internet.
3. Rellenan la Ficha de evaluación del uso de corpus a la vez que traducen para poder después evaluar cada paso que han dado en la resolución de 5 problemas de traducción del TO preseleccionados por el docente.
4. Cuando terminen, rellenan la Ficha de autoevaluación y la Ficha de valoración de la unidad didáctica.

Dinámica de grupos: individual

Duración estimada: 2 horas y media

MATERIALES:

Fichas

- Ficha de evaluación del uso de corpus al traducir
- Ficha de autoevaluación

- Ficha de valoración de la unidad

Textos

- Texto 17: "Roper MHE14XMD"

Recursos electrónicos

- URL de conversor de cantidades en línea [*Ask Numbers*: <http://www.asknumbers.com/>]
- URL de buscador de imágenes [*Google Imágenes*: <http://images.google.es/>]
- Corpus comparable bilingüe compilado entre todos ("Corpus descripciones técnicas microondas")
- Programa informático *WordSmith Tools*
- Programa informático *InfoRapid*

EVALUACIÓN:

- Cada estudiante muestra al resto cómo ha solucionado un determinado problema con la ayuda del corpus.
- Los alumnos entregan su Diario de aprendizaje rellenado a lo largo de la unidad.

COMENTARIOS:

- Se recuerda la importancia de tener un espíritu crítico a la hora de considerar los textos paralelos disponibles.

③ EVALUACIÓN DE LA UNIDAD DIDÁCTICA

1. Valoración realizada por los estudiantes

a) Valoración de su aprendizaje: ficha de autoevaluación

La valoración de los estudiantes de su propio aprendizaje se refleja en las puntuaciones que aparecen en la tabla que sigue (Tabla 22).

- Puntuación media de las respuestas proporcionadas por cada alumno. De las notas medias de cada alumno, se observa una tendencia hacia lo positivo en general y también en particular en, por ejemplo, F.M., alumna agradecida y abierta a las nuevas metodologías docentes más adaptadas y centradas en el estudiante, y hacia lo negativo, en, por ejemplo, A.G. y M.O., muy escépticos respecto a la metodología docente empleada, especialmente el trabajo con fichas y el trabajo de observación del detalle.
- Puntuación media de las respuestas en cada pregunta. De entre estas puntuaciones, se puede destacar la seguridad que los alumnos han adquirido a la hora de identificar las características del género tratado en la unidad, así como de comprender la importancia que tiene nombrar y ordenar bien los archivos de un corpus para poder aprovecharlo bien después. Parece, sin embargo, que la localización de términos creativos entre textos técnicos no ha quedado tan fijada. Por otro lado, los estudiantes valoran especialmente la ayuda que les ha aportado el corpus a la hora de hallar equivalentes de traducción, si bien es cierto que no es tanto así cuando se trata de hacerse más conscientes de la naturaleza y calidad de los textos que lo componen.

FICHA DE AUTOEVALUACIÓN: RESULTADOS UNIDAD DIDÁCTICA "UTILIDAD DE UN CORPUS COMPARABLE BILINGÜE ESPECIALIZADO PARA TRADUCIR TEXTOS TÉCNICOS"

Estudiantes	H. A.	M. A.	O. A.	B. B.	N. B.	N. Bo.	R. C.	M. F.	J. F.	A. F.	A. G.	E. G.	A. Go.	M. J.	R. J.	E. K.	A. L. R.	R. L.	E. L. M.	A. M.	F. M.	L. M.	N. M.	M. Ma.	M. M.	M. O.	M. O.	J. O.	A. P.	J. L. P.	S. R.	M. R.	P. R.	C. S.	S. S.	B. V.	A. V.	T. A.	Media de cada pregunta	
	Soy capaz de...	7	7	8	7	8	9	7	8	8	7	7	8	10	8	8	9	8	8	9	10	10	8	8	9	7	7	8	7	9	8	7	9	8	6	6	8	9	8	8
buscar fuentes de donde recoger textos técnicos	9	9	10	7	9	8	7	10	7	6	8	8	9	9	8	10	10	8	10	10	10	10	8	6	9	6	9	6	7	8	7	9	8	9	8	8	8	8	8	8,4
identificar el género de la descripción técnica de microondas en Internet en inglés y español	9	7	9	8	10	10	9	8	9	8	4	9	8	9	7	10	9	9	8	7	9	10	9	10	6	10	6	6	9	7	7	7	9	8	7	8	6	7	8,1	
compilar un corpus bilingüe comparable (textos originales en LO y LM) de textos técnicos	9	10	10	7	10	9	9	10	8	5	9	9	10	10	7	10	10	10	10	10	10	10	9	8	10	8	10	6	9	9	7	6	10	8	9	8	10	7	8,7	
comprender la importancia de nombrar y ordenar bien los archivos de un corpus	7	8	9	9	10	8	8	8	7	7	7	8	8	8	6	9	8	7	8	10	10	7	7	7	7	8	8	7	9	8	7	9	8	7	7	6	7	7,8		
tener un espíritu crítico sobre las fuentes de textos técnicos	6	6	7	8	6	8	8	7	8	6	6	7	7	7	7	8	7	7	8	8	8	9	6	5	7	4	7	7	7	7	5	8	7	8	7	7	8	6,9		
encontrar elementos creativos en un corpus de textos técnicos	6	7	10	9	10	8	8	8	6	7	7	8	10	9	6	9	9	8	7	7	9	8	9	6	9	4	7	7	7	7	5	8	7	8	7	7	8	7,7		
encontrar equivalentes terminológicos a partir de la información contextual (las "pistas") que ofrecen las concordancias de un corpus de textos técnicos	9	7	10	9	10	8	8	8	6	7	7	8	10	9	6	9	9	9	8	7	9	8	9	6	9	4	7	7	9	6	8	7	7	8	7	8	7	7,7		
trabajar individualmente en la traducción especializada del inglés al español	9	9	7	8	9	8	9	10	7	8	8	9	8,5	9	6	8	8	8	9	7	7	9	8	8	9	8	9	8	8	9	8	8	9	9	6	6	7	7,9		
trabajar en pareja / en grupo en la traducción especializada del inglés al español	9	7	10	9	10	8	7	8	10	8	6	9	9	9	7	8	9	8	9,5	9	7	8	9	7	8	9	8	9	10	8	7	6	9	8	7	10	9	8,3		
Media de cada alumno	7,9	7,6	9,1	7,9	9	8,5	7,9	8,1	8,3	7,1	6,4	8,3	8,8	8,4	7	9,1	8,3	8,3	8,2	8,4	9,2	8,2	8,5	6,7	8,5	6,7	6,7	7,8	7,7	8,7	7,2	7,3	8,2	7,9	7,8	7,3	7,8	7,5		
Te ha ayudado el corpus a...	9	7	5	7	9	8	8	9	8	7	8	7	10	9	8	9	10	8	7	8	10	8	8	7	10	4	6	10	10	6	8	8	8	7	8	7	6	8	7,9	
encontrar equivalentes de traducción	8	8	10	9	10	10	9	8	9	7	8	9	10	10	8	9	9	8	8	7	10	10	9	7	9	7	9	6	8	7	10	7	9	9	8	8	8	7	8,5	

estar más seguro de una solución	7	8	8	10	10	9	8	10	7	9	10	9	10	10	7	10	5	8	7	8	6	7	9	8	7	9	8,3										
asimilar la falta de univocidad en traducción, incluso en los textos técnicos (p.ej. "trim kit" = "marco/kit de encastre/empotramiento")	9	10	9	9	8	9	8	10	8	7	9	7	8	7	9	6	10	9	9	8	8	8	9	8	7	9	8,1										
comprender la importancia de las colocaciones	8	9	10	10	9	8	8	10	6	8	8	9	9	8	7	10	10	9	7	7	10	10	7	8	8	8	8										
ser más consciente de la naturaleza de los textos utilizados en la fase de documentación (origen, fecha, original o traducción, etc.)	9	7	6	9	10	10	7	9	9	7	8	9	9	10	7	6	7	10	7	10	7	6	9	7	8	8	7,8										
Media de cada alumno	8,3	8,2	8,3	9	9,2	8,8	8,2	9,2	7,8	7	8	8,5	9,3	9,2	7,3	9	9,5	7,8	7,3	7,2	9,8	9,3	8,2	6,5	9,5	6	7,8	8,7	8,8	6,7	8	8,7	7,8	8,2	7,7	7	8

Tabla 22. Resultados de la Ficha de autoevaluación de los alumnos de la UD "Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos"; resultados del curso 2006-07.

Aparte de puntuar, los estudiantes también han querido destacar especialmente algunos aspectos relacionados con la unidad como, por ejemplo, la conveniencia de organizarse bien en la fase de documentación y ordenar correctamente los archivos de un proyecto, en este caso, de un corpus, para poder extraer el máximo provecho del mismo. Se han percatado de que, si bien antes observaban muy pocos textos paralelos, ahora consultan muchísimos más, en menos tiempo, y los aprovechan mejor gracias a los programas informáticos que han descubierto. Según los estudiantes mismos, ahora saben por dónde empezar a tratar la traducción de un texto técnico de un tema del que no tienen conocimientos previos, han aprendido a prestar atención a los términos registrados por una marca o sus preferencias en cuanto a la terminología; se han dado cuenta de que incluso los textos técnicos publicados por empresas profesionales contienen errores y que, reuniendo muchos textos, se pueden observar patrones en cuanto a frases, características de un producto, formato en el que se muestra la información. Con todo esto, han aprendido a ser más críticos con su trabajo y que no deben fiarse sólo de los diccionarios ni de su “intuición” en lo que respecta a equivalencias de términos técnicos, sino tomar decisiones basadas en corpus que les permitan sentirse más seguros. Las respuestas completas de donde se han extraído estas conclusiones pueden leerse en el *Volumen 2*.

b) De la unidad didáctica

Lo que sigue a continuación es la Ficha de valoración de la unidad didáctica que se pasó a los estudiantes y, tras ésta, los temas recurrentes que se observaron en sus respuestas, las cuales también pueden leerse al completo en el *Volumen 2*.

**FICHA DE VALORACIÓN
UNIDAD DIDÁCTICA "UTILIDAD DE UN CORPUS COMPARABLE BILINGÜE ESPECIALIZADO PARA TRADUCIR
TEXTOS TÉCNICOS"**

Comentarios sobre la unidad

- ¿Crees que utilizando corpus has aprendido también una nueva forma de documentarte para resolver problemas de traducción en un texto? Por favor, explicita al máximo.
.....
- ¿Crees que podrías traducir mejor textos técnicos del inglés al español por utilizar corpus? Por favor, explicita al máximo.
.....
- ¿Hay algo de esta unidad didáctica que te haya gustado especialmente entre los siguientes temas? Por favor, explicita al máximo.
 - Buscar textos paralelos
 - Contactar con empresas del sector
 - Identificar las características del género "descripción técnica"
 - Recoger textos paralelos
 - Identificar errores en las descripciones técnicas de microondas en español
 - Identificar términos creativos en descripciones técnicas
 - Aprovechar un corpus de textos paralelos con un programa de búsqueda en una unidad local (*WordSmith Tools* o *InfoRapid*) para encontrar equivalentes de traducción
- ¿Hay algo que no te haya gustado? Por favor, explicita al máximo.
.....
- ¿Cómo lo cambiarías?
.....
- Otros comentarios
.....

Documento 14. Ficha de valoración de la UD "Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos".

(i) Nueva forma de documentarse

Prácticamente todos los estudiantes afirman que han encontrado en el uso de corpus una manera de documentarse diferente a la que practicaban hasta el momento ya que desconocían los beneficios de guardar en forma de corpus los textos que consultaban. Utilizar corpus les ha permitido, entre otras cosas, hacer búsquedas y obtener resultados con menos ruido, ver frecuencias de uso, contrastar equivalencias al poder ver si los contextos en los dos idiomas coinciden, extraer significado a partir del propio contexto, constatar la falta de univocidad terminológica entre dos lenguas y, sobre todo,

no quedarse al nivel de la palabra sino del término compuesto, el sintagma, las agrupaciones o las colocaciones. En definitiva, los estudiantes coinciden en que utilizar corpus para traducir es útil y les ahorra tiempo, con lo que están convencidos de que lo usarán en su futuro profesional.

(ii) Mejoras a la hora de traducir textos técnicos

El mayor beneficio que aporta el uso de corpus es, según los estudiantes, la posibilidad de observar los términos en el contexto adecuado, lo cual puede también aportarles información semántica, fórmulas o equivalencias ya establecidas en el campo. A determinados niveles, un corpus como el creado y utilizado en esta unidad puede ofrecer unas pautas sobre cómo debe ser el texto meta en español en cuanto a las convenciones del género o el estilo que espera encontrar un experto en el campo, eso sí, siempre y cuando los textos del corpus hayan sido seleccionados adecuadamente. Los estudiantes reconocen que, ante un texto especializado, ni sus conocimientos sobre su lengua materna, ni algunos recursos lexicográficos pueden ayudar con cuestiones de uso, en cuyo caso utilizar un corpus les permite no sólo encontrar o contrastar equivalencias, sino agilizar el proceso de traducir dicho texto.

(iii) Aspectos destacados positivamente por los alumnos

Siguiendo un orden decreciente, a continuación se comentan los elementos que, a juzgar por las respuestas de los estudiantes, les han resultado más atractivos en la unidad:

- 1) Aprovechar un corpus de textos paralelos con un programa de búsqueda en una unidad local para encontrar equivalentes de traducción. No cabe duda de que lo que más interesante han encontrado en la unidad didáctica ha sido poder aprovechar con *InfoRapid* y *WordSmith Tools* el corpus de textos paralelos que ellos mismos han construido para encontrar equivalentes de traducción. Estos programas les han ayudado a superar el nivel de la palabra, las búsquedas en el diccionario o en buscadores generales en Internet, con el resultado de, cuando menos, un ahorro de tiempo al no tener que mirar uno por uno los textos paralelos consultados.
- 2) Identificar errores en las descripciones técnicas de microondas en español. Los estudiantes destacan la necesidad de concienciarse sobre los errores que se dan en este tipo de textos para evitarlos.

- 3) Identificar términos creativos en descripciones técnicas. Esta tarea ha gustado por deshacer el prejuicio de que la creatividad sólo se puede encontrar en publicidad, literatura, etc., y por concienciarles del cuidado que deben tener con determinadas palabras que pueden pertenecer exclusivamente a una empresa.
- 4) Buscar y recoger textos paralelos. Si bien no es una tarea desconocida para los estudiantes, sí han encontrado interesante el hecho de que el corpus se haya construido con la colaboración de todos y que se tratara la conveniencia de reunir muchos, tenerlos guardados en una unidad local y no depender de estar conectado a Internet, y agruparlos según criterios distintivos como el idioma, si son originales o traducidos.
- 5) Contactar con empresas del sector. A pesar de la frustración de algunos alumnos al no recibir respuesta por parte de la empresa de pequeños electrodomésticos que se les había asignado, sí han visto de manera positiva el ejercicio de ponerse en contacto con el mundo profesional ya que eso les ha obligado a escribir mensajes de correo electrónico concisos y serios.
- 6) Identificar las características del género “descripción técnica”. A pesar de reconocer los estudiantes que ésta es una parte clave de la unidad didáctica, ya que sin reconocer bien el género textual no sería posible realizar una traducción adecuada a las convenciones de la lengua de llegada, no han encontrado esta tarea tan atractiva como las otras. Aún así, sí han podido apreciar la utilidad de fijarse en el lenguaje como materia de estudio y no sólo como un texto que traducir donde lo importante es el resultado.

(iv) Sugerencias de mejora y comentarios

Si bien es cierto que los estudiantes han considerado interesante y muy útil prácticamente todo el contenido de esta unidad didáctica, también lo es que, como así lo han expresado algunos de ellos, por ejemplo, les desanimó no recibir respuesta de algunas empresas, con lo que sugieren contactar con las mismas por otros medios o para otros propósitos para que haya más implicación por su parte.

Otro comentario reiterado es el de que, a pesar de reconocer la utilidad de hacer fichas de ejercicios que les han ayudado a reflexionar y fijar lo aprendido, sí les hubiera gustado hacer más ejercicios de traducción en el sentido más tradicional o dedicar más tiempo a conocer los programas informáticos tratados, ahorrando tiempo en, por ejemplo, la identificación de las características del género, los términos creativos o los errores.

La búsqueda y recogida de textos paralelos se reconocen arduas pero muy fructíferas, aunque algún estudiante ha echado de menos el uso de recursos lexicográficos bilingües ya contruidos del tipo de los que usan normalmente.

En definitiva, los estudiantes admiten su sorpresa ante todo lo que, gracias al uso de corpus, han aprendido sobre la traducción y, además, sobre recursos que les facilitan la tarea de documentación, les ahorran tiempo y les ayudan a sentirse más seguros, incluso habiendo partido de una falta total de experiencia en traducción técnica.

2. Valoración realizada por el docente

a) Valoración del aprendizaje de los estudiantes

Los alumnos plasman en la Ficha de recogida de datos sobre el uso de corpus cómo han utilizado este recurso a la hora de resolver 5 problemas de traducción seleccionados del texto original en su Tarea final. La observación de esta ficha y la traducción producida por cada alumno nos permite rellenar la Ficha de cruce de datos, que mostramos con un resumen de todos los resultados. En el *Volumen 2* de esta tesis se pueden encontrar todos los resultados de la Ficha de cruce de datos de todos los estudiantes que siguieron la unidad en el curso 2007-08.

FICHA DE CRUCE DE DATOS: RESUMEN DE RESULTADOS		UNIDAD DIDÁCTICA: "UTILIDAD DE UN CORPUS COMPARABLE BILINGÜE ESPECIALIZADO PARA TRADUCIR TEXTOS TÉCNICOS"																														
Estudiantes	J. T.	J. S. G.	J. D. G.	J. A. B.	J. I. A.	J. I. B.	J. M. A.	R. C. A.	M. A. C.	A. M. C.	B. C. C.	T. T.	T. M.	I. M. S.	V. S. B.	M. N. B.	A. E. N.	A. D. M.	A. B. D.	A. L. Di.	S. F.	S. R. A.	M. O.	M. C.	J. C. L.	A. L. Z.	G. R.	M. Be.	Media de cada indicador (tras aplicar el %)	Media de cada indicador (sobre 10 puntos)		
Indicador 1: realiza búsquedas de calidad y económicas con un programa de análisis de corpus (50%)	4,5	3,5	4	5	3,5	5	5	4	5	5	5	4,5	5	4	5	4,5	5	4,5	5	4	5	4,5	5	4,5	5	4	4	5	4	4,6/5	9,2	
Indicador 2: justifica las opciones de traducción basándose en datos sacados del corpus (50%)	3	3,5	5	5	5	3,5	4	5	4,5	4,5	4	3,5	5	3	4,5	4	4	4	5	5	4,5	4	2	5	3	0	4	5	3	4/5	8	
Nota de cada alumno	7,5	7	9	10	8,5	8,5	9,5	8,5	9,5	9	10	8	10	10	7	9,5	8,5	9,5	9	10	9	8,5	9	6,5	10	7	4	9	10	7		

Tabla 23. Resumen de resultados de la Ficha de cruce de datos de la UD "Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos"; resultados del curso 2007-08.

- Puntuaciones medias de los alumnos. La nota media de todas las notas obtenidas en el uso de corpus para traducir los 5 problemas seleccionados es de 8,6 puntos sobre 10, lo cual es un muy buen resultado.
- Puntuaciones medias obtenidas por el conjunto de alumnos en cada indicador. Si nos fijamos en los dos indicadores de competencias evaluados: (1) realiza búsquedas de calidad y económicas con un programa de análisis de corpus y, (2) justifica las opciones de traducción basándose en datos sacados del corpus, se observa en los estudiantes una mayor competencia a la hora de aplicar estrategias de búsqueda de equivalentes en el corpus (nota media: 9,2) que a la hora de justificar sus decisiones (nota media: 8). En este sentido, algunos todavía recurren a la justificación basada en su intuición de nativos de la lengua meta o aducen que una solución es correcta porque aparece más frecuentemente que otra en el corpus. Si bien es cierto que, dados los resultados, se puede afirmar que los objetivos propuestos para la unidad y evaluados en este caso sí se han conseguido, también lo es que se observa una necesidad de insistir en la importancia de poder justificar con documentación, en este caso, con un corpus especializado, una elección de traducción.

b) Valoración de la unidad didáctica

Tras probar la unidad total o parcialmente en varias ocasiones y observar los comentarios y resultados de los estudiantes, creemos que con ésta se han conseguido las metas propuestas, especialmente, la creación de un corpus comparable bilingüe de un género textual técnico y su explotación para la traducción técnica utilizando funciones diversas de varios programas informáticos.

④ EVOLUCIÓN DE LA UNIDAD DIDÁCTICA

Modificaciones efectuadas

La versión que se presenta aquí recoge ya las sugerencias de los estudiantes y las modificaciones introducidas tras probar la unidad. Estas modificaciones se refieren a las siguientes Fichas o Tareas:

- Tarea 2 (“Identificar el género textual”): Se redujo el número de textos técnicos no pertenecientes al género de la unidad que se ofrecían con fines contrastivos.
- Tarea 4 (“Terminología diferente y terminología compartida”): La dificultad observada en los alumnos a la hora de asimilar la variedad terminológica en los textos técnicos, es decir, su incredulidad ante la falta de univocidad terminológica, llevó a desarrollar esta tarea, que también permite trabajar la terminología que se repite en el campo.
- Tarea 5 (“Identificación de la fiabilidad de los textos o la importancia de ser ordenado”): Al comprobar que algunos alumnos o bien no daban nombres transparentes a sus archivos o bien no se fijaban en los nombres de éstos a la hora de tomar decisiones respecto a la fiabilidad de los textos se añadió esta tarea con el fin de facilitar la interpretación de los resultados extraídos del corpus.
- Tarea final: Por último, tras las peticiones de los alumnos, se puso un mayor énfasis en la traducción del texto final, dedicándole una tarea completa como ejercicio de aplicación de lo aprendido.



UNIDAD DIDÁCTICA

Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés



Esta unidad tiene una intención totalmente reflexiva en tanto que introduce al estudiante en la traducción inversa al inglés planteándole una gran diversidad de pequeños ejercicios en los que se utilizan corpus variados y otros recursos. El estudiante, a la vez que se inicia en la traducción al inglés, lo hace también en algunos principios del trabajo con corpus, en el manejo de corpus en Internet y de programas de análisis de corpus, todo ello de forma contextualizada y buscando siempre estimular el razonamiento crítico.



Grupos de prueba: Esta unidad didáctica fue puesta en práctica en tres ocasiones con algunas variaciones, primero con estudiantes de último curso de licenciatura en Letras en Belo Horizonte, Brasil, y después en dos cursos de máster en traducción especializada en Barcelona, siempre en asignaturas de traducción especializada inversa (español-inglés B). En los tres casos el número de alumnos fue reducido (14, 6 y 4, respectivamente), lo que facilitó mucho la interacción y la evaluación. Las tareas, fichas y evaluación que se presentan aquí corresponden a las utilizadas en la segunda y definitiva puesta en práctica de la unidad didáctica en el curso 2007-08.

Los materiales de esta unidad se encuentran en el *Volumen 2* de esta tesis (págs. 555-612).



❶ PRESENTACIÓN DE LA UNIDAD DIDÁCTICA

“Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”

NIVEL DE USO DE CORPUS

- Iniciación

MATERIA

- Iniciación a la traducción inversa (español A – inglés B)

OBJETIVOS

Objetivos relacionados con el uso de corpus para traducir

- Conocer aspectos básicos sobre corpus
- Identificar los principios metodológicos básicos del trabajo con corpus
- Utilizar sistemas de búsqueda en corpus en línea a nivel básico
- Utilizar las funciones básicas de un programa informático de análisis de corpus
- Seleccionar el corpus apropiado según las necesidades de traducción
- Integrar conocimientos para resolver problemas de traducción utilizando corpus a nivel básico

Objetivos relacionados con la materia

- Identificar el funcionamiento de la traducción inversa
- Manejar técnicas de documentación básicas
- Escribir textos en la LM que son gramaticalmente apropiados y pragmáticamente adecuados
- Conocer diferencias de léxico entre la LO y la LM dentro de ciertos campos y registros
- Conocer diferentes tipos de textos

CONTENIDOS

- Diversos tipos de corpus (I. 1.1.)
- Corpus de inglés en Internet (*BNC, WordBanksOnline, Business Letter Corpus, TEC, Brown*) (I. 1.2.1.)
- Corpus de español en Internet (*CREA*) (I. 1.2.2.)
- Repositorios de textos en inglés (*Guardian Archive*) (I. 1.3.1.)
- La web como corpus (*Internet pública*) (I. 1.4.)
- Concienciación sobre la observación del contexto (I. 2.2.6.)
- Índices de frecuencia (I. 2.2.7.)
- Software gratuito de análisis de corpus (*KwicFinder*) (III. 1.10.1.)
- Software comercializado de análisis de corpus (*WordSmith Tools*) (III. 1.10.2.)
- Función *Concord* (de *WordSmith Tools*) (III. 2.1.1.)
- Función *Resort* (de *WordSmith Tools*) (III. 2.3.1.)
- Selección del corpus más apropiado (IV. 1.)



- Extracción de definiciones: hiperónimos, etc. (IV. 2.1.1.)
- Extracción de tipos: hipónimos, etc. (IV. 2.1.2.)
- Extracción de vocabulario general y términos de especialidad (IV. 2.2.1.)

ESTRUCTURA

- Tarea 1: "Desde la perspectiva de un nativo"
- Tarea 2: "Dar con el corpus (más) idóneo"
- Tarea 3: "Frecuencias"
- Tarea 4: "Mirar a la izquierda, a la derecha y a la izquierda otra vez"

EVALUACIÓN

Indicadores

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación,...

- Indicador 1: maneja distintos (tipos de) corpus que están disponibles en línea (de la lengua general y especializada)
- Indicador 2: selecciona el corpus apropiado
- Indicador 3: interpreta las frecuencias de aparición
- Indicador 4: reordena el contexto de una concordancia

Fichas de evaluación

- Ficha de autoevaluación
- Ficha de valoración de la unidad

MATERIALES

Fichas

- Worksheet 1a: "What we think we know about language"
- Worksheet 1b: "What we think we know about language"
- Worksheet 1c: "Additional exercises"
- Worksheet 2: "Selecting the appropriate corpus"
- Worksheet 3: "Interpreting frequencies"
- Worksheet 4a: "Look right, left and right again"
- Worksheet 4b: "Look right, left and right again"
- Ficha de autoevaluación
- Ficha de valoración de la unidad

Textos originales y paralelos

- Texto 1: "Vacío legal inexplicable"
- Texto 2: "Perfil biofísico fetal"
- Texto 3: "Exploración ecográfica"
- Texto 4: "The future is yummy"
- Texto 5: "Euroblanca"
- Texto 6: "El canto de Cabo Verde"
- Texto 7: "Clínicas veterinarias"



Textos de apoyo

- 1: Material de introducción a la unidad
- 2: Material de apoyo al programa *WordSmith Tools*

Descripciones

- Descripción 1: corpus ad hoc de textos sobre aneuploidías extraídos del *Medline*
- Descripción 2: corpus de textos sobre economía extraídos del *Financial Times*
- Descripción 3: corpus ad hoc de textos con instrucciones de doctor a paciente extraídos de Internet
- Descripción 4: corpus ad hoc de textos sobre inteligencia emocional extraídos de Internet

Recursos electrónicos

- URL de diccionario monolingüe general [*Onelook*: <http://www.onelook.com>]
- URL de diccionario bilingüe general [*Wordreference*: <http://wordreference.com> ; *Vox*: <http://www.diccionarios.com> ; *Babylon*: <http://www.lexicool.com/babylon-online-translation-dictionary.asp?IL=3>]
- URL de diccionario de dudas del español [*Panhispanico*: <http://www.rae.es>]
- URL de diccionario de medicina en inglés [*Medical Dictionary*: <http://www.medical-dictionary.com>]
- URL de servicio multilingüe de consultas terminológicas [*Termcat*: <http://www.termcat.cat/>]
- URL de base de datos terminológica o glosario bilingüe [*IATE*: <http://iate.europa.eu/> ; *AndyMiles*: <http://andymiles.com/>]
- URL de recurso de expresiones idiomáticas [<http://www.ccdmd.qc.ca/ri/expressions/>]
- URL de buscador general [*Google*: <http://www.google.com>]
- URL de herramienta de comparación de frecuencias en la Web [*Googlefight*: <http://googlefight.com/>]
- URL de dudas lingüísticas del inglés resueltas con corpus [<http://www.eisu2.bham.ac.uk/Webmaterials/kibbitzers/kibbitzer06.htm>]
- URL de corpus de español general [*CREA*: <http://www.rae.es>]
- URL de corpus de inglés general [*WordBanksOnline*: <http://www.collins.co.uk/Corpus/CorpusSearch.aspx> ; *British National Corpus -BNC*: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- URL de archivo en línea de diario británico [*Guardian Unlimited*: <http://www.guardian.co.uk/Archive/>]
- URL de corpus de correspondencia comercial [*Business Letter Corpus*: <http://ysomeya.hp.infoseek.co.jp/>]
- URL de corpus paralelo [*COMPARA*: <http://www.linguateca.pt/COMPARA/>]
- URL de corpus de inglés americano [*Brown Corpus*: http://www.lextutor.ca/concordancers/concord_e.html]
- URL de corpus de traducciones en inglés [*TEC*: <http://ronaldo.cs.tcd.ie/tec/jnlp/>]
- URL de recurso en línea para utilizar la web como corpus [*WebCorp*: <http://www.webcorp.org.uk/>]



- URL de programa informático para utilizar la web como corpus [*KwicFinder*: <http://www.kwicfinder.com/KWiCFinder.html>]
- Colocaciones ("colloc_interest_FT.bmp") ya extraídas de un corpus de textos periodísticos financieros en inglés (*Financial Times*) y su descripción
- Concordancias ("conc_doppler.bmp") ya extraídas de un corpus ad hoc de textos médicos sobre aneuploidías ("aneuploidies_9000.txt")
- Corpus ad hoc de aproximadamente 9000 resúmenes de artículos médicos sobre aneuploidías en inglés extraído del *Medline*
- Corpus ad hoc de unas 66.000 palabras con instrucciones dadas por doctores a pacientes extraído de páginas web
- Corpus ad hoc sobre inteligencia emocional de unas 112.000 palabras en español y unas 89.000 palabras en inglés
- Programa informático *WordSmith Tools*

DURACIÓN ESTIMADA

8 horas

DESARROLLO DE LA UNIDAD DIDÁCTICA

TAREA 1: DESDE LA PERSPECTIVA DE UN NATIVO

INTRODUCCIÓN:

Esta tarea busca que, a través de la utilización de corpus en línea, el estudiante empiece a reflexionar y tome conciencia de la naturaleza repetitiva y predecible de la lengua y del valor relativo de la intuición del hablante nativo.

DESARROLLO:

1. Los alumnos visualizan el material de introducción a la unidad, que es un archivo *PowerPoint*.
2. Seguidamente, leen las preguntas que contiene la Worksheet 1a ("What we think we know about language") en formato *PowerPoint*, que buscan hacer pensar al alumno sobre lo que sabe intuitivamente de su lengua materna y del inglés. A la vez, responden en la Worksheet 1b ("What we think we know about language"), en formato *MSWord*.
3. Tienen a su disposición la Worksheet 1c ("Additional exercises") para hacerla como refuerzo o ampliación de la tarea.

Dinámica de grupos: individual

Duración estimada: 2 horas

MATERIALES:

Fichas

- Worksheet 1a: "What we think we know about language"
- Worksheet 1b: "What we think we know about language"
- Worksheet 1c: "Additional exercises"

Textos de apoyo

- 1: Material de introducción a la unidad

Recursos electrónicos

- URL de diccionario bilingüe general [*Wordreference*: <http://wordreference.com>]
- URL de buscador general [*Google*: <http://www.google.com>]
- URL de recurso de expresiones idiomáticas [<http://www.ccdmd.qc.ca/ri/expressions/>]
- URL de servicio multilingüe de consultas terminológicas [*Termcat*: <http://www.termcat.cat/>]
- URL de corpus de español general [*CREA*: <http://www.rae.es>]
- URL de corpus de inglés general [*WordBanksOnline*: <http://www.collins.co.uk/Corpus/CorpusSearch.aspx> ; *British National Corpus -BNC-*: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- URL de archivo en línea de diario británico [*Guardian Unlimited*: <http://www.guardian.co.uk/Archive/>]

- URL de recurso en línea para utilizar la web como corpus [WebCorp: <http://www.webcorp.org.uk/>]
- URL de programa informático para utilizar la web como corpus [KwicFinder: <http://www.kwicfinder.com/KWiCFinder.html>]

EVALUACIÓN:

Se ponen en común las respuestas de los alumnos para que constaten que prácticamente todos han contestado lo mismo.

COMENTARIOS:

- A no ser que se indique lo contrario, para las Worksheet 1a y 1b, no se permite que nadie busque nada en ningún recurso ya que interesan respuestas intuitivas. Sólo se puede buscar para ratificar las respuestas una vez puestas en común y comentadas.

TAREA 2: DAR CON EL CORPUS (MÁS) IDÓNEO**INTRODUCCIÓN:**

En esta tarea se expone al estudiante a distintos tipos de corpus con el fin de que aprenda a seleccionar aquél más apropiado según las necesidades de traducción que tenga.

DESARROLLO:

1. Los alumnos rellenan la Worksheet 2 ("Selecting the appropriate corpus"), que contiene preguntas referentes al uso o traducción de algunos términos. Estos términos están contextualizados en los textos que se facilitan y las preguntas se pueden resolver utilizando los corpus indicados en la propia Worksheet 2.
2. Las parejas se juntan con otras parejas y comentan sus hallazgos.

Dinámica de grupos: en parejas

Duración estimada: 2 horas

MATERIALES:**Fichas**

- Worksheet 2: "Selecting the appropriate corpus"

Textos originales y paralelos

- Textos en español que contengan los términos sobre los que se pregunta (Texto 1: "Vacío legal inexplicable"; Texto 2: "Perfil biofísico fetal"; Texto 5: "Euroblanca"; Texto 6: "El canto de Cabo Verde")

Descripciones de corpus

- Descripción 1: Corpus ad hoc de textos sobre aneuploidías extraídos del *Medline*
- Descripción 2: Corpus de textos sobre economía extraídos del *Financial Times*

Recursos electrónicos

- URL de diccionario bilingüe general [*Wordreference*: <http://wordreference.com> ; *Vox*: <http://www.diccionarios.com> ; *Babylon*: <http://www.lexicool.com/babylon-online-translation-dictionary.asp?IL=3>]
- URL de diccionario monolingüe general [*Onelook*: <http://www.onelook.com>]
- URL de base de datos terminológica o glosario bilingüe [*IATE*: <http://iate.europa.eu/> ; *AndyMiles*: <http://andymiles.com/>]
- URL de dudas lingüísticas del inglés resueltas con corpus [<http://www.eisu2.bham.ac.uk/Webmaterials/kibbitzers/kibbitzer06.htm>]
- URL de corpus de inglés general [*WordBanksOnline*: <http://www.collins.co.uk/Corpus/CorpusSearch.aspx> ; *BNC*: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- URL de corpus de correspondencia comercial [*Business Letter Corpus*: <http://ysomeya.hp.infoseek.co.jp/>]
- URL de corpus paralelo [*COMPARA*: <http://www.linguateca.pt/COMPARA/>]
- Concordancias ["conc_doppler.bmp"] ya extraídas de un corpus ad hoc de textos médicos sobre aneuploidías ("aneuploidies_9000.txt")
- Colocaciones ["colloc_interest_FT.bmp"] ya extraídas de un corpus de textos periodísticos financieros en inglés (*Financial Times*)

TAREA 3: FRECUENCIAS**INTRODUCCIÓN:**

Esta tarea busca hacer reflexionar al estudiante sobre el significado de la frecuencia de aparición de palabras o expresiones en distintos corpus.

DESARROLLO:

1. Los alumnos rellenan la Worksheet 3 ("Interpreting frequencies"), que contiene preguntas referentes a la frecuencia de uso de determinadas palabras o expresiones resaltadas y contextualizadas en los textos que se les facilitan.
2. A la vez, se facilitan los corpus de los que pueden disponer para realizar este ejercicio y que ayudarán a los alumnos a observar el aspecto de la frecuencia de aparición de dichas palabras.
3. También se facilita el material de apoyo al programa *WordSmith Tools*.
4. Las parejas se juntan con otras parejas y comentan sus hallazgos.

Dinámica de grupos: en parejas

Duración estimada: 2 horas

MATERIALES:

Fichas

- Worksheet 3: "Interpreting frequencies"

Textos originales y paralelos

- Textos en español e inglés que contengan los términos sobre los que se pregunta (Texto 1: "Vacío legal inexplicable"; Texto 3: "Exploración ecográfica"; Texto 4: "The future is yummy"; Texto 7: "Clínicas veterinarias")

Textos de apoyo

- 2: Material de apoyo al programa *WordSmith Tools*

Descripciones

- Descripción 3: Corpus ad hoc de instrucciones dadas por doctores a pacientes en inglés
- Descripción 4: Corpus ad hoc sobre inteligencia emocional en inglés y español

Recursos electrónicos

- URL de diccionario de dudas del español [*Panhispanico*: <http://www.rae.es>]
- URL de diccionario monolingüe general [*Onelook*: <http://www.onelook.com>]
- URL de diccionario de medicina en inglés [*Medical Dictionary*: <http://www.medical-dictionary.com>]
- URL de buscador general [*Google*: <http://www.google.com>]
- URL de herramienta de comparación de frecuencias en la web [*Googlefight*: <http://googlefight.com/>]
- URL de corpus de español general [*CREA*: <http://www.rae.es>]
- URL de corpus de inglés general [*WordBanksOnline*: <http://www.collins.co.uk/Corpus/CorpusSearch.aspx>]; [*BNC*: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- URL de corpus de inglés americano [*Brown Corpus*: http://www.lexutor.ca/concordancers/concord_e.html]
- URL de corpus de traducciones en inglés [*TEC*: <http://ronaldo.cs.tcd.ie/tec/jnlp/>]
- URL de archivo en línea de diario británico [*Guardian Unlimited*: <http://www.guardian.co.uk/Archive/>]
- Corpus ad hoc de aproximadamente 9000 resúmenes de artículos médicos sobre aneuploidías en inglés extraído del *Medline*
- Corpus ad hoc de unas 66.000 palabras con instrucciones dadas por doctores a pacientes extraído de páginas web
- Corpus ad hoc sobre inteligencia emocional de unas 112.000 palabras en español y unas 89.000 palabras en inglés
- Programa informático *WordSmith Tools*

TAREA 4: MIRAR A LA IZQUIERDA, A LA DERECHA Y A LA IZQUIERDA OTRA VEZ

INTRODUCCIÓN:

En esta tarea se pone énfasis en la importancia del co(n)texto para extraer información. Observando concordancias extraídas de distintos corpus ad hoc, el alumno practicará la reordenación del contexto dependiendo de la información requerida.

DESARROLLO:

1. Los alumnos leen la Worksheet 4a ("Look right, left and right again") en formato *PowerPoint*, con preguntas referentes a las colocaciones, uso o traducción de algunos términos en cuestión, y responden en la Worksheet 4b ("Look right, left and right again"), en formato *MSWord*.
2. Se facilitan los textos con los términos contextualizados.
3. Se facilitan los corpus de los que pueden disponer para realizar este ejercicio y que ayudarán a los alumnos a practicar la reordenación del contexto en unas concordancias.
4. Las parejas se juntan con otras parejas y comentan sus hallazgos.

Dinámica de grupos: en parejas

Duración estimada: 2 horas

MATERIALES:

Fichas

- Worksheet 4a: "Look right, left and right again"
- Worksheet 4b: "Look right, left and right again"
- Ficha de autoevaluación
- Ficha de valoración de la unidad

Recursos electrónicos

- URL de corpus de inglés general [*WordBanksOnline*: <http://www.collins.co.uk/Corpus/CorpusSearch.aspx>]
- Corpus ad hoc comparable bilingüe (inglés y español) de unas 89.000 y 112.000 palabras respectivamente sobre inteligencia emocional
- Corpus ad hoc de aproximadamente 9000 resúmenes de artículos médicos sobre aneuploidías en inglés extraído del *Medline*
- Programa informático *WordSmith Tools*

❸ EVALUACIÓN DE LA UNIDAD DIDÁCTICA

1. Valoración realizada por los estudiantes

a) Valoración de su aprendizaje: ficha de autoevaluación

A continuación se recogen en una tabla las puntuaciones concedidas por los estudiantes a sí mismos en relación con las competencias adquiridas y, más adelante, su análisis.

FICHA DE AUTOEVALUACIÓN: RESULTADOS								
UNIDAD DIDÁCTICA "USAR CORPUS PARA CONCIENCIAR SOBRE EL FUNCIONAMIENTO DE LA TRADUCCIÓN INVERSA AL INGLÉS"								
Estudiantes		A.M.	C.S.M.	J.P.	L.V.	M.I.	P.L.P.	Media de cada pregunta
Soy capaz de...	ver la naturaleza repetitiva y predecible de la lengua	8	8	9	9	7	8	8,2
	ser crítico con el valor de la intuición del hablante nativo	8	8	9	8	8	8	8,2
	manejar distintos (tipos de) corpus que están disponibles en línea (de la lengua general y especializada)	7	7	9	8	9	7	7,8
	seleccionar el corpus apropiado según las necesidades de traducción	8	7	9	8	8	6	7,7
	interpretar las frecuencias de aparición	9	7	9	7	8	6	7,7
	reordenar el contexto de una concordancia según me interese mirar a la derecha o a la izquierda de la expresión de búsqueda	8	9	9	8	9	6	8,2
	trabajar individualmente en la traducción del español al inglés	8	6	9	8	9	8	8
	trabajar en pareja / en grupo en la traducción del español al inglés	7	9	9	8	9	8	8,3
	Media de cada alumno	7,8	7,6	9	8	8,4	7,1	
Te ha ayudado el corpus a ...	tener más conciencia sobre la traducción del español al inglés (inversa)	7	9	9	8	9	8	8,3
	traducir mejor	8	9	8	8	9	7	8,2
	encontrar los equivalentes de traducción	8	9	8	8	9	7	8,2
	estar más seguro de la solución	9	9	8	8	8	7	8,2
	comprender mejor el texto original	7	8	7	8	9	7	7,7
	comprender la importancia de las colocaciones (palabras que suelen acompañar a otras)	9	9	9	8	8	8	8,5
	ser más consciente de la naturaleza de los textos utilizados en la fase de documentación (origen, fecha, original o traducción, etc.)	9	6	8	7	8	7	7,5
Media de cada alumno	8,1	8,4	8,1	7,8	8,6	7,3		

Tabla 24. Resumen de las Fichas de autoevaluación de la UD "Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés".

- Puntuación media de las respuestas proporcionadas por cada alumno. Como se puede observar, el nivel de confianza de los estudiantes respecto a lo adquirido durante la unidad es bastante alto, con puntuaciones que van desde el 7,1 hasta el 9 en cuanto a sus capacidades.
- Puntuación media de las respuestas en cada pregunta. En este caso destaca la uniformidad de las puntuaciones, que se mueven entre el 7,5 y el 8,5, de las que merece la pena resaltar la máxima nota, que los alumnos conceden a la concienciación sobre la importancia que tienen las colocaciones, elemento de suma importancia para alcanzar naturalidad en una traducción inversa.

b) Valoración de la unidad didáctica: ficha de valoración

Los comentarios expresados por los estudiantes en las fichas de valoración, disponibles en el *Volumen 2*, dejan traslucir un alto grado de satisfacción con la unidad didáctica. Destacaríamos, sin embargo, el hecho de que algún alumno hubiera deseado profundizar más en el uso de corpus para la traducción inversa, es decir, caminar hacia un estadio más avanzado que contemplara incluso la propia construcción de corpus y un trabajo más independiente. En este sentido, sólo cabe recordar que es ésta una unidad de iniciación al uso de corpus y de iniciación a la traducción inversa, donde lo que se busca es establecer algunos conceptos y habilidades fundamentales guiando de cerca al estudiante.

2. Valoración realizada por el docente

a) Valoración del aprendizaje de los estudiantes

En esta unidad didáctica se ha optado por no incluir una tarea final que incluya una evaluación integrada de las competencias adquiridas a lo largo de la unidad. En vez de esta modalidad, la evaluación se ha realizado a través de la observación de la marcha de las tareas y la corrección de las Fichas o Worksheets 1, 2, 3 y 4, y sólo podemos añadir que los resultados han sido excelentes.

b) Valoración de la unidad didáctica

Exceptuando algunas pequeñas modificaciones en la redacción de algún enunciado, la unidad didáctica se ha mantenido bastante cercana a su versión inicial. Como se verá en la Evolución de la unidad, ésta se probó con tres grupos de estudiantes con diferentes lenguas de partida,

el portugués y el español, y la misma lengua de llegada, el inglés. A pesar de este hecho, y dejando constancia de que para cada grupo se buscaron ejemplos específicos en su lengua de partida, la unidad didáctica funcionó perfectamente y permitió que tanto unos como otros se iniciaran en el uso de corpus para traducir y en la traducción inversa al inglés.

④ EVOLUCIÓN DE LA UNIDAD DIDÁCTICA

Modificaciones efectuadas

Éstos son los principales cambios que se realizaron en Fichas o Tareas tras la puesta en práctica de la unidad:

- Tarea 1 (“Desde la perspectiva de un nativo”): Si bien es cierto que los alumnos con los que se probó esta unidad didáctica estaban todos en un nivel inicial tanto de uso de corpus como de traducción inversa, quedó patente que algunos de ellos requerían más apoyo o deseaban más práctica, razón por la cual se elaboró la Worksheet 1c, con ejercicios que insistían y ampliaban lo visto en esta tarea.
- Tarea 3 (“Frecuencias”): Una de las observaciones realizadas tras la primera prueba de la unidad fue la necesidad de aclarar conceptos y practicar el manejo y evaluación de frecuencias de aparición de palabras o expresiones, ya que constituían para los alumnos un argumento de peso, muchas veces erróneo, a la hora de justificar una solución de traducción (p.ej. “he optado por la traducción X porque vi que aparecía con más frecuencia que otras soluciones”). Este hecho nos llevó a incluir más ejercicios dentro de esta tarea, dedicada precisamente a concienciar sobre el significado y la relatividad de la frecuencia de uso en un corpus.



UNIDAD DIDÁCTICA

El corpus ad hoc como remedio para la traducción médica al inglés



Con esta unidad didáctica se pretende concienciar al alumno de que un corpus de textos paralelos es imprescindible para resolver algunos aspectos de la traducción inversa de textos especializados ya que los recursos estrictamente lexicográficos (diccionarios, glosarios), aunque especializados, no resuelven todos los tipos de problemas de traducción que puede plantear la traducción de un texto.



Grupos de prueba: Esta unidad se probó en un grupo numeroso de alumnos (50 alumnos aprox.) de 4º curso de la licenciatura de Traducción e Interpretación de la UAB, en una asignatura de traducción especializada inversa (español-inglés B), durante el curso 2004-05.

Los materiales de esta unidad se encuentran en el *Volumen 2* de esta tesis (págs. 615-638).



❶ PRESENTACIÓN DE LA UNIDAD DIDÁCTICA

“El corpus ad hoc como remedio para traducir textos médicos al inglés”

NIVEL DE USO DE CORPUS

- Consolidación

MATERIA

- Traducción especializada médica inversa (español A – inglés B)

OBJETIVOS

Objetivos relacionados con el uso de corpus

- Crear un corpus comparable (monolingüe)
- Utilizar software y recursos de ubicación y descarga de textos electrónicos desde Internet
- Utilizar funciones avanzadas de un programa de análisis de corpus
- Resolver problemas de traducción utilizando corpus a nivel avanzado

Objetivos relacionados con la materia

- Aprender y mejorar el léxico
- Aprender y mejorar la redacción
- Captar la clase de información que aporta cada recurso de documentación
- Utilizar recursos de documentación de todo tipo para resolver problemas de traducción especializada inversa (buscadores, diccionarios, glosarios, tesauros, bases de datos, bibliotecas digitales)
- Crear recursos de documentación para resolver problemas de traducción especializada inversa (corpus)
- Traducir textos médicos del español al inglés

CONTENIDOS

- El tamaño de un corpus (I. 2.1.5.)
- La representatividad de un corpus (I. 2.1.6.)
- Buscadores de textos de especialidad (*Medline, Scirus*) (II. 1.3.3.)
- Importancia de nombrar lógicamente los archivos de un corpus (II. 2.1.)
- Creación de un corpus comparable monolingüe (ad hoc) (II. 3.1.1.)
- Extracción de vocabulario general y términos de especialidad (IV. 2.2.1.)
- Extracción de colocaciones (IV. 2.2.2.)
- Extracción de convenciones de género (IV. 2.3.1.)



ESTRUCTURA

- Tarea 1: "Primer acercamiento al texto original"
- Tarea 2: "Resolver la comprensión del tema. Cuando está escrito en nuestro idioma y aún así no lo entendemos"
- Tarea 3: "Resolver el vocabulario específico del tema"
- Tarea 4: "Resolver el vocabulario no específico del tema"
- Tarea 5: "Resolver el resto de la traducción. Problemas que van más allá del nivel del vocabulario"

EVALUACIÓN

Indicadores

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación,...

- Indicador 1: crea un corpus comparable monolingüe de calidad
- Indicador 2: utiliza funciones de un programa de análisis de corpus para traducir
- Indicador 3: introduce búsquedas idóneas en el corpus
- Indicador 4: resuelve la traducción de elementos de la lengua general y especializada

Fichas de evaluación

- Informe breve sobre la utilidad de los corpus para traducir un texto

MATERIALES

Fichas

- Ficha 1: "Análisis del TO"
- Ficha 2: "Identificación de problemas potenciales y análisis de la necesidad de recursos"
- Ficha 3: "El TO no es el enemigo, sino la primera fuente de información"
- Ficha 4: "Lo que sí / no ha conseguido solucionar cada tipo de recurso"
- Ficha 5: "Ficha de búsqueda y recogida de textos paralelos (TPs) en inglés"
- Ficha 6: "Ejemplos de cómo utilizar un corpus de TPs para resolver problemas de género, colocaciones y uso"

Texto original y paralelos

- Texto 1: "Perfil biofísico fetal"
- Texto 2: "Estudio biométrico"
- Texto 3: "Resultados clínicos"
- Texto 4: "Exploración ecográfica"

Recursos electrónicos

- URLs de diccionarios monolingües / bilingües / multilingües generales [*WordReference*: <http://wordreference.com> ; *Collins*: <http://diccionario.reverso.net/> ; *Onelook*: <http://www.onelook.com> ; *ForeignWord*: <http://www.foreignword.com> ; *Vox*: <http://www.diccionarios.com>]
- URLs de diccionarios / tesauros monolingües generales [*Wordsmyth*: <http://www.wordsmyth.net> ; *Thesaurus*: <http://thesaurus.reference.com/> ; *Dictionary*: <http://dictionary.reference.com> ; *Hyperdictionary*: <http://www.hyperdictionary.com>]



- URLs de diccionarios / glosarios multilingües / bilingües especializados [*Multilingual Glossary of technical and popular medical terms*: <http://allserv.rug.ac.be/~rvdstich/eugloss/welcome.html> ; IATE: <http://iate.europa.eu> ; Termcat: <http://www.termcat.net>]
- URL de tesoro monolingüe especializado [*Medical Dictionary*: <http://www.medical-dictionary.com>]
- URLs de enciclopedias [*Enciclopedias*: <http://www.enciclopedias.com> ; *Medical Encyclopedia*: <http://www.umm.edu/ency/index.html> ; *Encarta*: <http://es.encarta.msn.com/>]
- URLs de buscadores generales [*Google*: <http://www.google.com> ; *AllTheWeb*: <http://www.alltheweb.com>]
- URL de un buscador de buscadores especializados [*Buscopio*: <http://www.buscopio.net>]
- URL de buscador de acrónimos [*AcronymFinder*: <http://www.acronymfinder.com>]
- URLs de buscadores de TPs de tema científico [*PubMed*: <http://www4.ncbi.nlm.nih.gov/entrez/query.fcgi> ; *Scirus*: <http://www.scirus.com>]
- URL de corpus de español general [CREA: <http://www.rae.es>]
- URL de corpus de inglés general [BNC: <http://sara.natcorp.ox.ac.uk/lookup.html>]
- Programa informático *WordSmith Tools*

DURACIÓN ESTIMADA

13 horas

2 DESARROLLO DE LAS TAREAS

TAREA 1: PRIMER ACERCAMIENTO AL TEXTO ORIGINAL (TO)

INTRODUCCIÓN:

En esta tarea se busca que los alumnos hagan un primer análisis del texto original (Texto 1: "Perfil biofísico fetal") e identifiquen los puntos problemáticos de su traducción y los recursos que necesitarían para solucionarlos.

DESARROLLO:

1. Se facilita el TO (Texto 1: "Perfil biofísico fetal") a los alumnos para que lo lean junto con dos textos paralelos (TPs) (Texto 2: "Estudio biométrico" y Texto 3: "Resultados clínicos").
2. Se divide la clase en grupos de 3 ó 4 personas y rellenan la Ficha 1 ("Análisis del TO") con las características principales del TO y el encargo.
3. Se procede a un intercambio de fichas entre grupos para observar las respuestas de los demás.
4. Los mismos grupos rellenan la Ficha 2 ("Identificación de problemas potenciales y análisis de la necesidad de recursos") con los tipos de problemas de traducción potenciales que aprecian en el TO y los tipos de recursos de documentación que creen que les ayudarían a resolverlos.
5. Se procede a un intercambio de fichas entre grupos para observar las respuestas de los demás.
6. Los alumnos hacen una 1ª versión de la traducción el TO individualmente y con los medios que ellos consideren oportunos para entregarla antes de la siguiente sesión.

Dinámica de grupos: en grupos de 3 ó 4 e individual

Duración estimada: 3 horas y media

MATERIALES:

Fichas

- Ficha 1: "Análisis del TO"
- Ficha 2: "Identificación de problemas potenciales y análisis de la necesidad de recursos"

Texto original y paralelos

- Texto original breve (250-275 palabras) en español de tema científico (Texto 1: "Perfil biofísico fetal")
- Textos paralelos extraídos de la misma publicación que el TO (Texto 2: "Estudio biométrico" y Texto 3: "Resultados clínicos")

EVALUACIÓN:

Cada grupo valorará y completará sus propias fichas al intercambiarlas con otro grupo. Además, tanto la Ficha 1 como la Ficha 2 se comentarán en voz alta para asegurarse de que todos tienen las mismas respuestas.

COMENTARIOS:

- Se debe garantizar que, tras la puesta en común de la Ficha 2, ésta contiene en su versión final todos los problemas que debe recoger y todos los posibles recursos para solucionarlos.
- Es necesario avisar a los alumnos de que, tras la 1ª traducción, se pedirá una 2ª versión mejorada y, finalmente, una 3ª versión definitiva de la traducción del TO.

TAREA 2: RESOLVER LA COMPRENSIÓN DEL TEMA. CUANDO ESTÁ ESCRITO EN NUESTRO IDIOMA Y AÚN ASÍ NO LO ENTENDEMOS

INTRODUCCIÓN:

En esta tarea se trabaja la comprensión del TO (Texto 1: "Perfil biofísico fetal"); se empiezan a recoger los primeros textos paralelos y se aprende sobre el tema. También se empieza a hacer una valoración de la utilidad de los distintos tipos de recursos de documentación.

DESARROLLO:

1. Una vez que los alumnos han hecho la 1ª versión de la traducción y se han enfrentado a los problemas reales del texto, se les guía para que vayan solucionándolos paso a paso. Se les insta a resolver los distintos tipos de problemas uno a uno, empezando por la comprensión del TO y el tema que trata.
2. Se les facilita la Ficha 3 ("El TO no es el enemigo, sino la primera fuente de información") para que, en lengua de partida, entiendan el significado del concepto principal del TO y tengan una visión más clara y estructurada del contenido del mismo. Se les sugieren algunas *URLs* de buscadores y enciclopedias en línea.
3. Cada alumno revisa la ficha de un compañero.
4. Se les anima a que, si encuentran algún documento especialmente relevante, lo guarden en una carpeta con un nombre significativo.
5. Se les facilita la Ficha 4 ("Lo que sí / no ha conseguido solucionar cada tipo de recurso") para que apunten lo que cada tipo de recurso les ha ayudado a solucionar y lo que todavía les falta por resolver.
6. Los alumnos tendrán que resaltar en amarillo cualquier cambio que realicen sobre la 1ª versión de su traducción.

Dinámica de grupos: individual

Duración estimada: 1 hora y media

MATERIALES:**Fichas**

- Ficha 3: "El TO no es el enemigo, sino la primera fuente de información"
- Ficha 4: "Lo que sí / no ha conseguido solucionar cada tipo de recurso"

Texto original

- Texto 1: "Perfil biofísico fetal"

Recursos electrónicos

- URLs de buscadores generales [*Google*: <http://www.google.com> ; *AllTheWeb*: <http://www.alltheweb.com>]
- URLs de enciclopedias [*Enciclopedias*: <http://www.enciclopedias.com> ; *Medical Encyclopedia*: <http://www.umm.edu/ency/index.html> ; *Encarta*: <http://es.encarta.msn.com/>]
- URL de un buscador de buscadores especializados [*Buscopio*: <http://www.buscopio.net>]

EVALUACIÓN:

En cuanto a la Ficha 3, al revisar la ficha de un compañero, cada alumno valorará y completará su propia ficha. Una de las conclusiones a la que debe llegar el alumno tras hacer la Ficha 3 es que el TO no es el enemigo a batir, sino la primera fuente de información con la que cuentan al empezar a traducir.

COMENTARIOS:

- Con excepción de la Tarea 1 en la que el trabajo de 1ª aproximación al TO se lleva a cabo en grupo, el resto del trabajo a partir de aquí será individual, si bien es cierto que habrá intercambio de fichas, puesta en común de soluciones, etc.
- Se debe insistir en que más adelante se les pedirá que entreguen una 2ª versión mejorada de la traducción y, después, una 3ª y última versión. Es necesario que guarden las 3 versiones por separado.

TAREA 3: RESOLVER EL VOCABULARIO ESPECÍFICO DEL TEMA

INTRODUCCIÓN:

Se continúa en el camino de solucionar los problemas encontrados en el TO (Texto 1: "Perfil biofísico fetal") y, para ello, en esta tarea se abordan recursos para resolver la traducción de elementos específicos del tema, en especial, términos y acrónimos.

DESARROLLO:

1. Se facilitan *URLs* de diccionarios y glosarios multilingües, bilingües y monolingües especializados, así como *URLs* de tesauros especializados. También se les introduce al uso de un buscador de acrónimos, previendo que no conocen este tipo de recurso.
2. Los alumnos utilizan estos recursos u otros del mismo tipo para resolver las dudas de traducción de tipo terminológico de las que no quedaron del todo seguros en la 1ª versión de su traducción. Se recuerda que deben resaltar en amarillo los cambios que hagan.
3. Recuperan la Ficha 4 ("Lo que sí / no ha conseguido solucionar cada tipo de recurso") para continuar rellenándola según avanzan en el uso de recursos.

Dinámica de grupos: individual

Duración estimada: 2 horas

MATERIALES:

Fichas

- Ficha 4: "Lo que sí / no ha conseguido solucionar cada tipo de recurso"

Texto original

- Texto 1: "Perfil biofísico fetal"

Recursos electrónicos

- *URLs* de diccionarios / glosarios multilingües / bilingües especializados [*Multilingual Glossary of technical and popular medical terms*: <http://allserv.rug.ac.be/~rvdstich/eugloss/welcome.html> ; *IATE*: <http://iate.europa.eu> ; *Termcat*: <http://www.termcat.net>]
- *URL* de diccionario monolingüe general [*Onelook*: <http://www.onelook.com>]
- *URL* de tesoro monolingüe especializado [*Medical Dictionary*: <http://www.medical-dictionary.com>]
- *URL* de buscador de acrónimos [*AcronymFinder*: <http://www.acronymfinder.com>]

COMENTARIOS:

- Se atrae la atención sobre el hecho de que quizás el diccionario bilingüe especializado no soluciona todos los problemas terminológicos del texto porque no contiene todos los términos que buscamos. Previendo que el diccionario bilingüe sería uno de los primeros recursos utilizados por los alumnos, el hecho de no encontrar un término sería positivo ya que les forzaría a recurrir a otros medios.

TAREA 4: RESOLVER EL VOCABULARIO NO ESPECÍFICO DEL TEMA

INTRODUCCIÓN:

En esta tarea se abordan recursos que preparan al alumno para resolver la traducción de elementos no específicos del tema, es decir, palabras y expresiones de la lengua general que ocasionan muchos errores en la traducción inversa especializada.

DESARROLLO:

1. Se facilitan *URLs* de diccionarios y tesauros multilingües, bilingües y monolingües generales, así como *URLs* de corpus monolingües de español e inglés general.
2. Los alumnos utilizan estos recursos u otros del mismo tipo para resolver las dudas de traducción relativas a la lengua general de las que no quedaron del todo seguros en la 1ª versión de su traducción.
3. Recuperan la Ficha 4 ("Lo que sí / no ha conseguido solucionar cada tipo de recurso") para continuar rellenándola según avanzan en el uso de recursos.
4. Se pide que hagan una 2ª versión de la traducción individualmente con los recursos puestos a su alcance para entregarla antes de la siguiente sesión. Todos los cambios realizados desde la 1ª versión deberán estar marcados en amarillo.

Dinámica de grupos: individual

Duración estimada: 2 horas

MATERIALES:

Fichas

- Ficha 4: "Lo que sí / no ha conseguido solucionar cada tipo de recurso"

Texto original

- Texto 1: "Perfil biofísico fetal"

Recursos electrónicos

- *URLs* de diccionarios monolingües / bilingües / multilingües generales [*WordReference*: <http://wordreference.com> ; *Collins*: <http://diccionario.reverso.net/> ; *Onelook*: <http://www.onelook.com> ; *ForeignWord*: <http://www.foreignword.com> ; *Vox*: <http://www.diccionarios.com>]
- *URLs* de diccionarios / tesauros monolingües generales [*Wordsmyth*: <http://www.wordsmyth.net> ; *Thesaurus*: <http://thesaurus.reference.com/> ; *Dictionary*: <http://dictionary.reference.com> ; *Hyperdictionary*: <http://www.hyperdictionary.com>]
- *URL* de corpus de español general [*CREA*: <http://www.rae.es>]
- *URL* de corpus de inglés general [*BNC*: <http://sara.natcorp.ox.ac.uk/lookup.html>]

TAREA 5: RESOLVER EL RESTO DE LA TRADUCCIÓN. PROBLEMAS QUE VAN MÁS ALLÁ DEL NIVEL DEL VOCABULARIO

INTRODUCCIÓN:

Esta tarea es clave en tanto que se concentra en la creación por parte del alumno de un corpus comparable monolingüe que debe servirle para resolver los problemas de traducción que van más allá del nivel del término en su TO (Texto 1: "Perfil biofísico fetal"). En este ejercicio se observa la necesidad de no sólo asegurar la calidad de los TPs que componen el corpus, sino también la cantidad.

DESARROLLO:

1. Haciendo referencia a la Ficha 4 ("Lo que sí / no ha conseguido solucionar cada tipo de recurso"), se recuerdan los tipos de problemas que quedan sin solucionar tras haber utilizado ya diversos recursos: problemas de convenciones del género, de colocaciones y de uso. La única solución para resolverlos es acudir a TPs.
2. Se introduce a los alumnos al uso de los buscadores de textos científicos *PubMed* y *Scirus*, previendo que no conocen este tipo de recurso. Se pone énfasis en *PubMed*, ya que permite la recogida de textos en grandes cantidades en poco tiempo a través de búsquedas por autor, revista o palabras clave.
3. Se sugiere que, durante 15 minutos, hagan una búsqueda cuidada de textos, seleccionando aquellos que consideran más útiles o de mejor calidad para ayudarles en su traducción.
4. A la vez, se les facilita la Ficha 5 ("Ficha de búsqueda y recogida de textos paralelos (TPs) en inglés") para que lleven un control del contenido y número de textos que están incluyendo en su corpus. Se indica que los textos deben guardarse en una carpeta individual en una unidad del ordenador y que los archivos deben tener nombres identificativos de su contenido (p.ej. "biophysical_profile_500.txt").
5. Con ayuda del programa de análisis de corpus *WordSmith Tools*, los alumnos realizan algunas búsquedas variadas relacionadas con su TO (Texto 1: "Perfil biofísico fetal") para extraer concordancias utilizando este corpus inicial.
6. Discuten con los compañeros más próximos los problemas que les plantea trabajar con este primer corpus y cómo los solucionarían.
7. Como el problema principal habrá sido el reducido tamaño del corpus, los alumnos deben introducir en *PubMed* muchas más búsquedas en inglés de aquellos términos más representativos del TO cuya traducción ya han solucionado con los recursos anteriores. También pueden apoyarse en los TPs facilitados en la Tarea 1 para seleccionar términos relacionados con el tema que les ocupa.
8. Recogen un gran número de TPs para crear un corpus ad hoc amplio y, a la vez, continúan rellenando la Ficha 5.
9. Se ilustra la forma como pueden utilizar un corpus ad hoc de TPs para resolver problemas que van más allá del vocabulario. Para ello, se les facilita otro TP (Texto 4:

“Exploración ecográfica) y otro corpus ya construido a partir del *PubMed* por la profesora (“aneuploid_9000.txt”) junto con su descripción (Descripción 1).

10. Se les facilita la Ficha 6 (“Ejemplos de cómo utilizar un corpus de TPs para resolver problemas de género, colocaciones y uso”) con indicaciones para que descubran la solución a un problema de colocación, otro de género y otro de uso.
11. Hacen una 3ª y última versión de la traducción individualmente ayudándose del recurso que han creado (su propio corpus ad hoc) para entregarla antes de la siguiente sesión. Todos los cambios realizados desde la 2ª versión deberán estar resaltados en azul.
12. Una vez han acabado con la traducción, terminan de cumplimentar la Ficha 4 (“Lo que sí / no ha conseguido solucionar cada tipo de recurso”).
13. Escriben un informe en el que tienen que explicar 5 problemas de traducción variados que hayan solucionado con el corpus creado y comentar la utilidad de la construcción y uso de dicho corpus en este encargo.

Dinámica de grupos: individual

Duración estimada: 4 horas

MATERIALES:

Fichas

- Ficha 4: “Lo que sí / no ha conseguido solucionar cada tipo de recurso”
- Ficha 5: “Ficha de búsqueda y recogida de textos paralelos (TPs) en inglés”
- Ficha 6: “Ejemplos de cómo utilizar un corpus de TPs para resolver problemas de género, colocaciones y uso”

Texto original y paralelos

- Texto 1: “Perfil biofísico fetal”
- Textos paralelos (Texto 2: “Estudio biométrico”; Texto 3: “Resultados clínicos”; Texto 4: “Exploración ecográfica”)

Recursos electrónicos

- URLs de buscadores de TPs de tema científico [*PubMed*: <http://www4.ncbi.nlm.nih.gov/entrez/query.fcgi> ; *Scirus*: <http://www.scirus.com>]
- Programa informático *WordSmith Tools*

EVALUACIÓN:

La Ficha 4, la Ficha 5 y el informe con los 5 problemas se han de entregar junto con la traducción final.

En la sesión siguiente, una vez entregados todos los documentos (traducción + Fichas 4 y 5 + informe) y revisados por la profesora los comentarios realizados al respecto de la utilidad del corpus ad hoc, se establecerá una discusión guiada sobre la Unidad didáctica: qué han aprendido, qué no ha quedado claro, pros y contras del uso de corpus y el resto de recursos.

COMENTARIOS:

Se debe insistir en que es posible, e incluso conveniente, que aumenten el corpus con más textos a medida que vayan necesiéndolo para realizar su traducción.

③ EVALUACIÓN DE LA UNIDAD DIDÁCTICA

1. Valoración realizada por los estudiantes

a) Valoración de su aprendizaje y de la unidad didáctica

De forma excepcional, al final de la presente unidad no se facilitó a los alumnos un cuestionario de autoevaluación y otro de valoración de la unidad didáctica, sino que se prefirió permitirles expresar su apreciación sobre la utilidad de la construcción y el uso de un corpus ad hoc para el encargo de traducción del Texto 1 (“Perfil biofísico fetal”). Los comentarios *in extenso* de los alumnos, incluidos dentro del informe que se les solicitó como ejercicio final de la unidad, se pueden encontrar en el *Volumen 2* de la presente tesis. Entre ellos se han identificado los siguientes temas recurrentes:

- **Calidad de la traducción resultante:** el 100% de los alumnos coinciden en que el corpus les ha ayudado a mejorar la calidad de su traducción, no sólo en cuestiones terminológicas, sino también las relacionadas con el uso, el orden y la interrelación de los términos, la sintaxis y las convenciones del género, en este caso del artículo científico. Sin embargo, los estudiantes también son capaces de entender que quizás un corpus no puede resolver todos los problemas de traducción que les plantee un texto ya que, por ejemplo, si no se tienen claros los conceptos médicos y no se comprende el texto original por completo, se podrían encontrar perdidos ante la gran cantidad de información que ofrece el corpus y de formas posibles de llegar hasta ella.
- **Calidad y cantidad de textos paralelos consultados:** dadas las características de la fuente (*Medline*) de donde los alumnos han extraído la práctica totalidad de los textos paralelos para sus corpus, éstos han podido disponer de información fiable sobre el tema en cuestión, el género, el estilo de redacción médica, la terminología, etc. Por otro lado, los alumnos han tomado conciencia no sólo sobre la importancia de la calidad de los textos consultados, sino también de la cantidad, y sobre que, para obtener un corpus tanto bueno como extenso, necesitaban introducir en el *Medline* búsquedas bastante acertadas.
- **Tiempo:** a pesar de la inversión de tiempo que supone construir un corpus extenso y fiable, los alumnos han reconocido que, finalmente, contar con uno les ha supuesto un ahorro de tiempo al poder hacer búsquedas y obtener resultados muy rápidamente.

- **Seguridad:** en este caso, los comentarios de los alumnos dejan patente una paradoja respecto al uso de corpus en tanto que éste les ha servido para confirmar soluciones de las que no estaban del todo seguros, aunque en ocasiones, ante la abundancia de opciones de traducción, les ha planteado más dudas de las que tenían al principio, si bien es cierto que, en éste último caso, los alumnos deberían haber seguido documentándose hasta tener la certeza sobre el uso de dichas opciones.
- **Programa *WordSmith Tools*:** de este programa los alumnos destacan sobre todo su facilidad de uso y las diferentes posibilidades de realizar búsquedas (p.ej. a partir de sólo una de las palabras de un sintagma), si bien reconocen que deben ser cautos para no hacer una búsqueda demasiado amplia, pero tampoco restringirla demasiado. Además, apuntan las distintas formas de ver y utilizar los resultados que permite el programa y el hecho de que, gracias a esta presentación de los resultados, ha sido posible que una búsqueda relacionada con una frase o un término les condujera a otras búsquedas o incluso a solucionar otras dudas aún no planteadas.
- **Proceso:** sin dudarlo, los estudiantes reconocen que el corpus ha sido indispensable para resolver muchos de los problemas que planteaba el texto, si bien este recurso, en ocasiones, tiene que complementarse con otros.

2. Valoración realizada por el docente

a) Valoración del aprendizaje de los estudiantes y de la unidad didáctica

Si bien es cierto que la unidad didáctica puede parecer lenta al principio y antojarse larga, se ha constatado que los alumnos responden finalmente muy bien ante los descubrimientos que realizan gracias al trabajo con corpus, como se puede observar en sus comentarios resumidos en el apartado anterior.

Así, los alumnos comienzan el proceso de traducción del Texto 1 con “desasosiego” cuando no entienden el texto original o no encuentran los términos que buscan con los recursos que manejan normalmente, sentimiento que se aplaca al sugerirles otros recursos terminológicos más apropiados y útiles. Sin embargo, aún habiendo encontrado los términos, los alumnos suelen no estar seguros de “si se dice así” en inglés, razón por la cual se les muestra cómo pueden ganar en seguridad recogiendo miles de TPs en minutos y que, gracias a un programa de análisis de corpus, pueden ir

directamente a buscar aquello que les interesa, sin necesidad de leer todos los textos. Finalmente, aunque encuentran que puede haber varias opciones de traducción para los mismos términos, son capaces de confrontar su traducción con lo que encuentran en el corpus y solucionan la mayoría de sus dudas: terminológicas, colocacionales, de estilo y género. El corpus les proporciona mayor seguridad a la hora de tomar decisiones y entienden que no hubieran podido solucionar la traducción sin un corpus de TPs, objetivo principal de la unidad didáctica.

④ EVOLUCIÓN DE LA UNIDAD DIDÁCTICA

Modificaciones efectuadas

Todas las tareas presentadas se llevaron a cabo en el aula, pero no exactamente de la forma dispuesta aquí. Tras su puesta en práctica, algunas tareas y fichas han cambiado con el propósito de ser más claras y útiles.

- Tarea 1, Ficha 2 (“Identificación de problemas potenciales y análisis de la necesidad de recursos”): Se observó enseguida que el enunciado de esta ficha daba lugar a confusión, por lo que se añadió la siguiente nota: “No se pide el nombre del recurso concreto (p.ej. *wordreference), sino el tipo (p.ej. diccionario mono- y bilingüe)”.
- Tarea 2, Ficha 3 (“El TO no es el enemigo, sino la primera fuente de información”): tras comprobar que, por una falta de análisis del contenido del TO, muchos erraban al intentar utilizar un solo verbo en inglés para traducir “sometido” (a una historia, a una evaluación, a unas pruebas, etc.), o que traducían “test de estimulación” o “test de sobrecarga” por otros tipos de test, se amplió considerablemente esta ficha para procurar una lectura más atenta y analítica del TO.



UNIDAD DIDÁCTICA

Corpus y memoria de traducción



En esta unidad se pone mucho énfasis en el manejo de software de diversos tipos con el fin de crear un corpus paralelo que después pueda utilizarse a modo de memoria de traducción. El estudiante participa en todo el proceso, desde digitalizar o descargar textos, alinearlos, importarlos desde un programa de traducción asistida hasta explotarlos para una traducción científico-técnica extrayendo terminología, colocaciones o fraseología.



Grupos de prueba: La unidad didáctica tal y como se muestra aquí se puso en práctica con un grupo de 4º curso de la licenciatura en Traducción e Interpretación de la UAB en una asignatura de traducción especializada inversa (español A - inglés B). Tras la comprobación de su funcionamiento, partes de esta unidad han sido utilizadas con otro grupo de la asignatura mencionada, así como con un grupo muy reducido de alumnos de máster en traducción especializada. Las tareas, fichas y evaluación que se presentan aquí corresponden a las utilizadas en la primera puesta en práctica de la unidad didáctica completa, curso 2006-07.

Los materiales de esta unidad se encuentran en el *Volumen 2* de esta tesis (págs. 639-688).



❶ PRESENTACIÓN DE LA UNIDAD DIDÁCTICA

“Corpus y memoria de traducción”

NIVEL DE USO DE CORPUS

- Consolidación

MATERIA

- Traducción especializada científico-técnica inversa (español A – inglés B)

OBJETIVOS DE LA UNIDAD

Objetivos relacionados con el uso de corpus

- Crear un corpus comparable (monolingüe o bilingüe) o paralelo
- Utilizar software y recursos de ubicación y descarga de textos electrónicos desde Internet
- Utilizar software de conversión de textos de papel a formato electrónico
- Utilizar funciones avanzadas de un programa de alineamiento de corpus
- Utilizar corpus como memorias de traducción
- Integrar conocimientos para resolver problemas de traducción utilizando corpus a nivel avanzado

Objetivos relacionados con la materia

- Identificar la importancia de la naturalidad en el discurso especializado en L2 (colocaciones, fraseología)
- Utilizar recursos de documentación de todo tipo para resolver problemas de traducción especializada inversa (buscadores, diccionarios, glosarios, tesauros, bases de datos, bibliotecas digitales)
- Crear recursos de documentación para resolver problemas de traducción especializada inversa (corpus, glosarios, memorias de traducción)
- Traducir textos científico-técnicos del español al inglés

Objetivos relacionados con competencias genéricas

- Gestionar la información (habilidad para buscar y analizar información proveniente de fuentes diversas)

CONTENIDOS

- Corpus de inglés y español en Internet (*IULA*) (I. 1.1.4.)
- Software de conversión de textos en papel a formato electrónico (III. 1.1.)
- Software de conversión de formato PDF / imagen a texto (III. 1.2.)
- Software de descarga de textos (*WebCopier*) (III. 1.3.2.)
- Software de alineación de textos (*WinAlign* de Trados) (III. 1.5.)
- Software de conversión de un corpus paralelo en memoria de traducción (*Translator's Workbench* de Trados) (III. 1.6.)
- Función *Concordance* (de *Translator's Workbench* de Trados) (III. 2.1.4.)



- Extracción de vocabulario general y términos de especialidad (IV. 2.2.1.)
- Extracción de colocaciones (IV. 2.2.2.)
- Extracción de fraseología (IV. 2.2.3.)

ESTRUCTURA DE LA UNIDAD

- Tarea 1: "Del papel al formato electrónico"
- Tarea 2: "(Más) posibilidades de búsquedas en un corpus"
- Tarea 3: "Crear un corpus ad hoc: Descarga de documentos y sitios web"
- Tarea 4: "Crear un corpus ad hoc: Aprovechar los archivos PDF"
- Tarea 5: "De memoria de traducción a corpus paralelo"

EVALUACIÓN

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación,...

- Indicador 1: maneja software de conversión de textos de papel a formato electrónico
- Indicador 2: realiza búsquedas en archivos PDF como si fueran un corpus
- Indicador 3: maneja distintos (tipos de) corpus que están disponibles en línea
- Indicador 4: maneja funciones avanzadas de un programa de análisis de corpus
- Indicador 5: maneja software y recursos de ubicación y descarga de textos electrónicos desde Internet
- Indicador 6: maneja software de conversión de archivos .pdf
- Indicador 7: maneja software de alineación de textos
- Indicador 8: convierte un corpus paralelo en una memoria de traducción
- Indicador 9: maneja la función de extraer concordancias en un programa de traducción asistida para utilizar una memoria de traducción a modo de corpus
- Indicador 10: encuentra equivalentes terminológicos en textos científico-técnicos para traducir del español al inglés
- Indicador 11: encuentra las colocaciones de un término en textos científico-técnicos para traducir del español al inglés
- Indicador 12: encuentra fraseología en textos científico-técnicos para traducir del español al inglés

Fichas de evaluación

- Ficha de autoevaluación
- Ficha de valoración de la unidad

MATERIALES

Fichas

- Worksheet 1a: "From paper to electronic format"
- Worksheet 1b: "From paper to electronic format"
- Worksheet 2: "Searching a corpus by lemma and tags"
- Ficha 1: "Descargar un sitio o página web con el programa *WebCopier*"
- Ficha 2: "Convertir un archivo PDF a uno de texto"
- Ficha 3: "Alinear textos y convertirlos en una memoria de traducción"



- Ficha de autoevaluación
- Ficha de valoración de la unidad

Textos originales y paralelos

- Texto 1: "Seroteca"
- Texto 2: "Triturus analyzer"
- Mismo texto de tema y/o género relevante en español (Texto 3) y en inglés (Texto 4)

Textos de apoyo

- Documento 1: "Texto con errores tras el escaneado"

Recursos electrónicos

- URL de sitio web bilingüe relacionado con el tema del texto original que sirva como fuente de varios textos paralelos para la traducción que deben realizar [Empresa *Grifols*: <http://www.grifols.com/>]
- URL de corpus paralelo multilingüe sobre informática [IULA: <http://bwananet.iula.upf.edu/index.htm>]
- URL de un programa de conversión de PDF (gratis o demo) [TEXTfromPDF: <http://www.textfrompdf.com/tfpdownload.htm>]
- URL de un navegador *offline* (gratis) [WebCopier: <http://www.softcatala.org/prog84.htm>]
- Escáner
- Software de reconocimiento óptico de caracteres (OCP)
- Software de alineación de textos *WinAlign* de Trados
- Software de memoria de traducción *Translator's Workbench* de Trados

DURACIÓN ESTIMADA

7 horas

2 DESARROLLO DE LA UNIDAD DIDÁCTICA

TAREA 1: DEL PAPEL AL FORMATO ELECTRÓNICO

INTRODUCCIÓN:

Esta primera tarea busca que el estudiante aprenda y practique la conversión de textos en papel a formato electrónico con la finalidad principal de que pueda incorporar a un corpus electrónico aquellos textos paralelos de calidad que haya encontrado en formato papel.

DESARROLLO:

1. Los alumnos leen la Worksheet 1a ("From paper to electronic format") en formato *PowerPoint*, que contiene una explicación sobre cómo utilizar un OCP y algunas preguntas sobre el proceso de conversión del formato papel al electrónico. A la vez, responden en la Worksheet 1b ("From paper to electronic format"), en formato *MSWord*.

Dinámica de trabajo: individual

Duración estimada: 1 hora

MATERIALES:

Fichas

- Worksheet 1a: "From paper to electronic format"
- Worksheet 1b: "From paper to electronic format"

Textos originales y paralelos

- Texto de contenido científico-técnico en español en formato papel que, al escanearlo, dé problemas con algunos formatos (p.ej. tablas, caracteres como la "ñ" o vocales acentuadas) (Texto 1: "Seroteca")

Textos de apoyo

- Documento que muestra errores tras ser escaneado (Documento 1: "Texto con errores tras el escaneado")

Recursos electrónicos

- Software de reconocimiento óptico de caracteres (OCP)
- Escáner

EVALUACIÓN:

El alumno entrega un texto paralelo al texto original previamente escaneado y limpiado.

COMENTARIOS:

- Los alumnos pueden hablar entre ellos mientras que visualizan el archivo *PowerPoint* para comprobar sus respuestas.

TAREA 2: (MÁS) POSIBILIDADES DE BÚSQUEDAS EN UN CORPUS

INTRODUCCIÓN:

En esta tarea los alumnos avanzan en el uso de un corpus etiquetado. Aquí practican búsquedas más complejas que las realizadas hasta el momento con corpus como el *BNC*, por ejemplo.

DESARROLLO:

1. Los alumnos leen el texto original (Texto 1: "Seroteca") que se traducirá durante esta unidad.
2. Rellenan la Worksheet 2 ("Searching a corpus by lemma and tags"), que contiene ejercicios con búsquedas que deben hacer en un corpus especializado utilizando etiquetas y con el que ya resolverán algunos problemas de traducción del TO.
3. Se invita a los alumnos a discutir entre ellos sus hallazgos y las implicaciones del corpus que han utilizado.

Dinámica de trabajo: individual

Duración estimada: 1 hora

MATERIALES:

Fichas

- Worksheet 2: "Searching a corpus by lemma and tags"

Textos

- Texto original en español ya escaneado y limpiado (Texto 1: "Seroteca")

Recursos electrónicos

- URL de corpus paralelo multilingüe sobre informática [IULA: <http://bwananet.iula.upf.edu/index.htm>]

COMENTARIOS:

- Antes de empezar con el ejercicio, en el que se realizarán búsquedas en un corpus especializado en informática, es necesario hacer ver a los alumnos que el texto original no sólo trata un tema, el de la biotecnología, sino que contiene varias expresiones del campo de la informática de usuario, para cuya traducción nos ayudaremos del mencionado corpus.

TAREA 3: CREAR UN CORPUS AD HOC: DESCARGA DE DOCUMENTOS Y SITIOS WEB**INTRODUCCIÓN:**

En esta tarea se busca que el alumno maneje software y recursos de descarga de textos desde Internet con el fin de utilizarlos más adelante como un corpus electrónico.

DESARROLLO:

1. Los alumnos descargan e instalan el navegador *offline WebCopier*.
2. Utilizándolo, descargan un sitio web relevante como fuente de información para la traducción del texto original, siguiendo los pasos indicados en la Ficha 1 ("Descargar un sitio o página web con el programa *WebCopier*") y contestan a las preguntas que ésta contiene sobre la utilidad y el proceso de descarga de un sitio web.

Dinámica de trabajo: en parejas

Duración estimada: 1 hora

MATERIALES:**Fichas**

- Ficha 1: "Descargar un sitio o página web con el programa *WebCopier*"

Textos

- Texto 1 ("Seroteca")

Recursos electrónicos

- Página web útil como fuente de varios textos paralelos para la traducción que deben realizar [Empresa *Grifols*: <http://www.grifols.com/>]
- URL de un navegador *offline* (gratuito) [*WebCopier*: <http://www.softcatala.org/prog84.htm>]

COMENTARIOS:

- Para repartirse el trabajo entre cada pareja, uno puede ayudar con los pasos que hay que seguir y el otro ejecutar las instrucciones en el programa de descarga.
- La Ficha 1 que aparece aquí, y las Fichas 2 y 3 que aparecerán más adelante están en español y no en inglés para que sean perfectamente integrables en unidades de traducción directa.

TAREA 4: CREAR UN CORPUS AD HOC: APROVECHAR LOS ARCHIVOS PDF**INTRODUCCIÓN:**

En esta tarea se busca que el alumno utilice software de conversión de archivos PDF a otros formatos más fácilmente manejables por programas de análisis de corpus.

DESARROLLO:

1. Se facilita un texto en formato PDF relevante como fuente de información para la traducción del texto original (Texto 2: "Triturus analyzer").
2. Los alumnos descargan e instalan un programa de conversión de PDF.
3. Siguen los pasos indicados en la Ficha 2 ("Convertir un archivo PDF a uno de texto") y contestan a las preguntas que ésta contiene sobre la utilidad de hacer este tipo de conversión de formato.

Dinámica de trabajo: en parejas

Duración estimada: 2 horas

MATERIALES:**Fichas**

- Ficha 2: "Convertir un archivo pdf a uno de texto"

Textos

- Texto 1 ("Seroteca")
- Texto en formato PDF válido para formar parte de un corpus ad hoc (Texto 2: "Triturus analyzer")

Recursos electrónicos

- URL de un programa de conversión de PDF (gratis o demo) [*TEXTfromPDF*: <http://www.textfrompdf.com/tfpdownload.htm>]

TAREA 5: DE MEMORIA DE TRADUCCIÓN A CORPUS PARALELO

INTRODUCCIÓN:

Esta tarea es compleja en tanto que requiere del alumno que concatene varios ejercicios que van desde descargar textos, pasando por alinearlos de forma semi-automática, para crear una memoria de traducción perfectamente alineada que pueda ser utilizada como un corpus paralelo y de la que se pueda extraer información útil para la traducción que le ocupa.

DESARROLLO:

1. Los alumnos descargan desde una web de tema relevante para la traducción un mismo documento en español y en inglés (Textos 3 y 4).
2. Con el software apropiado, en este caso *WinAlign* de Trados, alinean estos dos textos y los importan desde un programa de traducción asistida, *Translator's Workbench* de Trados, para crear una memoria de traducción. Para ello, los estudiantes siguen las indicaciones de la Ficha 3 ("Alinear textos y convertirlos en una memoria de traducción") y contestan a las preguntas y ejercicios que ésta contiene sobre la utilidad de realizar este proceso.

Dinámica de trabajo: en parejas

Duración estimada: 2 horas

MATERIALES:

Fichas

- Ficha 3: "Alinear textos y convertirlos en una memoria de traducción"

Textos originales y paralelos

- Texto 1 ("Seroteca")
- Mismo texto de tema y/o género relevante en español (Texto 3) y en inglés (Texto 4)

Recursos electrónicos

- URL de sitio web bilingüe relacionado con el tema del texto original [Empresa Grifols: <http://www.grifols.com>]
- Software de alineación de textos *WinAlign* de Trados
- Software de memoria de traducción *Translator's Workbench* de Trados

③ EVALUACIÓN DE LA UNIDAD DIDÁCTICA

1. Valoración realizada por los estudiantes

a) Valoración de su aprendizaje: ficha de autoevaluación

A continuación se recogen en una tabla las puntuaciones concedidas por los estudiantes a sí mismos en relación con lo que han aprendido y una valoración de los resultados.

- Puntuación media de las respuestas proporcionadas por cada alumno. Las puntuaciones muestran un buen nivel de confianza por parte de los estudiantes respecto a lo adquirido durante la unidad con notas que van del 7 al 10 en las dos partes de la ficha.
- Puntuación media de las respuestas en cada pregunta. Si bien se observa una gran homogeneidad en los resultados de todos los contenidos por los que se ha preguntado en la ficha, en la primera parte de la tabla destaca la nota más alta (8,9) concedida a una capacidad novedosa y útil, manejar software y recursos de ubicación y descarga de textos electrónicos desde Internet, y la nota más baja (7,4) concedida al indicador más general (“Soy capaz de traducir textos científico-técnicos del español al inglés”). En la segunda parte de la tabla, entre las notas más altas (9,2), resaltamos la concedida a la contribución del uso de corpus a que los alumnos tengan una mayor seguridad a la hora de dar una solución de traducción, aspecto de gran relevancia al tratarse de una unidad de traducción inversa.

FICHA DE AUTOEVALUACIÓN: RESULTADOS															
UNIDAD DIDÁCTICA "CORPUS Y MEMORIA DE TRADUCCIÓN"															
Estudiantes		A.P.	A.M.	D.L.	E.B.	E.S.	I.M.	J.D.	M.G.	MA.G.	N.M.	S.M.	V.F.	Y.J.	Media de cada pregunta
Soy capaz de...	manejar software de conversión de textos de papel a formato electrónico	10	8	5	10	9	9	7	7	8	6	8	8	9	8
	realizar búsquedas en archivos PDF como si fueran un corpus	8	7	10	10	7	9	9	7	9	9	9	7	8	8,4
	manejar distintos (tipos de) corpus que están disponibles en línea (p.ej. IULA)	8	7	10	10	8	8	9	6	9	9	9	8	8	8,4
	manejar funciones avanzadas de un programa de análisis de corpus (p.ej. extraer colocaciones, agrupaciones, realizar búsquedas por lema, por categoría gramatical, etc.)	10	9	4	8	7	9	9,5	8	9	8	9	10	8	8,3
	manejar software y recursos de ubicación y descarga de textos electrónicos desde Internet (p.ej. Buscopio, Medline, Scirus, WebCopier, Flashgot)	8	8	8	10	9	10	9	7	9	10	8	10	10	8,9
	Manejar software de conversión de archivos .pdf (p.ej. TextfromPdf)	9	7	4	10	9	9	9,5	10	9	10	8,5	8	10	8,7
	manejar software de alineación de textos (p.ej. WinAlign)	7	6		8	6	9	9,5	5	7	8	8	8	10	7,6
	convertir un corpus paralelo (texto original + texto traducido) en una memoria de traducción (p.ej. WinAlign)	7	6		8	6	10	9,5		7	7	8	8	8	7,7
	manejar la función de extraer concordancias en un programa de traducción asistida para utilizar una memoria de traducción a modo de corpus (p.ej. en Translator's Workbench)	7	7	8	8	6	10	9,5		7	5	9	10	8	7,9
	trabajar individualmente en la traducción especializada del español al inglés	9	8	8	9	8	10	9	5	7	7	8,5	9	10	8,3
	trabajar en pareja / en grupo en la traducción especializada del español al inglés	8	8	7	7	8	10	9	7	7	10	8,5	9	10	8,3
	traducir textos científico-técnicos del español al inglés	7	7	6	8	8	8	7	5	7	7	8	9	9	7,4
	buscar, observar y encontrar equivalentes terminológicos en textos científico-técnicos para traducir del español al inglés	7	8	8	7,5	10	10	9,5	8	9	8	8,5	9	10	8,6
	buscar, observar y encontrar las colocaciones de un término en textos científico-técnicos para traducir del español al inglés	8	9	7	8	9	9	9	8	9	8	8,5	9	10	8,6
	buscar, observar y encontrar fraseología en textos científico-técnicos para traducir del español al inglés	7	8	7	8,5	10	9	9	7	9	6	8,5	9	10	8,3
Media de cada alumno		8	7,5	7	8,7	8	9,3	9	7	8,1	7,9	8,5	8,7	9,2	

Te ha ayudado el corpus a ...	traducir mejor	10	7	8	10	10	10	10	7	10	9	9	10	10	9,2
	encontrar los equivalentes de traducción	5	8	7	9	10	10	10	7	9	6	9	10	9	8,4
	estar más seguro de la solución	9	8	10	9	10	10	10	8	10	7	9	10	10	9,2
	comprender mejor el texto original	6	7	5	7	7	10	10	7	9	3	8	9	9	7,5
	comprender la importancia de las colocaciones	8	9	8	10	7	10	10	7	10	9	9	10	9	8,9
	ser más consciente de la naturaleza de los textos utilizados en la fase de documentación (origen, fecha, original o traducción, etc.)	5	7	9	9	7	10	10	7	9	7	9	8	10	8,2
	Media de cada alumno	7,2	7,7	7,8	9	8,5	10	10	7,2	9,5	6,8	8,8	9,5	9,5	

Tabla 25. Resumen de las Fichas de autoevaluación de la UD “Corpus y memoria de traducción”.

Aparte de los aspectos en los que les ha ayudado trabajar con corpus por los que se les pregunta en el cuestionario, algunos alumnos han añadido otros que se pueden leer al completo en el *Volumen 2*. Por ejemplo, D.L., apunta que trabajar con corpus le ha ayudado a “comprender algo mejor la importancia de una acertada elección de software a la hora de afrontar la traducción especializada”; E.S. ha podido ver que en un texto científico (y al contrario de lo que ella pensaba) puede haber más de una traducción correcta, y a E.B le ha ayudado a “determinar el orden de los elementos que conforman un sintagma (especialmente cuando se trata de terminología o fraseología habitual) y a determinar qué tiempos y modos verbales deben usarse en cada caso (*we introduced/was introduced? Has been introduced/Was introduced?/Is introduced?* etc.)”, aspecto éste en el que coincide con la alumna Y.J., que era una nativa inglesa. J.D. apunta que también le ha ayudado a “dudar más de las primeras traducciones intuitivas que uno puede hacer y a insistir en la verificación de las soluciones que se extraen de las búsquedas en corpus, puesto que a veces hay un alto grado de concordancias distintas, y hay que cerciorarse de las diferencias entre una y otra, por ejemplo *hCG is administered* y *hCG is injected*. Ambas aparecían en un alto número de ocasiones, pero no sólo se podía confiar en el corpus, había que comprobar en diccionarios y otros recursos cual era en realidad la diferencia”. S.M. añade que “trabajar con corpus me ha ayudado a mejorar mis conocimientos y mi capacidad para escribir textos en inglés: buscando la solución a un único problema, he acabado leyendo multitud de frases y de posibilidades distintas que, en general, desconocía; además, utilizar corpus me ha servido (y creo que me servirá en el futuro) para traducir de una manera más rápida y sencilla que si empleara casi exclusivamente diccionarios bilingües especializados (que es lo que solía hacer antes de cursar esta asignatura)”. Por último, M.A.G. apunta que utilizar corpus le ha ayudado en varios aspectos, como “ahorrar infinidad de tiempo por el hecho de poder consultar multitud de archivos simultáneamente, sentirme mucho más segura a la hora de abordar una traducción inversa, y analizar los textos paralelos desde distintos puntos de vista, sacándoles el máximo partido (concordancias, colocaciones, fraseología, etc.)”.

b) Valoración de la unidad didáctica: ficha de valoración

Los comentarios relevantes realizados por los estudiantes en relación con la unidad se ofrecen en el *Volumen 2* de esta tesis. Otros comentarios se han obviado ya que estaban orientados a la asignatura completa donde esta unidad didáctica se puso en práctica y no sólo a la unidad que nos ocupa y de la que se pretendía obtener una valoración.

Los comentarios de los estudiantes en la Ficha de valoración permiten constatar la buena recepción que ha tenido la presente unidad didáctica.

(i) Nueva forma de documentarse

Todos los alumnos sin excepción consideran que utilizando corpus han aprendido no sólo una forma de documentarse nueva, sino también más eficaz y rápida que contrasta con sus antiguas prácticas de sólo consultar diccionarios o navegar por páginas web de forma no sistemática.

(ii) Mejoras a la hora de traducir textos del español al inglés

La práctica totalidad de los alumnos afirman haber mejorado la calidad de sus traducciones al inglés o, al menos, haber aumentado su confianza y su capacidad de justificar sus soluciones. Aquellos que no consideran que han mejorado basan su respuesta en que sus resultados en una traducción inversa siempre serán peores que en una directa, argumento con el que nosotros sólo estamos parcialmente de acuerdo.

(iii) Destacado positivamente por los alumnos

Destaca sobremanera la utilidad que han apreciado en las búsquedas por lema o por categoría gramatical en corpus, así como las búsquedas en una memoria de traducción utilizado a modo de corpus paralelo.

(iv) Sugerencias de mejora y comentarios

Esta unidad destaca especialmente por lo positivo de los resultados, la satisfacción de los estudiantes y la adecuación de objetivos y materiales. Los alumnos alaban el sentido práctico y útil de la unidad y sólo apuntan que les hubiera gustado acceder a estos contenidos en cursos anteriores y dedicarle más tiempo.

2. Valoración realizada por el docente

a) Valoración del aprendizaje de los estudiantes

En esta unidad didáctica se ha optado por no incluir una tarea final que incluya una evaluación integrada de las competencias adquiridas a lo largo de la unidad. En vez de esta modalidad, la evaluación se ha realizado a través de la observación de la marcha de las tareas y la corrección de las Worksheets 1b y 2 y las Fichas 1, 2 y 3. Los resultados han sido excelentes.

b) Valoración de la unidad didáctica

La unidad didáctica que se muestra es muy cercana a su versión inicial, salvo algunas correcciones de detalle. Ésta funcionó perfectamente tanto en su forma completa como cuando se utilizaron sólo partes de la misma, y permitió que los estudiantes que la trabajaron crearan corpus y supieran buscar dentro de los mismos y que también establecieran la relación entre el uso de corpus y las memorias de traducción.

④ EVOLUCIÓN DE LA UNIDAD DIDÁCTICA

Modificaciones efectuadas

La idea de diseñar una unidad didáctica que una el uso de corpus con las memorias de traducción nace de la impresión que se obtiene del pequeño estudio sobre el uso de corpus en agencias de traducción (cf. supra. 2.4.), donde se observa que éstas, a pesar de no utilizar expresamente corpus electrónicos, sí guardan los textos relacionados con los encargos de traducción (originales y su traducción, paralelos, de apoyo) y los aprovechan con memorias de traducción, sobre todo la del paquete Trados.

Apuntamos aquí algunas recomendaciones en relación a Fichas o Tareas:

- Tarea 1 (“Del papel al formato electrónico”) y Tarea 4 (“Crear un corpus ad hoc: Aprovechar los archivos PDF”): Dependiendo del nivel de los alumnos en cuanto al manejo del ordenador, se podría prescindir de tareas como la de escanear un texto en papel o la de convertir un archivo PDF a texto plano, por ejemplo.
- Tarea extra: Recomendaríamos la inclusión de una tarea más para aquellos que no tengan demasiada práctica en el uso de corpus. Esta tarea consistiría en utilizar a modo de corpus un solo documento en formato PDF o varios documentos PDF guardados en un mismo directorio para obtener información a través de la extracción de “concordancias”. Se pediría encontrar determinada información haciendo búsquedas simples o complejas de un término o expresión. El estudiante obtendría varios segmentos en los que aparecería dicho término o expresión en contexto y podría navegar a la vez por el o los archivos PDF para observar un contexto más amplio. Un ejercicio podría ser buscar toda la información posible sobre un término (p.ej. “plate”) y sus derivados, o cómo funciona un verbo (p.ej. “stimulate”) en todos los archivos PDF que describen los productos relacionados con el analizador *Triturus* y que cuelgan de la URL

<http://www.triturus.com/products/productlist.html>, una vez descargados a una unidad local. Esta mini-tarea o ejercicio vendría después de la Tarea 1 y cubriría el contenido III. 2.1.3. [Función *Búsqueda avanzada* (de *Adobe Reader*)].



UNIDAD DIDÁCTICA

Sólo utilizar corpus; utilizar corpus solo



En esta unidad didáctica se pone todo el énfasis en la explotación de recursos en línea no lexicográficos, como por ejemplo libros especializados que están disponibles en Internet, o textos que se pueden descargar para construir un corpus ad hoc. Se busca mejorar la calidad de las búsquedas en un corpus y de los textos consultados aplicando tanto habilidades ya adquiridas en unidades anteriores como otras nuevas. En este caso se trata de un proyecto de traducción extenso e individual en el que se busca un alto grado de autonomía.



Grupos de prueba: La unidad original "Dermovet, 1ª puesta en práctica" se probó en un grupo de 26 alumnos de 4º curso de licenciatura en Traducción e Interpretación de la UAB, en una asignatura de traducción especializada inversa (español A - inglés B), durante el año académico 2002-03. Cronológicamente, ésta es la primera unidad didáctica que se probó como tal, es decir, ya no como ejercicios sueltos sino como tareas consecutivas y conducentes a una tarea final. La evaluación pertenece a la versión primitiva, a partir de la cual se introdujeron todas las modificaciones, pero las tareas y fichas que se presentan aquí corresponden a las utilizadas en la última versión de la unidad didáctica, una vez revisada.

Los materiales de esta unidad se encuentran en el *Volumen 2* de esta tesis (págs. 689-728).



❶ PRESENTACIÓN DE LA UNIDAD DIDÁCTICA

“Sólo utilizar corpus; utilizar corpus solo”

NIVEL DE USO DE CORPUS

- Consolidación

MATERIA

- Traducción especializada médica inversa (español A – inglés B)

OBJETIVOS

Objetivos relacionados con el uso de corpus para traducir

- Crear un corpus comparable (monolingüe)
- Utilizar software y recursos de ubicación y descarga de textos electrónicos desde Internet
- Utilizar funciones avanzadas de un programa de análisis de corpus
- Integrar conocimientos para resolver problemas de traducción utilizando corpus a nivel avanzado

Objetivos relacionados con la materia

- Utilizar recursos de documentación de todo tipo para resolver problemas de traducción especializada (buscadores, bibliotecas digitales, corpus, etc.)
- Crear recursos de documentación para resolver problemas de traducción (corpus, glosarios, etc.)
- Traducir textos científico-técnicos

Objetivos relacionados con competencias genéricas

- Gestionar la información (buscar y analizar información proveniente de fuentes diversas)

CONTENIDOS

- Buscadores (*Google, Amazon*) (II. 1.3.1.)
- Buscadores de textos de especialidad (*Medline, Scirus, FreeBooks4Doctors*) (II. 1.3.3.)
- Ayuda en Internet (*Intute: Virtual Training Suite*) (II. 1.4.1.)
- Software de descarga de textos (*WebCopier; Free Download Manager; Flashgot*) (III. 1.3.2.)
- Software gratuito de análisis de corpus (*AntConc*) (III. 1.10.1.)
- Función *Concordance* (de *AntConc*) (III. 2.1.2.)
- Función *Kwic sort* (de *AntConc*) (III. 2.3.2.)
- Función *Collocates* (de *AntConc*) (III. 2.4.1.)
- Listas de palabras por orden de frecuencia (de *AntConc*) (III. 2.7.2.)
- Extracción de definiciones: hiperónimos, etc. (IV. 2.1.1.)
- Extracción de tipos: hipónimos, etc. (IV. 2.1.2.)
- Extracción de vocabulario general y términos de especialidad (IV. 2.2.1.)
- Extracción de colocaciones (IV. 2.2.2.)
- Extracción de enunciados ilocutivos: ordenar (dar instrucciones, indicaciones...) (IV. 2.2.4.)



ESTRUCTURA

- Tarea 1: "Navegando dentro del texto original, un sitio web"
- Tarea 2: "Navegando dentro de textos paralelos"
- Tarea 3: "Cómo mejorar las búsquedas en Internet y encontrar mejores textos"
- Tarea 4: "Traducir mucho buscando poco"

EVALUACIÓN

Indicadores

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación,...

- Indicador 1: selecciona un corpus adecuado
- Indicador 2: busca una palabra/cadena de palabras adecuada
- Indicador 3: limita la búsqueda
- Indicador 4: ordena los resultados
- Indicador 5: usa las funciones del programa de análisis de corpus
- Indicador 6: emplea un tiempo razonable en resolver un problema

Fichas de evaluación

- Ficha de recogida de datos sobre el uso de corpus al traducir
- Ficha de autoevaluación
- Ficha de valoración de la unidad

MATERIALES

Fichas

- Worksheet 1: "Searching inside the source text"
- Worksheet 2: "Searching inside books, online and offline"
- Worksheet 3: "Help with searching for quality texts on the Internet"
- Worksheet 4: "Techniques for searching in a comparable target-language corpus"

Textos

- Texto 1: "Dermovet"

Textos de apoyo

- 1: Material de apoyo a los programas *TagEditor* y *Translator's Workbench* de Trados

Recursos electrónicos

- URL de un portal de venta de libros en Internet que permite acceder a parte de los contenidos de los mismos [Amazon: <http://www.amazon.co.uk>]
- URL de un buscador de libros sobre medicina [Freebooks4doctors: <http://freebooks4doctors.com>]
- URL de un manual de medicina [Merck Manual: <http://www.merck.com/mmpe/sec10.html>]
- URL de tutorial en línea para mejorar la búsqueda de documentos en Internet [RDN Virtual Training Suite: <http://www.vts.rdn.ac.uk/>]
- URLs de buscadores de TPs de tema científico [PubMed: <http://www4.ncbi.nlm.nih.gov/entrez/query.fcgi> ; Scirus: <http://www.scirus.com>]



- Programa informático de descarga de webs y navegación *offline* (*WebCopier*)
- Programa informático de análisis de corpus (*AntConc*)
- Programa informático de descarga de webs (*Flashgot; Free Download Manager*)
- Programas informáticos de traducción asistida (*TagEditor* y *Translator's Workbench* de Trados)

DURACIÓN ESTIMADA

15 horas

DESARROLLO DE LA UNIDAD DIDÁCTICA

TAREA 1: NAVEGANDO DENTRO DEL TEXTO ORIGINAL, UN SITIO WEB

INTRODUCCIÓN:

Esta tarea sirve para introducir al estudiante en un proyecto extenso que realizará individualmente y con un mayor grado de autonomía respecto a otras unidades. El texto original es un sitio web completo y, debido a su naturaleza híbrida (textos con distintos tonos, funciones, grados de especialización), para su traducción se utilizarán varios corpus, empezando por considerar como corpus el propio TO.

DESARROLLO:

1. Se presenta el encargo de traducción que cada alumno deberá realizar. Se trata de la traducción al inglés de todo un sitio web personal dedicado a la dermatología veterinaria (Texto 1: "Dermovet"). Los lectores potenciales de la traducción son, especialmente, hablantes de inglés que tengan perros o gatos y puedan necesitar los servicios de un veterinario especializado en dermatología.
2. Los alumnos navegan rápidamente (10 min.) por el sitio web para observar el tono, registro, tema y distribución del contenido en las diferentes páginas web.
3. Con el programa *WebCopier* descargan todo el sitio web, es decir, el TO.
4. Con el programa de análisis de corpus *AntConc* extraen una lista de palabras del TO por orden de frecuencia para observar qué palabras se repiten o son clave en todo el sitio web y pueden servir para después buscar textos paralelos en inglés.
5. Rellenan la Worksheet 1 ("Searching inside the source text"), para lo cual, a partir de la lista de palabras, extraen concordancias de aquéllas que les parezcan candidatas a términos o interesantes como colocaciones. Para observar patrones que se repitan pueden reordenar el contexto de las concordancias, a partir de las cuales también pueden observar contextos más amplios.

Dinámica de grupos: individual

Duración estimada: 2 horas

MATERIALES:

Fichas

- Worksheet 1: "Searching inside the source text"

Textos

- Texto 1: "Dermovet"

Recursos electrónicos

- Programa informático de descarga de webs y navegación *offline* (*WebCopier*)
- Programa informático de análisis de corpus (*AntConc*)

TAREA 2: NAVEGANDO DENTRO DE TEXTOS PARALELOS

INTRODUCCIÓN:

En esta tarea se busca ampliar la batería de recursos en línea útiles para muchas situaciones de traducción en las que los estudiantes pueden necesitar documentarse. Se trata de que accedan a libros especializados que se ofrecen en Internet de forma gratuita y, por un lado, asimilen cómo pueden extraer concordancias de los mismos y, por otro lado, practiquen cómo pueden descargarse su contenido para que éste forme parte de un corpus ad hoc.

DESARROLLO:

1. Los alumnos hacen la Worksheet 2 ("Searching inside books, online and offline"), con la que aprenden, en primer lugar, a utilizar la herramienta de extracción de concordancias de libros de *Amazon* con el fin de obtener información tanto conceptual como lingüística útil para la traducción del sitio web de Dermovet. En segundo lugar, practican la descarga de contenido web, en este caso, de libros enteros o secciones de libros que ofrece el sitio *FreeBooks4Doctors* con el fin de empezar a recopilar textos para construir un corpus ad hoc.

Dinámica de grupos: individual

Duración estimada: 2 horas

MATERIALES:

Fichas

- Worksheet 2: "Searching inside books, online and offline"

Recursos electrónicos

- URL de un portal de venta de libros en Internet que permite acceder a parte de los contenidos de los mismos [*Amazon*: <http://www.amazon.co.uk>]
- URL de un buscador de libros sobre medicina [*Freebooks4doctors*: <http://freebooks4doctors.com>]
- URL de un manual de medicina [*Merck Manual*: <http://www.merck.com/mmpe/sec10.html>]
- Programa informático de descarga de webs (*Flashgot*; *Free Download Manager*)

TAREA 3: CÓMO MEJORAR LAS BÚSQUEDAS EN INTERNET Y ENCONTRAR MEJORES TEXTOS

INTRODUCCIÓN:

Aparte de los textos publicados, sean éstos libros o artículos de revistas especializadas, es probable que los estudiantes recurran a buscadores generales en Internet para consultar textos y, como es el caso, para construir su propio corpus especializado. Esta tarea se introduce con el fin de mejorar en los estudiantes la calidad de las búsquedas y, por ende, los textos que consultan y recogen de la Web.

DESARROLLO:

1. Los alumnos rellenan la Worksheet 3 ("Help with searching for good quality texts on the Internet") a la vez que siguen un tutorial en línea que muestra cómo hacer búsquedas en Internet más ajustadas a la información que se necesita.

Dinámica de grupos: individual

Duración estimada: 1 hora

MATERIALES:

Fichas

- Worksheet 3: "Help with searching for quality texts on the Internet"

Recursos electrónicos

- URL de tutorial en línea para mejorar la búsqueda de documentos en Internet [*RDN Virtual Training Suite*: <http://www.vts.rdn.ac.uk/>]

TAREA 4: TRADUCIR MUCHO BUSCANDO POCO

INTRODUCCIÓN:

En esta tarea el estudiante sigue recogiendo textos para su corpus ad hoc que, debido a la naturaleza híbrida del texto original, será variado en cuanto a la función y registro de los textos que contenga. Esta tarea también busca que el estudiante, cuando se enfrente a un problema de traducción, se esfuerce en pensar búsquedas de calidad y económicas para realizar en un corpus y así no multiplique sus búsquedas innecesariamente. Para esto, deberá pensar bien la información que quiere obtener del corpus y, normalmente, empezar sus búsquedas por lo que ya sabe para llegar a descubrir lo que no sabe, es decir, introducir como búsqueda aquellos elementos de cuya traducción ya está seguro para obtener ayuda en la traducción de los que no lo está.

DESARROLLO:

1. Los alumnos rellenan la Worksheet 4 ("Techniques for searching in a comparable target-language corpus"), que contiene preguntas que les llevan a reflexionar sobre sus conocimientos de inglés en un ámbito especializado y las formas como pueden buscar soluciones en un corpus.
2. Para poder hacer el primer ejercicio de la Worksheet 4 cada alumno tiene que recoger un texto donde aparezcan instrucciones para pacientes y, si es posible, pacientes que sean animales. Se juntan los textos de todos para obtener un pequeño corpus ad hoc de instrucciones a pacientes.
3. Para poder hacer el segundo ejercicio cada alumno recoge textos paralelos, entendidos en un sentido amplio, de fuentes ya conocidas para ellos como *Medline*, *Scirus*, etc., que añadirán a los ya recogidos en la Tarea 2. Como palabras clave en inglés para buscar nuevos textos se pueden utilizar aquéllas obtenidas también en la Tarea 2.
4. Traducen todo el sitio web Dermovet individualmente con los programas *TagEditor* y *Translator's Workbench* de Trados, de los que se les facilita un dossier explicativo. *TagEditor* es un programa que les permite traducir páginas web sin tener que ocuparse de las etiquetas HTML; *Translator's Workbench* es una memoria de traducción que les permite reaprovechar la traducción de segmentos que se repitan en el texto original, como así ocurre en este encargo con todo el contenido que aparece en los marcos izquierdo y derecho de las páginas web de Dermovet.
5. Utilizan el corpus ad hoc que han creado como principal recurso de documentación.

Dinámica de grupos: individual

Duración estimada: 10 horas

MATERIALES:**Fichas**

- Worksheet 4: "Techniques for searching in a comparable target-language corpus"

Texto original

- Texto 1: "Dermovet"

Textos de apoyo

- Material de apoyo a los programas *TagEditor* y *Translator's Workbench* de Trados

Recursos electrónicos

- URLs de buscadores de TPs de tema científico [PubMed: <http://www4.ncbi.nlm.nih.gov/entrez/query.fcgi> ; Scirus: <http://www.scirus.com>]
- Programa informático *AntConc*
- Programas informáticos de traducción asistida (*TagEditor* y *Translator's Workbench* de Trados)

❸ EVALUACIÓN DE LA UNIDAD DIDÁCTICA

1. Valoración realizada por los estudiantes

a) Valoración de su aprendizaje: ficha de autoevaluación

A continuación se recogen en una tabla las puntuaciones concedidas por los estudiantes a sí mismos en relación con las competencias adquiridas. Es necesario puntualizar que la evaluación que sigue es el resultado de la puesta en práctica, en el curso 2002-03, de una unidad didáctica de contenido similar al que se ha presentado en el apartado anterior, pero con una organización y distribución muy diferentes. De hecho, las puntuaciones, comentarios y sugerencias recogidos tras esa primera prueba, dieron lugar a modificaciones en la unidad, que quedó como finalmente se ha expuesto.

De la unidad didáctica primitiva “Dermovet, 1ª puesta en práctica”, los estudiantes valoraron de forma especialmente positiva:

- La capacidad que habían adquirido para utilizar un corpus monolingüe en lengua meta como ayuda a la hora de traducir un texto, que finalmente es el objetivo principal que se quería lograr.
- La seguridad que les aportó la utilización de este recurso y metodología a la hora de dar soluciones, como es usual en el trabajo con corpus en traducción inversa.

La menor puntuación, aún sin ser negativa, fue para:

- La identificación de qué tipo de documentos ayudan a traducir qué parte del TO.

De entre los alumnos, C.V. destacó por su confianza en cuanto a lo aprendido, lo cual se reflejaría después en unos resultados de traducción excelentes, y, en contraposición, M.B destacó por mostrar un grado de satisfacción mucho más bajo.

FICHA DE AUTOEVALUACIÓN: RESULTADOS
UNIDAD DIDÁCTICA "DERMOVET, 1ª PUESTA EN PRÁCTICA"

Estudiantes		A. T.	C. V.	C. N.	D. P.	D. V.	D. M.	E. R.	E. L.	I. G.	L. P.	L. S.	L. R.	M. Fo.	M. B.	M. F.	M. R.	M. M.	M. Mu.	M. Mz.	M. Mt.	N. R.	P. M.	P. R.	S. N.	S. B.	V. V.	Media de cada pregunta
Soy capaz de...	analizar un texto original y ver los problemas que puede plantear su traducción	8	8	7	10	8	9	10	7	9	7	8	5	8	6	7	7	7	7	6	9	8	8	8	8	8	8	7,7
	analizar las necesidades de documentación específicas según las características del TO y el encargo de traducción	7	8,5	8	8	7	8	10	7	9	6	9	7	7	6	9	6,5	7,5	7	7	9	9	8	8	9	7	8	7,8
	encontrar recursos y documentos en formato electrónico para ayudarme a traducir una página web de tema científico	9	8,5	8	7	7	8	8	6	8	7	8	9	7	7	7	8	7	7	7	7	7	7	8	9	7	8	7,6
	evaluar los recursos y documentos hallados y seleccionados para su inclusión en un corpus	8	8	9	8	7	8	9	7	8	7	9	6	7	6	8	6,5	6,5	8	7	7	7	9	9	7	6	9	7,6
	compilar un corpus monolingüe en lengua meta	8	8	8	8	8	9	8	7	8	6	8	7	7	6,5	8	7	7	8	7	8	8	8	9	8	7	9	7,7
	utilizar un corpus monolingüe en lengua meta como ayuda para traducir un texto	9	8,5	8	8	9	7	9	8	8	8	7	8	8	7	7	7	8,5	8	6,5	8	7	7	9	8	8	9	7,9

	identificar qué tipo de documentos ayudan a traducir qué parte del TO	8	7,5	7	8	7	8	8	6	8	6	8	7	6	6	6	7	8	5	6	7	8	7	9	8	7	9	7,2
	encontrar equivalentes de traducción utilizando el corpus que he creado sirviéndome de diversas estrategias de búsqueda	8	8,5	8	8	8	7	7	8	8	8	8	8	8	7	8	8	7	6	7	8	7	9	8	8	7	8	7,7
	Media de cada alumno	8,1	8,9	7,9	8,1	7,6	8	8,6	7	8,3	6,9	8,1	7,1	7,3	6,4	7,5	7,1	7,3	7	6,7	7,9	7,6	7,9	8,5	8,1	7,1	8,5	

Estudiantes		A. T.	C. V.	C. N.	D. P.	D. V.	D. M.	E. R.	E. L.	I. G.	L. P.	L. S.	L. R.	M. Fo.	M. B.	M. F.	M. R.	M. M.	M. Mu.	M. Mz.	M. Mt.	N. R.	P. M.	P. R.	S. N.	S. B.	V. V.	Media de cada pregunta
Te ha ayudado el corpus a ...	traducir mejor	7	9,5	6	8	8	8	8	6	6	6	4	8	8	8	8	7	7,5	8	8	8	8	9	8	8	8	7	7,5
	encontrar los equivalentes de traducción	8	9	8	8	8	7	7	8	8	9	9	8	8	8	9	7	8,5	8	8	10	7	9	8	9	9	8	8,2
	estar más seguro de la solución	8	9,5	9	9	9	9	8	8	9	8	6	9	8	7	9	7	9	9	7	9	9	8	8	6	9	8	8,3
	comprender la importancia de las colocaciones (palabras que suelen acompañar a otras)	8	9,5	7	9	9	9	9	6	8	7	7	9	7	8	8	8	9	7	7	8	9	8	9	8	9	8	8,1
	ser más consciente de la naturaleza de los textos utilizados en la fase de documentación (origen, fecha, original o traducción, etc.)	10	8	7	7	6	7	8	7	7	6	8	9	6	7	8	9	8,5	7	7	6	6	7	8	8	7	8	7,4
	Media de cada alumno	8,2	9,1	7,4	8,2	8	8	8	7	7,6	7,2	6,8	8,6	7,4	7,6	8,4	7,6	8,5	7,8	7,4	8,2	7,8	8,2	8,2	7,8	8,4	7,8	

Tabla 26. Resultados de la Ficha de autoevaluación de los alumnos de la UD “Dermovet, 1ª puesta en práctica”.

Otros aspectos que los estudiantes quisieron destacar del uso de corpus fueron:

- La rapidez que aportó a su proceso de traducción desafiando incluso la idea previa de que construir un corpus les haría perder tiempo.
- La mejora en la sintaxis y estilo científico en inglés, ahora más genuinos al poder confirmar sus intuiciones con los textos del corpus, del mismo tipo o tema que su texto original.
- La seguridad, de nuevo, que les proporcionó utilizar corpus al poder, entre otras cosas, ver la frecuencia de un término, confirmar su ortografía cuando existen variantes, confirmar su uso cuando existen sinónimos, etc.

b) Valoración de la unidad didáctica: ficha de valoración

La ficha que sigue fue la que se pasó a los estudiantes que realizaron la unidad didáctica en el curso 2002-03, es decir, con su diseño original.

FICHA DE VALORACIÓN: UNIDAD DIDÁCTICA "DERMOVET, 1ª PUESTA EN PRÁCTICA"
<p>Comentarios sobre la unidad</p> <ul style="list-style-type: none"> • ¿Crees que utilizando corpus has aprendido también una nueva forma de documentarte para resolver problemas de traducción en un texto? Por favor, explicita al máximo. • ¿Hay algo de esta unidad didáctica que te haya gustado especialmente entre los siguientes temas? Por favor, explicita al máximo. • ¿Hay algo que no te haya gustado? Por favor, explicita al máximo. • ¿Cómo lo cambiarías? • Otros comentarios

Documento 13. Ficha de valoración de la UD "Dermovet, 1ª puesta en práctica".

Los temas más destacados surgidos en las respuestas de los alumnos fueron los siguientes:

(i) Nueva forma de documentarse

Se podría decir que el reconocimiento por parte de los estudiantes de la conveniencia de utilizar corpus para traducir es, en este caso, doble, en el sentido de que muchos admitieron haber usado anteriormente buscadores como *Google* y textos paralelos para documentarse, pero no de la forma optimizada que permite la metodología y herramientas de corpus. Por otra parte, el descubrimiento de esta posibilidad de optimización del uso de textos paralelos les derrumbó algunos prejuicios como que los diccionarios o glosarios especializados solucionaban todos los problemas, o que construir un corpus ad hoc no merecía la pena por la inversión de tiempo que requiere. En este sentido, muchos pudieron observar que la inversión inicial de tiempo en crear un corpus de textos de calidad tenía su recompensa más adelante cuando podían obtener del mismo no sólo términos especializados, sino también definiciones, colocaciones, e incluso la solución a otros problemas de traducción que no se estaban buscando.

(ii) Aspectos destacados positivamente por los alumnos

Aprender a utilizar el programa de edición de páginas web *TagEditor* y el programa de análisis de corpus *WordSmith Tools* fue, con diferencia, lo más destacado por los alumnos, especialmente porque el encargo de traducir una página web permitió que se diera un aprendizaje contextualizado de estas herramientas. Junto a esto, construir un corpus de textos paralelos y descubrir cuánto provecho se podía sacar de los mismos, dieron a la unidad didáctica un carácter práctico que los alumnos también reconocieron como positivo.

(iii) Sugerencias de mejora y comentarios

En este caso también está claro cuál fue el aspecto más destacado por los estudiantes en un sentido negativo, la falta de tiempo, especialmente para realizar la traducción, aunque sí que se dedicó mucho tiempo a su preparación con abundantes fichas de ejercicios. Las sugerencias fueron en la dirección de unificar y simplificar alguna ficha, dejando muy claro su objetivo e intentando minimizar los problemas técnicos y de disponibilidad.

Por último, a pesar de sus puntualizaciones, los estudiantes fueron capaces de reconocer la progresión en el trabajo y en las fichas y que todo ello había de conducirles hacia la asimilación de una metodología y una mejor resolución de la traducción final.

2. Valoración realizada por el docente

a) Valoración del aprendizaje de los estudiantes

La Ficha de recogida de datos sobre el uso de corpus, rellena por los estudiantes a la vez que realizaban la traducción de una parte del sitio web de Dermovet en el curso 2002-03, es la que ofreció el input para rellenar la Ficha de cruce de datos. En el *Volumen 2* se pueden encontrar tanto la Ficha de recogida de datos como todos los resultados de la Ficha de cruce de datos de todos los estudiantes que siguieron la unidad en el curso 2002-03. Lo que se ofrece a continuación es un resumen de todos los resultados de las Fichas de cruce de datos y un breve análisis.

- Puntuaciones medias de los alumnos. Se observa una gran heterogeneidad entre las notas que denota una asimilación muy desigual de los contenidos, si bien es cierto que la mayoría de las puntuaciones son buenas.
- Puntuaciones medias obtenidas por el conjunto de alumnos en cada indicador. Estos resultados permiten identificar los indicadores de competencias no alcanzadas plenamente que son, por orden de importancia:
 - Limitar la búsqueda de una forma conveniente (hacer búsquedas truncadas, con palabra de contexto), habilidad que casi no se ha desarrollado.
 - Ordenar los resultados de unas concordancias de una forma conveniente.

Por otro lado, si se consulta el *Volumen 2* y la leyenda de la Ficha de cruce de datos, también se puede observar que entre los problemas por los que se preguntó a los alumnos, que eran de diverso tipo (terminológico, pragmático, de coherencia, etc.), el problema cuya resolución ofreció mayor dificultad fue sin duda el relacionado con la coherencia y cohesión (Problema 4), hecho que nos ha parecido digno de mención aquí.

FICHA DE CRUCE DE DATOS: RESUMEN DE RESULTADOS																		
UNIDAD DIDÁCTICA: "DERMOVET, 1ª PUESTA EN PRÁCTICA"																		
Estudiantes	C.V. + D.M.	C.N. + M.M.	D.P.	D.V.	E.R. + N.R.	E.L.	I.G. + A.T.	L.P.	L.S.	L.R.	M.Fo.	M.R. + P.M.	M.Mas.	M.B. + M.M.	P.R. + V.V.	S.N.	Media de cada indicador (tras aplicar el %)	Media de cada indicador (sobre 10 puntos)
Indicador 1: selecciona un corpus adecuado a sus necesidades (20%)	0,6	1,6	1,8	1,8	1,8	1,8	1,6	1,8	1,8	1,8	1,8	2	2	2	2	1,4	1,7/2	8,5
Indicador 2: busca una palabra/cadena de palabras adecuada a sus necesidades (20%)	0,6	1,6	1,8	1,6	1,6	1,6	1,2	1,8	1,8	1,6	1,8	2	1,6	1,4	2	1,8	1,6/2	8
Indicador 3: limita la búsqueda de una forma conveniente (20%)	0	0	0	0	1,6		0		0	1	0,4	2	1	0	1,2	1	0,6/2	3
Indicador 4: ordena los resultados de una forma conveniente (20%)	2	0	1	2	1,2	2	0	2	0	0	1,4	2	2	1	2	0	1,2/2	6
Indicador 5: usa las funciones del programa de una forma conveniente (10%)	0,8	0,7	1	1	1	1	0	1	0,6	0	0,9	1	1	1	1	0	0,8/1	8
Indicador 6: emplea un tiempo razonable en resolver un problema (10%)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1/1	10
Nota de cada alumno	5	4,9	6,6	7,4	8,2	9,2	3,8	9,5	5,2	5,4	7,3	10	8,6	6,4	9,2	5,2		

Tabla 27. Resumen de resultados de la Ficha de cruce de datos de la UD "Dermovet, 1ª puesta en práctica"; resultados del curso 2002-03.

b) Valoración de la unidad didáctica

Las puntuaciones otorgadas por los estudiantes en la Ficha de autoevaluación, sus comentarios de la Ficha de valoración de la unidad didáctica y los resultados observados en las Fichas de cruce de datos obligan a considerar la conveniencia de la unidad como se probó en un primer momento. Si bien es cierto que la mayoría de los objetivos se cubren y alcanzan con la unidad probada, también lo es que ésta necesita algunas modificaciones profundas en cuanto a la estructuración de ejercicios y la distribución del tiempo.

4 EVOLUCIÓN DE LA UNIDAD DIDÁCTICA

Modificaciones efectuadas

Una vez puesta en práctica y vistos los resultados, positivos en general, pero que dejan entrever fallos de organización, se procedió a introducir cambios importantes en la unidad didáctica para llegar a la que se ha mostrado en los apartados ❶ y ❷.

Así, en cuanto a los siguientes aspectos detectados o apuntados como mejorables, se tomaron las siguientes medidas:

- Cuestiones de disponibilidad de recursos y software.

La nueva unidad didáctica está basada casi completamente en recursos en línea y software gratuito (el programa *AntConc* para las Tareas 1 y 4, *WebCopier* en la Tarea 1, *Amazon*, *FreeBooks4Doctors* y un gestor de descarga *-Flashgot-* en la Tarea 2, un tutorial en línea en la Tarea 3, y *Medline* y *Scirus* en la Tarea 4, en la que también se utiliza el *TagEditor*, de Trados, que es el único que ya no se puede descargar desde la Red.

- Práctica en la reorganización del contexto de unas concordancias.

Vista la necesidad de algunos alumnos de recordar la función *Resort/Kwic Sort*, se introduce su uso en la Worksheet 1 de la Tarea 1.

- Práctica en la limitación de las búsquedas.

Vistos los resultados extraídos de las Fichas de cruce de datos, en la Worksheet 4 se pone énfasis sobre diversas formas de realizar búsquedas, a través de palabras clave, de palabras con asterisco, de palabras de contexto, para acabar precisamente con un resumen de estrategias de búsqueda en un corpus monolingüe.

- Tiempo.

Para reorganizar el tiempo dedicado a las distintas tareas de la unidad didáctica, se reducen las fichas quedando en una por tarea, y se aumenta considerablemente el tiempo destinado a la traducción del texto original. Éste, dado el interés por la localización de páginas web por parte de los alumnos, ya no es sólo una parte de una página, sino todo el sitio web.

CONCLUSIONES

El objetivo general de esta tesis ha sido el de integrar el uso de corpus electrónicos en la enseñanza de la traducción. Tomando como punto de referencia los objetivos específicos planteados al inicio del trabajo, se pueden obtener los resultados que siguen.

1. ELABORACIÓN DEL MARCO CONCEPTUAL DE LA LINGÜÍSTICA DE CORPUS Y SUS CAMPOS DE APLICACIÓN

1. Establecimiento del marco teórico de la lingüística de corpus

En este trabajo hemos establecido el marco teórico de la lingüística de corpus (cfr. supra *Capítulo 1*). Hemos expuesto las fases del trabajo con corpus y revisado los campos donde se ha empleado la metodología de corpus hasta la fecha, poniendo especial énfasis en su aplicación en la traducción y campos afines (cfr. supra *Capítulo 3*). Como metodología de estudio de la lengua que es, la metodología de corpus electrónicos, en líneas generales, ha venido a aportar al estudio de la lengua y la traducción unas bases empíricas sólidas fundamentadas en la autenticidad, la cantidad y la calidad de los textos que conforman dichos corpus, además de un software que permite el análisis de datos de una forma repetible y más objetiva, aunque siempre interpretable.

2. Elaboración del estado de la cuestión de la existencia y uso de corpus electrónicos en España

Hemos dado cuenta de los corpus principales que existen en España o que incluyen la lengua española (cfr. supra *Capítulo 2: 2.1. y 2.2.*) y de los trabajos de investigación elaborados en nuestro país utilizando esta metodología en el campo de la traducción (cfr. supra *Capítulo 2: 2.3.*).

Además, con el fin de aportar datos que dieran soporte a la selección de objetivos y materiales para la elaboración de las unidades didácticas, hemos presentado los resultados de un sondeo sobre el uso de corpus electrónicos en 16 agencias de traducción españolas para comprobar si estos recursos se utilizan en el mundo profesional y de qué manera (cfr. supra *Capítulo 2: 2.4.*).

3. Descripción de los tipos de corpus electrónicos más útiles para la formación de traductores

El campo de aplicación que más nos interesa, el uso de corpus en la formación de traductores es, hasta la fecha, escaso y anecdótico. Aquí hemos dado cuenta de los tipos de corpus más útiles para la didáctica de la traducción (cfr. supra *Capítulo 4: 4.2.*) y de las pruebas que se han realizado hasta el momento para comprobar la utilidad de su uso (cfr. supra *Capítulo 4: 4.3.*). También hemos recopilado y aportado ejemplos propios de las estrategias más útiles de búsqueda de equivalentes de traducción en un corpus especializado monolingüe en lengua meta (cfr. supra *Capítulo 4: 4.2.5.*).

2. ELABORACIÓN DEL MARCO CONCEPTUAL DE LA DIDÁCTICA DE LA TRADUCCIÓN APROPIADO PARA ENSEÑAR A UTILIZAR CORPUS PARA TRADUCIR

1. Ubicación de la competencia traductora y su adquisición en la enseñanza de la traducción

Hemos ubicado la competencia traductora y su adquisición en la enseñanza de la traducción y dado cuenta de los estudios principales realizados sobre ambas. Nos hemos centrado en la investigación del grupo PACTE que, (1) para la competencia traductora, contempla la existencia de cinco subcompetencias, entre las cuales hemos destacado la subcompetencia instrumental por la relación que guarda con nuestro trabajo; y (2) para la adquisición de la competencia traductora, contempla el dinamismo de un proceso de desarrollo e integración de las subcompetencias (cfr. supra *Capítulo 5: 5.1.*).

2. Revisión de los enfoques de la didáctica de la traducción

Hemos revisado la evolución de los enfoques en la didáctica de la traducción escrita ya que es el otro marco de referencia en el que se ubica este trabajo (cfr. supra *Capítulo 5: 5.3.*). En este sentido, nos hemos centrado especialmente en revisar la enseñanza por objetivos de aprendizaje, el enfoque por tareas de traducción y la formación por competencias en traducción, dado que son éstos los enfoques que hemos integrado en nuestra propuesta. El enfoque por tareas y la formación por competencias comparten esa metodología activa que sitúa al alumno en el centro del proceso de aprendizaje, que es flexible y se adapta a los estudiantes y sus circunstancias, que integra todos los elementos principales de la enseñanza (estudiantes, profesores, objetivos, medios,

evaluación), y que permite una organización coherente y estructurada del aprendizaje, fijándose como meta la adquisición de competencias, no de contenidos.

3. Realización de un estudio entre traductores sobre la práctica profesional de la traducción inversa

Con el fin de aportar datos que dieran soporte a la selección de materias para la elaboración de las unidades didácticas, hemos presentado los resultados de un sondeo sobre la práctica profesional de la traducción inversa español-inglés entre 35 traductores autónomos con el fin de comprobar la oferta y demanda de estos servicios en España (cfr. supra *Capítulo 6: 6.2.4.1*).

3. DISEÑO DE UNA PROPUESTA DE USO DE CORPUS ELECTRÓNICOS EN LA ENSEÑANZA DE LA TRADUCCIÓN

En esta tesis hemos planteado el marco metodológico y de diseño curricular para el desarrollo de una competencia específica, “uso de corpus electrónicos para traducir”, que hemos ubicado dentro de la subcompetencia instrumental de la competencia traductora tal y como las define el grupo PACTE.

1. Identificación de cuatro componentes de la competencia “uso de corpus electrónicos para traducir”

Hemos identificado 4 componentes de la competencia “uso de corpus para traducir” (cfr. supra *Capítulo 6: 6.1*):

- (1) Identificar los principios básicos relacionados con corpus para resolver problemas de traducción.
- (2) Crear corpus para traducir.
- (3) Manejar software relacionado con corpus para traducir.
- (4) Resolver problemas de traducción utilizando corpus.

2. Planteamiento de tres tipos de objetivos para la propuesta didáctica

Aparte de los objetivos relacionados con el uso de corpus para traducir, hemos planteado otra clase de objetivos. El resultado ha sido 3 tipos de objetivos diferentes (cfr. supra *Capítulo 6: 6.2*):

- (1) Objetivos relacionados con el uso de corpus para traducir.
- (2) Objetivos relacionados con materias de traducción.
- (3) Objetivos relacionados con competencias genéricas.

3. Identificación de cuatro bloques de contenidos y desarrollo

A partir de los 4 componentes de la competencia de uso corpus para traducir hemos desarrollado en detalle 4 bloques de contenidos (cfr. supra *Capítulo 6: 6.3*):

- (1) Principios básicos relacionados con corpus.
 - Corpus de interés para el traductor
 - Tipos de corpus
 - Corpus gratuitos en Internet
 - Repositorios de textos gratuitos en Internet
 - La Web como corpus (*Internet pública*)
 - Principios básicos del trabajo con corpus
 - Construcción de corpus
 - Uso de corpus
- (2) Construcción de corpus.
 - Fuentes de donde recoger textos
 - Fuentes en papel.
 - Fuentes en formato electrónico excepto Internet
 - Fuentes en Internet
 - Espíritu crítico sobre las fuentes
 - Gestión
 - Importancia de nombrar lógicamente los archivos de un corpus
 - Importancia de ordenar lógicamente los archivos de un corpus
 - Construcción
 - Creación de corpus
 - Cálculo de la representatividad de un corpus
- (3) Software relacionado con corpus y operaciones.
 - Tipos de software
 - Software de conversión de textos en papel a formato electrónico
 - Software de conversión de formato PDF / imagen a texto
 - Software de ubicación o descarga de textos desde Internet
 - Software de etiquetado de textos
 - Software de alineación de textos
 - Software de conversión de un corpus paralelo en memoria de traducción
 - Software de extracción de terminología
 - Software de cálculo de la representatividad de un corpus
 - Software de extracción de concordancias de Internet
 - Software de análisis de corpus: programas de instalación y acceso local

- Software de análisis de corpus: funciones ofrecidas por el software de acceso local
 - Extracción de concordancias
 - Visualización del contexto
 - Reordenación del contexto
 - Extracción de colocaciones
 - Extracción de agrupaciones
 - Extracción de la progresión de aparición de una expresión
 - Extracción de listas de palabras
 - Alineación de textos
- (4) Resolución de problemas de traducción utilizando corpus.
- Selección del corpus más apropiado
 - Resolución de problemas de traducción de la lengua general y especializada
 - Información conceptual de la lengua general y especializada
 - Elementos microtextuales de la lengua general y especializada
 - Elementos macrotextuales
 - Elementos de discrepancia entre dos lenguas

4. Elaboración de siete unidades didácticas

De estos bloques de contenidos hemos seleccionado y agrupado algunos, lo que ha dado como resultado 7 unidades didácticas (cfr. *supra Parte III*):

- (1) “Introducción al uso de corpus electrónicos para traducir”
- (2) “Contrastes de uso en un corpus paralelo”
- (3) “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”
- (4) “Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”
- (5) “El corpus ad hoc como remedio para la traducción médica al inglés”
- (6) “Corpus y memoria de traducción”
- (7) “Sólo utilizar corpus; utilizar corpus solo”

Todas estas unidades presentan una amplia variedad de materias, contenidos y materiales.

a) Diversidad de materias

Las materias en las que las 7 unidades didácticas se ha puesto en práctica es variada:

- (1) Iniciación a la traducción directa (inglés B/C - español A).
- (2) Iniciación a la traducción inversa (español A - inglés B).
- (3) Traducción especializada técnica directa (inglés B - español A).
- (4) Traducción especializada científico-técnica inversa (español A - inglés B).

b) Diversidad de contenidos

Las 7 unidades diseñadas se han programado a través de tareas diversas. Estas tareas se pueden clasificar según el bloque de contenidos que tratan.

(1) Principios básicos relacionados con corpus.

Tareas relacionadas:

- “Los resultados que ofrece el corpus dependen de los ingredientes que contenga”
- “Un corpus más grande no es siempre un corpus mejor”
- “El ojo humano puede fallar”
- “El corpus como recurso documental más completo que otros”
- “Cómo hacer preguntas a un corpus”
- “Identificación de la fiabilidad de los textos o la importancia de ser ordenado”
- “Desde la perspectiva de un nativo”
- “Dar con el corpus (más idóneo)”
- “Frecuencias”
- “Mirar a la izquierda, a la derecha y a la izquierda otra vez”

(2) Construcción de corpus.

Tareas relacionadas:

- “De qué y cómo está hecho un corpus paralelo”
- “Encontrar fuentes y recursos”
- “Identificar el género textual”
- “Compilación de un corpus comparable bilingüe”
- “Del papel al formato electrónico”
- “Crear un corpus ad hoc: Descarga de documentos y sitios web”
- “Crear un corpus ad hoc: Aprovechar los archivos PDF”
- “De memoria de traducción a corpus paralelo”
- “Cómo mejorar las búsquedas en Internet y encontrar mejores textos”

(3) Software relacionado con corpus y operaciones.

Tareas relacionadas:

- “Funciones básicas de un programa de análisis de corpus para empezar a ver la utilidad de este recurso”
- La práctica totalidad de las tareas incluye algún contenido de software para su realización, aunque el énfasis no esté en la herramienta en sí, sino en la adquisición de principios sobre el uso de corpus, la creación de corpus o la resolución de problemas.

(4) Resolución de problemas de traducción utilizando corpus.

Tareas relacionadas:

- “Explotación real de un corpus en la traducción de un texto”
- “Uso de la raya o guión largo en inglés y español”
- “Búsqueda de equivalentes para ‘this’”
- “Búsqueda de equivalentes para la voz pasiva”
- “Encontrar equivalentes para diversos elementos de discrepancia”
- “Terminología diferente y terminología compartida”
- “Creatividad en un corpus técnico”
- “Aprovechar las pistas que da el corpus en lengua original (LO)”
- “Encontrar equivalencias utilizando el recurso creado”
- “Primer acercamiento al texto original”
- “Resolver la comprensión del tema. Cuando está escrito en nuestro idioma y aún así no lo entendemos”
- “Resolver el vocabulario específico del tema”
- “Resolver el vocabulario no específico del tema”
- “Resolver el resto de la traducción. Problemas que van más allá del nivel del vocabulario”
- “(Más) posibilidades de búsquedas en un corpus”
- “Navegando dentro del texto original, un sitio web”
- “Navegando dentro de textos paralelos”
- “Traducir mucho buscando poco”

c) Diversidad de materiales

Aparte de las tareas, en el marco de las 7 unidades mencionadas también hemos utilizado y elaborado materiales diversos, como se puede apreciar a continuación.

(1) “Introducción al uso de corpus electrónicos para traducir”

- Tareas = 7
- Fichas de trabajo = 8
- Fichas de evaluación = 3 (recogida de datos sobre el uso de corpus, autoevaluación, valoración)
- Textos originales y paralelos = 6-7
- Textos de apoyo = 1
- Concordancias = 1
- Recursos electrónicos:
 - *URLs* = 8
 - recursos lexicográficos = 5
 - corpus = 2

- otros = 1
- Corpus facilitados y su descripción = 2
- Programas informáticos = 1

(2) “Contrastes de uso en un corpus paralelo”

- Tareas = 5
- Fichas de trabajo = 4
- Fichas de evaluación = 3 (ejercicio de rellenar huecos, autoevaluación, valoración)
- Textos originales y paralelos = 2
- Recursos electrónicos:
 - URLs* = 3
 - recursos lexicográficos = 1
 - corpus = 2
- Programas informáticos = 1

(3) “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”

- Tareas = 8
- Fichas de trabajo = 6
- Fichas de evaluación = 3 (recogida de datos sobre el uso de corpus, autoevaluación, valoración)
- Textos originales y paralelos = 17
- Textos de apoyo = 1
- Concordancias = 1
- Listas de palabras = 1
- Recursos electrónicos:
 - URLs* = 2
 - otros = 2
- Corpus facilitados y su descripción = 1
- Programas informáticos = 3

(4) “Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”

- Tareas = 4
- Fichas de trabajo = 4
- Fichas de evaluación = 2 (autoevaluación, valoración)
- Textos originales y paralelos = 7
- Textos de apoyo = 2
- Concordancias = 1
- Colocaciones = 1
- Recursos electrónicos:

URLs = 22

- recursos lexicográficos = 9
- corpus = 9
- otros = 4

Corpus facilitados y su descripción = 3

Programas informáticos = 2

(5) “El corpus ad hoc como remedio para la traducción médica al inglés”

- Tareas = 5
- Fichas de trabajo = 6
- Fichas de evaluación = 1 (informe de valoración)
- Textos originales y paralelos = 4
- Recursos electrónicos:

URLs = 24

- recursos lexicográficos = 16
- corpus = 2
- otros = 6

Programas informáticos = 1

(6) “Corpus y memoria de traducción”

- Tareas = 5
- Fichas de trabajo = 5
- Fichas de evaluación = 2 (autoevaluación, valoración)
- Textos originales y paralelos = 4
- Textos de apoyo = 1
- Recursos electrónicos:

URLs = 2

- corpus = 1
- otros = 1

Programas informáticos = 4

(7) “Sólo utilizar corpus; utilizar corpus solo”

- Tareas = 4
- Fichas de trabajo = 4
- Fichas de evaluación = 3 (recogida de datos sobre el uso de corpus, valoración, autoevaluación)
- Textos originales y paralelos = 1
- Textos de apoyo = 1
- Recursos electrónicos:

URLs = 6

- otros = 6

Programas informáticos = 4

Finalmente, también hemos incluido una propuesta de evaluación de la misma, cuyas conclusiones se ofrecen en el siguiente apartado.

4. EVALUACIÓN DE LA PROPUESTA DISEÑADA DE USO DE CORPUS ELECTRÓNICOS EN LA ENSEÑANZA DE LA TRADUCCIÓN

1. Elaboración de indicadores de evaluación de la competencia “uso de corpus electrónicos para traducir”

Para el desarrollo de la competencia “uso de corpus para traducir” y las unidades didácticas diseñadas hemos elaborado los siguientes indicadores (cfr. supra *Capítulo 6: 6.7.2.*).

Al final del proceso de aprendizaje, el estudiante, según sus necesidades de documentación, ...

- 1: selecciona el corpus idóneo
 - 2: maneja distintos (tipos de) corpus que están disponibles en línea
 - 3: realiza búsquedas en archivos PDF como si fueran un corpus
 - 4: realiza una búsqueda de calidad (palabra/cadena de palabras idónea; búsqueda truncada, búsqueda con palabra de contexto)
 - 5: reordena los resultados de unas concordancias
 - 6: usa las funciones idóneas del programa informático
 - 7: selecciona el programa informático idóneo
 - 8: compila un corpus ad hoc de calidad
 - 9: maneja software de conversión de archivos PDF
 - 10: ubica y descarga textos electrónicos desde Internet
 - 11: maneja software de alineación de textos
 - 12: convierte un corpus paralelo en una memoria de traducción
 - 13: extrae concordancias bilingües en un programa de traducción asistida para utilizar una memoria de traducción a modo de corpus
 - 14: encuentra equivalentes a elementos de discrepancia entre el inglés y el español
 - 15: interpreta las frecuencias de aparición
 - 16: resuelve la traducción de elementos de la lengua general y especializada
 - 17: encuentra equivalentes terminológicos/colocaciones/fraseología en textos científico-técnicos
 - 18: justifica correctamente su opción de traducción basándose en datos extraídos del corpus
 - 19: emplea un tiempo razonable en resolver un problema
- - Aceptabilidad del equivalente propuesto

2. Diseño de herramientas de evaluación

Dentro de la propuesta de evaluación realizada hemos diseñado varias herramientas con dos fines (cfr. supra *Capítulo 6: 6.8.4*):

- (1) Herramientas para la evaluación del aprendizaje.
 - a. Diario de aprendizaje
 - b. Fichas de autoevaluación (con variantes, según la unidad didáctica)
 - c. Fichas de recogida de datos sobre el uso de corpus al traducir (con variantes, según la unidad didáctica)
 - d. Fichas de cruce de datos (con variantes, según la unidad didáctica)
- (2) Herramientas para la evaluación de la enseñanza.
 - a. Ficha de registro de observación
 - b. Fichas de valoración de la unidad didáctica

3. Datos recogidos

De las 7 unidades probadas en los 6 grupos y las 4 asignaturas descritos (cfr. supra *Introducción*), hemos recogido:

- 44 diarios de aprendizaje.
- 98 fichas de autoevaluación.
- 53 fichas de recogida de datos sobre el uso de corpus al traducir.
- 53 fichas de cruce de datos.
- 7 fichas de registro de observación (una por cada UD).
- 96 fichas de valoración de la unidad didáctica.

4. Tendencias observadas

De las fichas, tanto de las que recogían valoraciones numéricas como impresiones, hemos obtenido los siguientes datos y tendencias. Hay que resaltar que estas tendencias han podido observarse a pesar de que algunas de las preguntas de las fichas variaban dependiendo de los objetivos de cada unidad didáctica:

- Diarios de aprendizaje

De los 44 diarios de aprendizaje recogidos, queremos destacar los siguientes aspectos:

- 1) Es muy importante explicitar el objetivo del ejercicio o tarea e insistir sobre lo que se está aprendiendo (bien durante la tarea, bien al final de la misma, a modo de conclusión o reflexión) para que los estudiantes sean conscientes de su aprendizaje y fijen lo aprendido.
- 2) Los estudiantes valoran de manera especialmente positiva el uso de recursos electrónicos y, en concreto, el uso de corpus en combinación con algún programa de análisis por permitirles consultar muchos textos a la vez y ahorrar tiempo. Así, por ejemplo, la estudiante B.C. apunta en su diario que “programas como *Inforapid* o *WordSmithTools* [...] son un gran apoyo para los traductores y además nos permiten ahorrar mucho tiempo. Para las traducciones de textos técnicos como es el caso del texto del microondas, este tipo de programas ha sido de gran ayuda para poder llegar a la solución exacta, ya que podíamos consultar textos de diferentes compañías de electrodomésticos y ver las diferentes propuestas”.
- 3) Los estudiantes han asimilado la intención de esta propuesta didáctica. Por ejemplo, el estudiante I.M., también de la UD dedicada a la traducción técnica directa anota que “...más que aprender cuestiones de vocabulario o léxico de especialidad, he aprendido un poco más a documentarme, a moverme, a darle mil vueltas a una traducción y siempre ver algo que no me acaba de convencer, a mirar en mil fuentes distintas, a no quedarme con la primera opción que encuentre sino a comparar con otras fuentes, en definitiva, considero que más bien he adquirido aptitudes y desarrollado habilidades”.

➤ **Fichas de autoevaluación**

De las 98 fichas de autoevaluación recogidas y en lo que respecta a las capacidades (apartado “Soy capaz de...”, con variantes en cada ficha según los contenidos tratados en cada unidad didáctica), los alumnos admiten haber ganado más confianza en los siguientes aspectos, que destacan por su concreción:

- 1) Manejo de software de ubicación y descarga de documentos desde Internet.
- 2) Distinción entre las búsquedas para las que es más apropiado utilizar un corpus grande y general, y las que requieren de un corpus pequeño y especializado.
- 3) Comprensión de la importancia de nombrar y ordenar los archivos de un corpus de una forma lógica transparente.
- 4) Conversión de archivos en formato PDF a formato TXT para su posterior aprovechamiento.

En lo que respecta a la utilidad de los corpus (apartado “Te ha ayudado el corpus a...”, con preguntas más parecidas o coincidentes entre unas unidades didácticas y otras), se han observado tendencias claras. Así, entre los elementos más valorados por los alumnos están:

- 1) Estar más seguro de una solución de traducción. Por ejemplo, en las unidades “Introducción al uso de corpus electrónicos para traducir”, donde este aspecto obtuvo una nota media de 7,8; “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos” (8,3), “Corpus y memoria de traducción” (9,2) y “Dermovet, 1ª puesta en práctica” (8,3).
- 2) Comprender la importancia de las colocaciones. Por ejemplo, en las unidades “Introducción al uso de corpus electrónicos para traducir” (8,5) y “Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés” (8,5).
- 3) Encontrar equivalentes de traducción. Por ejemplo, en las unidades “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos” (8,5) y “Dermovet, 1ª puesta en práctica” (8,2), excepto en la unidad “Introducción al uso de corpus electrónicos para traducir”, donde fue uno de los elementos menos puntuados (5,7).

Por otro lado, los aspectos menos valorados por los alumnos (aunque con notas muy altas) en cuanto al uso de corpus fueron:

- 1) Traducir mejor. Por ejemplo, en las unidades “Introducción al uso de corpus electrónicos para traducir” (6,5), “Contrastes de uso en un corpus paralelo” (7,4) y “Dermovet, 1ª puesta en práctica” (7,5), excepto en la unidad “Corpus y memoria de traducción”, donde fue el elemento con mayor puntuación (9,2). Nótese aquí que aunque los alumnos no consideren que el uso de corpus les ha ayudado a traducir mejor de una forma global en estas unidades, sí consideran, como se puede observar en los elementos más valorados, que estos recursos les han ayudado a encontrar equivalentes de traducción.
- 2) Ser más consciente de la naturaleza de los textos utilizados en la fase de documentación. Por ejemplo, en las unidades “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos” (7,8), “Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés” (7,5), “Dermovet, 1ª puesta en práctica” (7,4), y “Corpus y memoria de traducción” (8,2).

- 3) Comprender mejor el texto original. Por ejemplo, en las unidades “Corpus y memoria de traducción” (7,5) y “Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés” (7,7). Nótese aquí que en estas unidades el uso de corpus no estaba orientado hacia este propósito, por lo que quizás la pregunta no debería haberse incluido en la ficha.

➤ **Fichas de recogida de datos sobre el uso de corpus al traducir**

La información de las 53 fichas obtenidas proporciona datos sobre los pasos que ha seguido el estudiante al utilizar un corpus para resolver un problema de traducción. Estos datos se toman para después cruzarlos con la aceptabilidad de la solución y el valor de los indicadores de evaluación en la ficha de cruce de datos. Tras la primera puesta en práctica de la primera UD, se constató lo siguiente:

- 1) La necesidad de pedir al estudiante que rellenase esta ficha a la vez que traducía un texto utilizando corpus ya que si se le solicitaba este ejercicio después de la traducción, se perdían datos.

➤ **Fichas de cruce de datos**

En estas fichas se han cruzado los datos sobre cómo ha usado un corpus el estudiante a la hora de resolver un problema de traducción, la aceptabilidad del resultado y el valor asignado a cada indicador evaluado, con el fin de obtener una nota numérica indicativa de la eficiencia del alumno en el uso de corpus al traducir. El resultado del análisis de las 53 fichas de cruce de datos obtenidas es el siguiente:

- 1) La nota media en el uso de corpus de todos los alumnos que utilizaron corpus cuando traducían es un 8. Para llegar a este resultado hemos calculado la nota media de las puntuaciones obtenidas de todas las fichas de cruce de datos de las unidades “Introducción al uso de corpus electrónicos para traducir”, “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos” y “Sólo utilizar corpus; utilizar corpus solo”, que es en las que se realizó este tipo de evaluación. Tras haber evaluado las fichas de recogida de datos sobre el uso de corpus al traducir, hemos dividido la suma de las todas puntuaciones obtenidas entre el número de personas evaluadas, lo cual ha dado un resultado de 8 puntos sobre 10, lo cual consideramos una valoración muy positiva.

➤ **Fichas de registro de observación**

De las 7 fichas cumplimentadas por la docente, una para cada unidad didáctica probada, se han podido extraer las siguientes conclusiones:

- 1) Al principio es necesario guiar muy de cerca al estudiante en su aprendizaje para que éste no se sienta abrumado por la cantidad de datos que puede llegar a obtener, especialmente cuando la metodología de la lingüística de corpus le resulta nueva. Así, por ejemplo, en la ficha de registro con fecha 04/03/03 y perteneciente a la UD de introducción, se encuentra el siguiente comentario: “Empleamos 2 horas en repasar lo que hemos visto, que acaben de rellenar la Ficha 1, la comenten entre ellos y lleguen a una conclusión”.
- 2) El alumno necesita tiempo para asimilar lo que es un corpus y para entender esta nueva forma de aproximarse a la lengua. Por ejemplo, en la ficha del 08/04/03 de la misma UD aparece este comentario: “Pensé que sería muy fácil pero ocupamos todo el tiempo con el ejercicio de la Tarea 5 e incluso hacemos una parte todos juntos en voz alta”. En relación con esto, apuntamos también en la ficha las soluciones que pueden ser de utilidad y que después se probarán:
 - Mostrar al estudiante cómo es un corpus por dentro, es decir, enseñarle los archivos que lo componen.
 - Ver la fuente de donde se han extraído los textos que componen un corpus (p.ej. mirar la edición en Internet del *Financial Times* para entender el tipo de textos que puede contener un corpus ad hoc compuesto por textos de este diario).

➤ Fichas de valoración de la unidad didáctica

Aunque en su momento ya se comentaron los puntos destacables de las fichas de cada unidad, del total de 96 fichas de valoración recogidas (*Volumen 2*) se han podido extraer algunas tendencias comunes a todas las unidades.

En lo que respecta a la novedad de esta metodología, las conclusiones a las que ha llegado la mayoría de los estudiantes son:

- 1) La utilización de corpus electrónicos para documentarse ha supuesto una novedad para la práctica totalidad de los estudiantes. Los estudiantes reconocen que, hasta el momento, sus habilidades de documentación se basaban principalmente en consultar diccionarios y páginas web pero no de una forma eficaz. La aplicación de la metodología de corpus, acompañada de ejercicios para optimizar la búsqueda de documentos e información, y apoyada por programas informáticos, les ha permitido resolver problemas de traducción al poder manejar una gran cantidad de textos paralelos, hacer

búsquedas y obtener resultados sin ruido, observar términos en contexto, descubrir su significado, observar su uso con la ayuda de datos de frecuencia, y, sobre todo, no quedarse en el nivel de la palabra sino del término compuesto, el sintagma, las agrupaciones, las colocaciones o las convenciones de género.

- 2) El uso de corpus puede ser indispensable para resolver algún tipo de problemas como, por ejemplo, cuestiones de frecuencia de uso o convenciones de escritura o género.
- 3) Una búsqueda puede dar pistas o conducir hacia la resolución de dudas que ni siquiera estaban planteadas. Esto se puede dar cuando se realiza una búsqueda y se observan los resultados de una forma amplia, es decir, no centrándose exclusivamente en aquella palabra o elemento microtextual que ocupa la atención en ese momento, sino mirando más allá para aprovechar al máximo unos resultados.
- 4) No sólo es importante la calidad de los textos consultados, sino también de la cantidad, en especial para poder extraer patrones de uso.
- 5) La construcción de corpus resulta útil y conduce a un proceso de documentación más eficaz. Aunque a priori la inversión de tiempo en crear un corpus de textos de calidad parece excesiva, los estudiantes reconocen que, a la hora de traducir textos especializados, compensa.
- 6) La utilización de corpus en el aula ha desvelado nuevos usos de este recurso en traducción, como es evitar errores causados por la cercanía de dos lenguas al poder observar los contextos de uso. Este caso se pudo observar en la UD de introducción al uso de corpus, en el que una alumna cuya lengua materna era el catalán encontró ayuda en los corpus para salvar algunos calcos con el castellano.

En lo que respecta a lo que los alumnos han destacado de forma más positiva del uso de corpus, las conclusiones son:

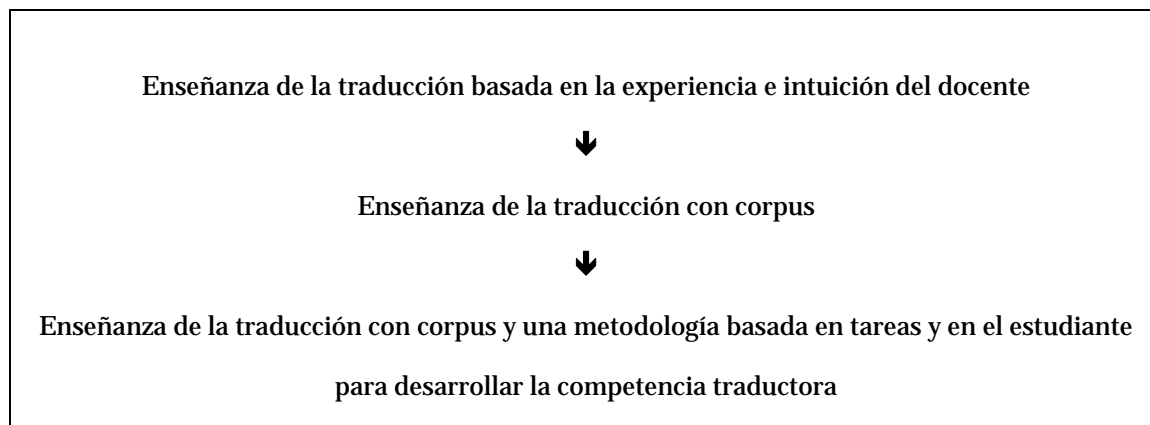
- 1) El enfoque es innovador, con una clara orientación hacia el alumno y en un excelente clima de cooperación entre docente y estudiantes.
- 2) Los contenidos se transmiten de una forma eminentemente práctica y el aprendizaje de recursos y herramientas es significativo, contextualizado en materias de traducción y llevado a cabo en aulas multimedia. De entre los programas informáticos más valorados están *WordSmith Tools*, *AntConc*, *InfoRapid* y *TagEditor* de Trados.

- 3) El uso de corpus ayuda a mejorar la calidad de las traducciones ya que permite encontrar equivalentes de traducción superando el nivel de la palabra, las búsquedas en el diccionario o en buscadores generales en Internet
- 4) El uso de corpus ayuda a sentirse más seguro de una solución, aunque en algunas ocasiones también provoca que el estudiante se cuestione una solución que ya había aportado.
- 5) El uso de corpus supone un ahorro de tiempo a la hora de traducir una vez se ha acabado el proceso de compilación.
- 6) El uso de corpus permite aprender sobre las lenguas implicadas en una traducción y sobre la traducción en sí misma.
- 7) Algunas de las habilidades adquiridas son especialmente útiles, como la identificación de errores y de términos creativos en corpus, las búsquedas por lema o por categoría gramatical o búsquedas en una memoria de traducción.

En lo que respecta a las sugerencias realizadas, destacan las siguientes:

- 1) Mayor práctica de traducción de textos en sentido tradicional. En ocasiones, los estudiantes, a pesar de reconocer la utilidad de las fichas y ejercicios previos a una traducción o en lugar de ésta, les gustaría traducir más textos completos.
- 2) Construcción de corpus propios y mayor independencia de trabajo, en aquellas unidades de iniciación en el uso de corpus para traducir, lo cual lleva a plantearse desdibujar las líneas entre la fase de iniciación y consolidación en el uso de corpus.
- 3) Los estudiantes deberían haber estado expuestos a contenidos relacionados con corpus desde el comienzo de sus estudios de traducción. Hemos comprobado, para nuestra satisfacción, que la práctica totalidad de los alumnos valoran muy positivamente los contenidos y organización de todas y cada una de las unidades didácticas e incluso enfatizan la utilidad de lo aprendido y lo adquirido para sus circunstancias presentes y futuras, a la vez que lamentan no haber estado expuestos a la metodología de corpus con anterioridad en sus estudios. Así, por ejemplo, la alumna MA.G., en la UD “Corpus y memoria de traducción”, apunta: “No puedo decir nada que no sea positivo, ya que esta asignatura me ha hecho sentir mucho más segura a la hora de abordar una traducción inversa. Lo único que me gustaría comentar es que es una pena que recursos tan interesantes como los que hemos aprendido no se hayan introducido en cursos anteriores”.

Como se puede comprobar, la propuesta diseñada y probada en el aula ha sido evaluada y, por todo lo expuesto, se puede concluir que la propuesta parece funcionar tanto a nivel de satisfacción del alumnado con lo que han adquirido como a nivel de los resultados del aprendizaje. Por lo tanto, parece apropiado sugerir la necesidad de que se dé la siguiente evolución:



PERSPECTIVAS

Las vías que se abren hacia el futuro con nuestra investigación son:

- (1) En lo académico, adaptar la presente propuesta a otras materias de traducción como, por ejemplo, la traducción especializada económica y jurídica.
- (2) Por otro lado, creemos importante hacer llegar la filosofía de esta propuesta a un profesorado de traducción que basa gran parte de su docencia y evaluación en su intuición y aún no utiliza la metodología de corpus, que ayudaría a sistematizar su enseñanza. En este sentido, esbozamos aquí el diseño de una herramienta orientada a los profesores, especialmente de traducción, que combinen ésta con el uso de recursos electrónicos como corpus en sus clases. Esta herramienta, que hemos bautizado como “ADICT” (Aplicación Didáctica de Corpus en Traducción), sería una aplicación informática basada en lenguaje HTML o en Excel y serviría de guía para el profesor en la elaboración de unidades didácticas en las que se utilicen corpus, yendo desde la selección de competencias o subcompetencias que se desean trabajar hasta los indicadores de la evaluación, pasando por los contenidos o los materiales. Se trataría de una aplicación sencilla, personalizable y que acompañaría al profesor paso a paso en el diseño de una unidad didáctica o una tarea relacionada con el uso de corpus para traducir.

A continuación mostramos algunas de las pantallas de este posible proyecto de futuro:

BIENVENIDO A


ADICT

(Aplicación Didáctica de Corpus a la Traducción)

ADICT es una aplicación basada en lenguaje HTML puesta al servicio de profesores de traducción. Por el momento, está dirigida sobre todo a aquellos que trabajan en asignaturas...

- de inglés y español
- de iniciación a la traducción o traducción especializada
- de directa o inversa
- donde se integra el uso de las TIC
- ... y con un potencial para la aplicación de corpus electrónicos

Esperamos que te vuelvas ADICT-o a esta aplicación!!



ENTRA

ESTACIÓN DE TRABAJO DEL PROFESOR

~ Opciones ~

- [Listar todas las unidades didácticas y tareas ya diseñadas](#)
- [Buscar unidades didácticas / tareas ya diseñadas](#)
- [Crear nuevas unidades didácticas / tareas](#)

ESTACIÓN DE TRABAJO DEL PROFESOR

~ Crear una unidad didáctica ~

Título de la unidad

Objetivo general de la unidad

Competencias que se quieren trabajar

Indicadores de la competencia para la evaluación:

Materiales

CORPUS GENERALES:

BYC - British National Corpus
 BoE - Bank of English
 CREA - Corpus de Referencia del Español Actual
 CORDE - Corpus Diacrónico del Español
 Corpus del español
 COMPARA

[añadir corpus](#)

CORPUS ESPECIALIZADOS:

Business Letter Corpus
 Corpus Técnico de DULC
 CLUSO - Corpus Lingüístico da Universidade de Vigo
 Project Gutenberg

[añadir corpus](#)

Desarrollo

Tiempo estimado de duración

Comentarios

- (3) Por último, no olvidamos las enormes posibilidades que ofrece la metodología de corpus en la enseñanza a distancia, como ya se ha apuntado en esta tesis. Por las características del trabajo con corpus, en formato electrónico, con la posibilidad de utilizar recursos y software gratuito y las condiciones ideales para que se dé un aprendizaje por descubrimiento, que puede hacerse de una forma guiada o autónoma, la enseñanza de la traducción con corpus electrónicos puede llegar a ofrecerse en lugares donde no existe tradición en la enseñanza de la traducción.

En el mundo profesional se podría difundir el uso de la metodología de corpus a nivel de agencias de traducción y la formación de su personal en la recopilación y preparación de textos para alimentar corpus que deberían saber manejar. Tanto en el sector académico como el profesional, resta seguir buscando maneras y recursos que permitan la construcción de corpus de calidad en muy poco tiempo y software gratuito o libre para manejarlos fácilmente.



Como punto final, nos gustaría resaltar uno de los principios básicos del uso de corpus que también se aplica al uso de corpus en la enseñanza de la traducción. Independientemente de la calidad y tamaño de un corpus o la potencia del software con el que se maneje, la función del profesor, estudiante o investigador a la hora de interpretar aquello que nos ofrece un corpus es insustituible. Aquél que desee obtener ayuda de un corpus deberá haber sido capaz primero de plantearse las preguntas adecuadas y hacerlo de tal forma que de la consulta a dicho corpus se puedan extraer resultados interpretables.

GLOSARIO

- **AGRUPACIÓN (*CLUSTER*)**

“Clusters are words which are found repeatedly together in each others' company, in sequence” (Scott 2008).

- **COLIGACIÓN o patrones gramaticales (*COLLIGATION*)**

“[It is] (1) the grammatical company a word keeps, or avoids keeping, either within its own group or at a higher rank; (2) the grammatical functions that a word's group prefers or avoids; (3) the place in a sequence that a word prefers or avoids” (Hoey 2004).

- **COLOCACIÓN (*COLLOCATION*)**

“Collocations are groups of words which frequently appear in the same context” (Oakes 1998: 149).

- **COMPONENTE DE UN CORPUS**

“Conjunto de textos homogéneos derivados de un corpus o un subcorpus” (Corpas 2001: 158).

- **CONCORDANCIA (*CONCORDANCE; KWIC CONCORDANCE*)**

“[It] is a list of unconnected lines of text, which have been summoned by the concordance program from a computer corpus” (Partington 1998).

- **CORPUS**

“Un corpus es un conjunto homogéneo de documentos lingüísticos de cualquier tipo (orales, escritos, literarios, coloquiales, etc.) que se toman como modelo de un estado o nivel de lengua predeterminado, al cual representan o se pretende que representen”. (Alvar Ezquerro, Blanco Rodríguez y Pérez Lagos 1994: 9-10).

“A collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language” (Sinclair 1996).

“...any computer-readable collection of texts or transcripts which can be accessed and interrogated selectively using text-retrieval or concordancing software” (Aston 1997: 205).

- **CORPUS ANOTADO**

Corpus cuyos documentos han sido etiquetados lingüística o metatextualmente, ya sea de forma manual o de forma automática (Corpas 2001: 159).

- **CORPUS CANÓNICO**

“Corpus que contiene las obras completas de un determinado autor” (Corpas 2001: 158).

- **CORPUS COMPARABLE**

“Corpus que, en relación a otro u otros corpus de lenguas distintas, incluyen tipos similares de textos originales. (...) Se habla de corpus comparable bilingüe cuando se incluyen dos lenguas, y corpus comparable multilingüe, cuando hay más de dos” (Corpas 2001: 158).

- **CORPUS DE REFERENCIA**

“Corpus que no incorpora documentos enteros, sino fragmentados, ya que el interés no reside en el texto en sí, sino en el estadio de lengua representado” (Corpas 2001: 158).

- **CORPUS AD HOC, DESECHABLE O “DE FABRICACIÓN PROPIA” (*AD HOC CORPUS; DISPOSABLE CORPUS; DIY CORPUS*)**

“These are small, specialized corpora created ad-hoc to serve the needs of the translator for a specific translation project, and their value lies not only in their analysis but even more so in their creation” (Zanettin 2002: 239).

- **CORPUS DIACRÓNICO**

“Corpus que incluye textos pertenecientes a etapas sucesivas en el tiempo con objeto de estudiar la evolución lingüística” (Corpas 2001: 158).

- **CORPUS DOCUMENTADO**

“Corpus en el que cada documento textual lleva asociado un archivo DTD (Document Type Definition) o una cabecera (header) descriptiva de su procedencia” (Corpas 2001: 159).

- **CORPUS EQUILIBRADO**

“Corpus que contiene diversas variedades de lengua en porcentajes similares (Corpas 2001: 158).

- **CORPUS ESPECIALIZADO**

“Corpus representativo de un tipo particular de lengua, que incluye textos pertenecientes exclusivamente a tal variedad” (Corpas 2001: 158).

- **CORPUS GENERAL**

“Corpus representativo de la lengua común, que incluye una gran variedad de textos producidos en situaciones comunicativas cotidianas” (Corpas 2001: 158).

- **CORPUS GENÉRICO**

“Corpus representativo de un determinado género, que incluye textos pertenecientes a dicho género” (Corpas 2001: 158).

- **CORPUS GRANDE O EXTENSO**

“Corpus cuya extensión no está delimitada voluntariamente y que recoge, típicamente, un número de palabras muy elevado, sin tener en cuenta aspectos como el equilibrio o la representatividad” (Corpas 2001: 158).

- **CORPUS LÉXICO**

“Corpus que recoge documentos formados por fragmentos muy pequeños de texto de igual longitud” (Corpas 2001: 159).

- **CORPUS MONITOR**

“Corpus que mantiene un volumen textual constante, pero actualizado, al ir añadiendo elementos nuevos a la vez que elimina cantidades equivalentes de material antiguo” (Corpas 2001: 158).

- **CORPUS NO ANOTADO**

“Corpus en formato de sólo texto o ASCII, que no ha sido etiquetado de forma alguna y, por tanto, presenta un alto grado de simplicidad” (Corpas 2001: 159).

- **CORPUS NO DOCUMENTADO**

“Corpus cuyos documentos integrantes carecen de tales archivos asociados o cabeceras” (Corpas 2001: 159).

- **CORPUS PARALELO**

“Corpus formado por una serie de textos en la lengua de origen junto con sus traducciones en una (o varias) lengua(s) meta. Se habla de corpus paralelo bilingüe cuando hay sólo dos lenguas implicadas, y de corpus paralelo multilingüe cuando se trata de más de dos” (Corpas 2001: 158).

- **CORPUS PERIÓDICO O CRONOLÓGICO**

“Corpus que incluye textos producidos durante un determinado periodo de tiempo” (Corpas 2001: 158).

- **CORPUS PIRAMIDAL**

“Corpus que contiene textos distribuidos por niveles, los cuales se caracterizan por aumentar progresivamente la complejidad de las variedades temáticas, en detrimento del número de textos incluidos en cada variedad” (Corpas 2001: 158).

- **CORPUS TEXTUAL**

“Corpus que incluye textos completos con objeto de ofrecer una muestra representativa de la lengua común, así como de sus variedades más importantes, de manera que pueda ser utilizado como base para la elaboración de gramáticas, diccionarios y otras obras de referencia” (Corpas 2001: 158).

- **ELICITACIÓN (*ELICITATION*)**

En lingüística, “native speaker judgements about the acceptability of grammatical and vocabulary usage” (Flowerdew 1996: 103).

- **MANEJO DE CORPUS EN TRADUCCIÓN (*CORPUS MANAGEMENT IN TRANSLATION*)**

“[It is] the knowledge and skills needed in the compilation and use of corpus information for individual translation assignments” (Varantola 2003).

- **PREFERENCIA COLOCACIONAL (*COLLOCATIONAL PRIMING*)**

Preferencia que una palabra o expresión muestra por la presencia de otras palabras determinadas a su alrededor. Cambia con el tiempo y es sensible al campo temático en el que aparece la palabra o expresión (adaptado de Hoey 2004).

- **PREFERENCIA SEMÁNTICA (*SEMANTIC PREFERENCE*)**

“Semantic preference” follows “collocation” and “colligation” as a level of increasing abstraction in Sinclair’s description of extended units of meaning. It indicates that certain lexical items tend to be accompanied not only by particular words (“collocation”) or word classes (“colligation”) but also by items that, although belonging to different grammatical classes, still remain within the same semantic field. An example is the phrase naked eye, which is almost always preceded by a word or phrase having to do with visibility (Sinclair 1996, en Palumbo 2002).

- **PROSODIA SEMÁNTICA (*SEMANTIC PROSODY*)**

“Semantic prosody refers to the connotation, be it positive, negative, or neutral, that is expressed by a lexical item in association with its collocates” (Bowker 2000: 35).

- **SUBCORPUS**

“Selección de textos derivada de un corpus más general y complejo” (Corpas 2001: 158).

- **TIPOS (*TYPES*) y FORMAS (*TOKENS*)**

“[...] a ‘type’ is the base form of a word, as you might find it listed in a dictionary, whereas a ‘token’ is the actual word, each occurrence of which can be counted” (Hughes 1996: 163).

ÍNDICE DE TABLAS, DOCUMENTOS Y FIGURAS

Tabla 1. Aspectos principales del diseño de corpus (adaptado de Torruella y Llisteri 1999).....	25
Tabla 2. Requisitos de un corpus para su uso en la traducción especializada (Bowker 1996)	103
Tabla 3. Estrategias de búsqueda de información en un corpus monolingüe en lengua meta.....	107
Tabla 4. Estructuración de la unidad didáctica (Hurtado 1999).....	133
Tabla 5. Estructuración de la tarea de traducción (Hurtado 1999)	133
Tabla 6. Categorías de competencias específicas en la formación de traductores (Hurtado 2007)	141
Tabla 7. Competencia traductora y uso de corpus electrónicos para traducir	144
Tabla 8. Objetivos generales y específicos de la iniciación en el uso de corpus para traducir	149
Tabla 9. Objetivos generales y específicos de la consolidación en el uso de corpus para traducir	150
Tabla 10. Contenido disciplinario asociado a la competencia “uso de corpus electrónicos para traducir”	155
Tabla 11. Tipos de información requerida y recursos donde encontrarla (Montalt 2005)	169
Tabla 12. Variedad de algunos elementos incluidos en las unidades didácticas elaboradas.....	174
Tabla 13. Progresión de las unidades didácticas según el nivel de uso de corpus	175
Tabla 14. Progresión de las unidades didácticas según la dirección y el nivel de competencia en traducción.....	176
Tabla 15. Ejemplo de selección de materiales de la unidad didáctica “El corpus ad hoc como remedio para la traducción médica al inglés”	190
Tabla 16. Indicadores susceptibles de ser evaluados respecto al uso de corpus para traducir.....	198
Tabla 17. Elementos incluidos en la presentación de cada unidad didáctica.....	226
Tabla 18. Resultados de la Ficha de autoevaluación de la UD “Introducción al uso de corpus electrónicos para traducir”	246
Tabla 19. Resumen de resultados de la Ficha de cruce de datos de la UD “Introducción al uso de corpus electrónicos para traducir”	248
Tabla 20. Resultados de la Ficha de autoevaluación de la UD “Contrastes de uso en un corpus paralelo”	262
Tabla 21. Notas medias resultantes de puntuar la Ficha 5 de la UD “Contrastes de uso en un corpus paralelo”	264
Tabla 22. Resultados de la Ficha de autoevaluación de los alumnos de la UD “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”	281
Tabla 23. Resumen de resultados de la Ficha de cruce de datos de la UD “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”	287
Tabla 24. Resumen de las Fichas de autoevaluación de la UD “Usar corpus para concienciar sobre el funcionamiento de la traducción inversa al inglés”	290
Tabla 25. Resumen de las Fichas de autoevaluación de la UD “Corpus y memoria de traducción” ..	329
Tabla 26. Resultados de la Ficha de autoevaluación de los alumnos de la UD “Dermovet, 1ª puesta en práctica”	345
Tabla 27. Resumen de resultados de la Ficha de cruce de datos de la UD “Dermovet, 1ª puesta en práctica”	349



Documento 1. Cuestionario dirigido a empresas de traducción	56
Documento 2. Respuestas al cuestionario dirigido a empresas de traducción.....	57
Documento 3. Cuestionario sobre la práctica de la traducción inversa (español-inglés) en España ..	152
Documento 4. Diario de aprendizaje del estudiante	201
Documento 5. Diario de aprendizaje cumplimentado por un alumno	202
Documento 6. Ficha de autoevaluación cumplimentada por una alumna	203
Documento 7. Ficha de recogida de datos sobre el uso de corpus cuando se traduce	205
Documento 8. Ficha de cruce de datos.....	206
Documento 9. Texto original para traducir en la tarea final de la unidad “Dermovet, 1ª puesta en práctica”	211
Documento 10. Ficha de recogida de datos cumplimentada por dos estudiantes.....	213
Documento 11. Ficha de cruce de datos cumplimentada por la docente para la evaluación del uso de corpus de dos alumnas	214
Documento 12. Ficha de registro de observación cumplimentada por la docente	217
Documento 13. Ficha de valoración cumplimentada por un estudiante	217
Documento 14. Ficha de valoración de la UD “Utilidad de un corpus comparable bilingüe especializado para traducir textos técnicos”	283
Documento 13. Ficha de valoración de la UD “Dermovet, 1ª puesta en práctica”.....	346



Figura 1. Modelo holístico de la competencia traductora (PACTE 2003)	121
Figura 2. Modelo dinámico de la adquisición de la competencia traductora (PACTE 2000)	123
Figura 3. Tipo de clientes que encargan traducciones inversas español-inglés en España.....	152

BIBLIOGRAFÍA

Las referencias bibliográficas de esta tesis están organizadas en tres grandes categorías:

- Lingüística de corpus y temas afines
- Traducción, estudios sobre la traducción, didáctica de lenguas y traducción y competencia traductora
- Conferencias no escritas

LINGÜÍSTICA DE CORPUS Y TEMAS AFINES

AARTS, J.; Meijs, W. (eds.). (1984). *Corpus linguistics: recent developments in the use of computer corpora in English language research*. Ámsterdam: Rodopi.

AARTS, J.; Meijs, W. (eds.). (1986). *Corpus Linguistics II*. Ámsterdam: Rodopi.

ADOLPHS, S. (2008). *Corpus and context: investigating pragmatic functions in spoken discourse*. Ámsterdam; Filadelfia: John Benjamins. (Studies in Corpus Linguistics; 30).

AHMAD, K.; Holmes-Higgin, P.; Abidi, S.R. (1994). "A description of texts in a corpus: 'virtual' and 'real' corpora". En: W. Martín; *et al.* (eds.). *EURALEX-94: proceedings*. Ámsterdam: Vrije Universiteit, pp. 390-402.

AIJMER, K. (1984). "Sort of and kind of in English conversation". En: *Studia Linguística* (ISSN 0039-3193). 38(1), pp. 117-128.

ALTENBERG, B. (1984). "Causal linking in spoken and written English". En: *Studia Linguística* (ISSN 0039-3193). 38(1), pp. 20-69.

ALVAR EZQUERRA, M.; Blanco Rodríguez, M.J.; Pérez Lagos, F. (1994). "Diseño de un corpus español en el marco de un corpus europeo". En: M. Alvar Ezquerra; J.A. Villena Ponsoda (coord.) *Estudios para un corpus del español*. Málaga: Universidad de Málaga, pp. 9-30.

ARCHER, D.; *et al.* (eds.) (2003). *Proceedings of the Corpus Linguistics 2003 conference*. Lancaster: Lancaster University. (UCREL Technical Papers; 16). URL: <http://ucrel.lancs.ac.uk/publications/CL2003/papers/> [Consulta: 11 de julio de 2008].

ASTON, G. (1995). "Corpora in language pedagogy: matching theory and practice". En: G. Cook y B. Seidlhofer (eds.). *Principle and practice in applied linguistics. Studies in honour of H.G. Widdowson*. Oxford: Oxford University Press, pp. 257-270.

ASTON, G. (1997a). "Enriching the learning environment: corpora in ELT". En: A. Wichmann; *et al.* (eds.). *Teaching and Language Corpora*. Londres; Nueva York: Longman, pp. 51-64.

- ASTON, G. (1997b). "Small and large corpora in language learning". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.). *PALC'97: Practical Applications in Language Corpora*. Lodz: Lodz University Press, pp. 51-62.
- ASTON, G. (1997c). "Involving learners in developing learning methods: exploiting text corpora in self-access". En: P. Benson; P. Voller (eds.). *Autonomy & independence in Language Learning*. Londres; Nueva York: Longman, pp. 204-214.
- ASTON, G. (1999). "Corpus Use and Learning to Translate". En: *Textus: English studies in Italy: rivista dell'Associazione italiana di anglistica* (ISSN 1824-3967). 12(2), pp. 289-314.
- ASTON, G. (2000a). "Corpora and language teaching". En: L. Burnard; T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective*. Bern: Peter Lang, pp. 7-17.
- ASTON, G. (2000b). "I corpora come risorse per la traduzione e per l'apprendimento". En: S. Bernardini; F. Zanettin (eds.). *I corpora nella didattica della traduzione: Corpus Use and Learning to Translate*. Bologna: CLUEB, pp. 21-29.
- ASTON, G. (ed.). (2001a). *Learning with corpora*. Bologna: CLUEB.
- ASTON, G. (2001b). "Learning with corpora: an overview". En: G. Aston (ed.). *Learning with corpora*. Bologna: CLUEB, pp. 7-45.
- ASTON, G. (2002). "The learner as corpus designer". En: B. Ketteman; G. Marko (eds.). *Teaching and learning by doing corpus analysis: Proceedings of the Fourth International Conference on Teaching and Language Corpora, Graz 19-24 July, 2000*. Ámsterdam; Nueva York: Rodopi, pp. 9-25.
- ATKINS, S.; Clear, J.; Ostler, J. (1992). "Corpus design criteria". En: *Literary and Linguistic Computing* (ISSN 0268-1145). 7(1), pp. 1-16.
- ATWELL, E.; *et al.* (2003). "Detecting student copying in a corpus of science laboratory reports [en línea]: simple and smart approaches". En: D. Archer; *et al.* (eds.). *Proceedings of the Corpus Linguistics 2003 conference*. Lancaster: Lancaster University. (UCREL Technical Papers; 16). URL: <http://ucrel.lancs.ac.uk/publications/CL2003/papers/atwellgent.pdf> [Consulta: 29 de junio de 2008].
- AUSTERMÜHL, F. (2001). *Electronic tools for translators*. Manchester: St. Jerome. (Translation Practices Explained Series).
- BADIA, T.; *et al.* (1999). "LETRAC curriculum modules" [en línea]. En: IAI. *LETRAC: Language Engineering for Translator Curricula*. [S.l.]: IAI, 26-01-2006. URL: <http://www.iai.uni-sb.de/docs/D3.pdf> [Consulta: 31 de marzo de 2007].
- BAKER, M. (1993). "Corpus Linguistics and Translation Studies: Implications and Applications". En: M. Baker; G. Francis; E. Tognini-Bonelli (eds.). *Text and Technology: in Honour of John Sinclair*. Ámsterdam; Filadelfia: John Benjamins, pp. 233-250.
- BAKER, M. (1995). "Corpora in translation studies: an overview and some suggestions for future research". En: *Target* (ISSN 0924-1884). 7(2), pp. 223-243.

- BAKER, M. (1996). "Corpus-based Translation Studies: the Challenges that Lie Ahead". En: H. Somers (ed.). *Terminology, LSP and Translation Studies in Language Engineering, in Honour of Juan C. Sager*. Ámsterdam; Filadelfia: John Benjamins, pp. 175-186.
- BAKER, M. (2000). "Towards a Methodology for Investigating the Style of a Literary Translator". En: *Target*. 12(2), pp. 241-266.
- BAKER, P. (2006). *Using corpora in discourse analysis*. Londres: Continuum.
- BEDNAREK, M. (2008). *Emotion talk across corpora*. Houndmills; [etc.]: Palgrave Macmillan.
- BEEBY, A.; Rodríguez Inés, P.; Sánchez-Gijón, P. (eds.). (2008). *Corpus use and translating: corpus use for learning to translate and learning corpus use to translate*. Ámsterdam: John Benjamins.
- BEECHING, K. (2003). "Synchronic and diachronic variation [en línea]: the how and why of sociolinguistic corpora" En: D. Archer; et al. (eds.) *Proceedings of the Corpus Linguistics 2003 conference*. Lancaster: Lancaster University. (UCREL Technical Papers; 16). URL: http://ucrel.lancs.ac.uk/publications/CL2003/papers/beeching_abstract.pdf [Consulta: 30 de junio de 2008].
- BERNARDINI, S. (2000a). "Systematising serendipity: proposals for concordancing large corpora with language learners". En: L. Burnard; T. McEnery (eds.). *Rethinking Language Pedagogy from a Corpus Perspective*. Bern: Peter Lang, pp. 225-234.
- BERNARDINI, S. (2000b). *Competence, capacity, corpora*. Bolonia: CLUEB.
- BERNARDINI, S. (2000c). "A 'trainee' translator's perspective on corpora". En: S. Bernardini; F. Zanettin (eds.). *I corpora nella didattica della traduzione: Corpus Use and Learning to Translate*. Bolonia: CLUEB.
- BERNARDINI, S. (2001). "'Spoilt for choice': a learner explores general language corpora". En: G. Aston (ed.). *Learning with corpora*. Bolonia: CLUEB, pp. 220-249.
- BERNARDINI, S. (2002). "Educating translators for the challenges of the new millenium: the potential of parallel bi-directional corpora". En: B. Maia, J. Haller; M. Ulrych (eds.). *Training the language services provider for the new millennium*. Oporto: AstraFlup, pp. 173-186.
- BERNARDINI, S. (2003). "Designing a corpus for translation and language teaching: The CEXI experience". En: *TESOL Quarterly* (Special issue on Corpus Linguistics in TESOL) (ISSN 0039-8322). 37(3), pp. 528-537.
- BERNARDINI, S.; Zanettin, F. (eds.). (2000). *I corpora nella didattica della traduzione: Corpus Use and Learning to Translate*. Bolonia: CLUEB.
- BERNARDINI, S.; Stewart, D.; Zanettin, F. (2003). "Corpora in translator education: an introduction". En: F. Zanettin; S. Bernardini; D. Stewart (eds.). *Corpora in translator education*. Manchester: St. Jerome, pp. 1-13.

- BIBER, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- BIBER, D.; *et al.* (1994). "Corpus-based approaches to issues in applied linguistics". En: *Applied Linguistics* (ISSN 0142-6001). 15(2), pp. 169-189.
- BIBER, D.; Burges, J. (2000). "Historical change in the language use of women and men". En: *Journal of English Linguistics* (Special Issue on Historical Corpora) (ISSN 0075-4242). 28(1), pp. 21-37.
- BOWKER, L. (1996). "Towards a corpus-based approach to terminography". En: *Terminology* (ISSN 0929-9971). 3(1), pp. 27-52.
- BOWKER, L. (1998). "Using specialized monolingual native-language corpora as a translation resource [en línea]: a pilot study". En: *Meta* (ISSN 0026-0452). 43(4), pp. 631-651. También disponible en URL: <http://www.erudit.org/revue/meta/1998/v43/n4/002134ar.html>, y en URL: <http://www.erudit.org/revue/meta/1998/v43/n4/002134ar.pdf> [Consulta: 30 de junio de 2008].
- BOWKER, L. (1999). "Using a corpus to assess student translations: a pilot study". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.). *PALC'99: Practical Applications in Language Corpora. Papers from the International Conference at the University of Łódź, 15-18 April 1999*. Bern: Peter Lang, pp. 529-540.
- BOWKER, L. (2000). "Towards a methodology for exploiting specialized target language corpora as translation resources". En: *International Journal of Corpus Linguistics* (ISSN 1384-6655). 5(1), pp. 17-52.
- BOWKER, L.; Pearson, J. (2002). *Working with specialized language: A practical guide to using corpora*. Londres; Nueva York: Routledge.
- BOWKER, L.; Bennison, P. (2003). "Student translation archive: Design, development and application". En: F. Zanettin; S. Bernardini; D. Stewart (eds.). *Corpora in translator education*. Manchester: St. Jerome, pp. 103-117.
- BRODINE, R. (2001). "Integrating corpus work into an academic reading course". En: G. Aston (ed.). *Learning with corpora*. Bolonia: CLUEB, pp. 138-176.
- BRUNET, E. (1989). "Hyperbase: logiciel documentaire et statistique pour l'exploitation des grands corpus". En: W. McCarty (ed.). *Tools for Humanists, 1989: A Guidebook to the Software and Hardware Fair Held in Conjunction with The Dynamic Text: The 9th International Conference on Computers and the Humanities (ICCH) and the 16th International Association for Literacy and Linguistic Computing (ALLC) Conference*. Toronto: Centre for Computing in the Humanities, pp.33-36.
- BURNARD, L.; McEnery, T. (eds.). (2000). *Rethinking language pedagogy from a corpus perspective*. Bern: Peter Lang.
- CABRÉ CASTELLVÍ, M.T. (1999). "Fuentes de información terminológica para el traductor". En: M. Pinto; J.A. Córdón (eds.). *Técnicas documentales aplicadas a la traducción*. Madrid: Síntesis, pp. 19-40.

- CANDEL MORA, M.A. (2002). *Terminología aplicada a la traducción técnica de informes sobre la gestión de la I+D*. [Tesis doctoral]. Universidad Politécnica de Valencia, Departamento de Filología Inglesa y Alemana. Valencia.
- CHANDLER, B. ; Tribble, C. (1989). *Longman mini-concordancer*. Londres: Longman.
- CHOMSKY, N.A. (1957). *Syntactic structures*. La Haya: Mouton.
- CHOMSKY, N.A. (1962). "A transformational approach to syntax". En: A.A. Hill (ed.). *The third Texas conference on problems of linguistic analysis in English*. Austin: University of Texas Press, pp. 124-158.
- CHRIST, O. (1994). "A modular and flexible architecture for an integrated corpus query system". En: *Proceedings of COMPLEX '94: 3rd Conference on computational lexicography and text research*, Budapest, pp. 23-32.
- CHURCH, K.; et al. (1991). "Using statistics in lexical analysis". En: U. Zernik (ed.). *Lexical acquisition: exploiting online resources to build a lexicon*. Hillsdale: Lawrence Erlbaum Associates, pp. 115-164.
- CIVIT, M. (2003). *Criterios de etiquetación y desambiguación morfosintáctica de corpus en español*. Alicante: Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN). (Colección de Monografías de la Sociedad Española para el Procesamiento del Lenguaje Natural -SEPLN-; 3). También disponible en URL: <http://www.sepln.org/monografiasSEPLN/monografiaCivit.pdf> [Consulta: 30 de junio de 2008].
- CIVIT, M.; Martí, M.A.; Padró, L. (2003). "Using hybrid probabilistic-linguistic knowledge to improve pos-tagging performance" [en línea]. En: *Proceedings of the Corpus Linguistics 2003*. Lancaster: Lancaster University. (UCREL Technical Papers; 16). URL: <http://ucrel.lancs.ac.uk/publications/CL2003/papers/torruella.pdf> [Consulta: 3 de abril de 2007].
- CORPAS PASTOR, G. (2001). "Compilación de un corpus *ad hoc* para la enseñanza de la traducción inversa especializada" [en línea]. En: *Trans* (ISSN 1137-2311). 5, pp. 155-184. También disponible en URL: http://www.trans.uma.es/Trans_5/t5_155-184_GCorpas.pdf [Consulta: 30 de junio de 2008].
- CORPAS PASTOR, G. (2002). "Traducir con corpus: de la teoría a la práctica". En: J. García Palacios; M.T. Fuentes (eds.). *Texto, terminología y traducción*. Salamanca: Almar, pp. 189-226.
- CORPAS PASTOR, G.; Postigo Pinazo, M.E. (2002). "Aplicaciones del corpus para la redacción en inglés de textos científicos originales o traducidos: a propósito de la sigla *ACTH*". En: M.D. Fernández de la Torre Madueño (ed.). *Estudio del léxico: análisis y docencia*. Málaga: Servicio de Publicaciones de la Universidad de Málaga, pp. 39-71.
- CORPAS PASTOR, G.; Seghiri Domínguez, M. (2006). *El concepto de representatividad en la lingüística del corpus: aproximaciones teóricas y metodológicas*. Documento técnico BFF2003-04616 MCYT/TI-DT-2006-1.

- CORPAS PASTOR, G.; Seghiri Domínguez, M. (2008). "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: corpus use for learning to translate and learning corpus use to translate*. Ámsterdam: John Benjamins.
- CRUDEN, A. (1769). *Complete concordance to the Old and New Testaments*. 3rd ed. Londres: Lutterworth Press.
- CURADO FUENTES, A.; García Berzosa, M.J. (2000). "Dealing with small representative corpora in ESP". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.). *PALC'99: Practical Applications in Language Corpora: papers from the international conference at the University of Lodz, 15-18 April 1999*. Frankfurt am Main; Nueva York: Peter Lang. (Lodz studies in language series; 1), pp. 475-488.
- CURZAN, A. (2000). "English historical corpora in the classroom". En: *Journal of English Linguistics* (Special Issue on Historical Corpora) (ISSN 0075-4242). 28(1), pp. 77-89.
- EATON, H. (1940). *Semantic frequency list for English, French, German and Spanish*. Chicago: Chicago University Press.
- ELLIS, A.J. (1889). *On early English pronunciation. Part V: The existing phonology of English dialects compared with that of West Saxon*. Nueva York: Greenwood Press.
- EUROPEAN LANGUAGE RESOURCES ASSOCIATION (ELRA) (1997). *Multilingual Corpora for Co-operation-MLCC* [recurso digital]. Ver. 1.0. París. 2 CD-ROM. Programa informático.
- ESTOUP, J.B. (1916). *Gammes stenographiques*. París: Institut Stenographique de France.
- FAUTH, J.; Schmid, H. (2003). "Detecting gender-preferential patterns of linguistic features in face-to-face communication" [en línea]. En: D. Archer; et al. (eds.). *Proceedings of the Corpus Linguistics 2003 conference*. Lancaster: Lancaster University. (UCREL Technical Papers; 16). URL: http://ucrel.lancs.ac.uk/publications/CL2003/papers/fauth_abstract.pdf [Consulta: 30 de junio de 2008].
- FERNÁNDEZ SÁNCHEZ, F. (2005). *El folleto de cursos de idiomas para extranjeros: análisis contrastivo (alemán-español) por tipos de emisor y subtextos*. [Tesis doctoral]. Universitat Pompeu Fabra, Departament de Traducció i Filologia. Barcelona.
- FILLMORE, C.J. (1992). "'Corpus linguistics' or 'Computer-aided armchair linguistics'". En: J. Svartvik (ed.). *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4-8 August 1991*. Berlín: Mouton de Gruyter, pp. 35-60.
- FLIGELSTONE, S. (1993). "Some reflections on the question of teaching, from a corpus linguistics perspective". En: *ICAME Journal: Computers in English Linguistics* (ISSN 0801-5775). 17, pp. 97-109.

- FLOWERDEW, J. (1996). "Concordancing in language teaching". En: M.C. Pennington (ed.). *The Power of Call*. Houston: Athelstan, pp. 97-113.
- FOUCOU, P.; Kübler, N. (2000). "A web-based environment for teaching technical English". En: L. Burnard; T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective*. Bern: Peter Lang, pp. 65-73.
- FRANKENBERG-GARCIA, A. (2001). "COMPARA [en línea]: the Portuguese-English parallel corpus". En: *Actas do I Congresso Internacional de Estudos Anglo-Portugueses, Universidade Nova de Lisboa, 6-8 de Mayo de 2001*. URL: <http://www.fcsh.unl.pt/congressocean/actas-proceedings.htm> [Consulta: 29 de junio de 2008].
- FRASER, J. (1999). "The translator and the word: the pros and cons of dictionaries in translation". En: G. Anderman; M. Rogers (eds.) *Word, text, translation: Liber Amicorum for Peter Newmark*. Clevedon: Multilingual Matters, pp. 25-34.
- FRIEDBICHLER, I.; Friedbichler, M. (2000). "The potential of domain-specific target-language corpora for the translator's workbench". En: S. Bernardini; F. Zanettin (eds.). *I corpora nella didattica della traduzione: Corpus Use and Learning to Translate*. Bolonia: CLUEB, pp. 107-116.
- FRIES, C.C. (1952). *The structure of English: an introduction to the construction of sentences*. Nueva York: Harcourt-Brace.
- FRIES, C.C.; Traver, A.A. (1940). *English word lists*. Washington D.C.: American Council of Education.
- GARSIDE, R. (1987). "The CLAWS word-tagging system". En: R. Garside; G. Leech; G. Sampson (eds.). *The computational analysis of English: a corpus-based approach*. Londres: Longman, pp. 30-41.
- GAVIOLI, L. (1997). "Exploring texts through the concordancer: guiding the learner". En: A. Wichmann; *et al.* (eds.). *Teaching and Language Corpora*. Londres; Nueva York: Longman, pp. 83-99.
- GAVIOLI, L. (2001). "The learner as researcher: introducing corpus concordancing in the classroom". En: G. Aston (ed.). *Learning with corpora*. Bolonia: CLUEB, pp. 108-137.
- GAVIOLI, L.; Zanettin, F. (2000). "Comparable corpora and translation: a pedagogic perspective". En: S. Bernardini; F. Zanettin (eds.) *I corpora nella didattica della traduzione: Corpus Use and Learning to Translate*. Bolonia: CLUEB.
- GOUGENHEIM, G.; *et al.* (1956). *L'Elaboration du français élémentaire*. París: Didier.
- GRANGER, S. (1999). "Use of tenses by advanced EFL learners: evidence from an error-tagged computer corpus". En: H. Hasselgård; S. Oksefjell (eds.). *Out of corpora: Studies in honour of Stig Johansson*. Ámsterdam: Rodopi, pp. 191-202.
- GREENE, B.B.; Rubin, G.M. (1971). *Automated grammatical tagging of English*. Providence: Brown University.

- HOEY, M. (2004). "Lexical priming and the properties of text". En: A. Partington; J. Morley; L. Haarman (eds.). *Corpora and discourse*. Bern: Peter Lang, pp. 385-412. También disponible en: [URL: http://www.monabaker.com/tsresources/LexicalPrimingandthePropertiesofText.htm](http://www.monabaker.com/tsresources/LexicalPrimingandthePropertiesofText.htm) [Consulta: 30 de junio de 2008].
- HOEY, M.; *et al.* (2007). *Text, discourse and corpora*. Londres: Continuum. (Studies in corpus and discourse).
- HONEYFIELD, J. (1989). "A typology of exercises based on computer-generated concordance material". En: *Guidelines: a periodical for classroom language teachers* (ISSN 0129-7767). 2(1), pp. 42-50.
- HUGHES, R. (1996). *English in speech and writing: investigating language and literature*. Londres; Nueva York: Routledge.
- HULTSIJN, J.H. (1992). "Retention of inferred and given word meanings: experiments in incidental vocabulary learning". En: P.J.L. Arnaud; H. Béjoint (eds.). *Vocabulary and applied linguistics*. Londres: Macmillan, pp. 113-125.
- JABBOUR, G. (1997). "Where is the writer in a frequency list?: Using a corpus of medical research articles in teaching". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.). *PALC'97: Practical Applications in Language Corpora*. Lodz: Lodz University Press, pp. 178-188.
- JACOB, I.; *et al.* (2003). "Automatic feeding of translation memory tools" [en línea]. En: D. Archer; *et al.* (eds.). *Proceedings of the Corpus Linguistics 2003 conference*. Lancaster: Lancaster University. (UCREL Technical Papers; 16). URL: <http://ucrel.lancs.ac.uk/publications/CL2003/papers/jacob.pdf> [Consulta: 30 de junio de 2008].
- JOHANSSON, S.; Hofland, K. (1994). "Towards an English-Norwegian parallel corpus". En: U. Fries; *et al.* (eds.). *Creating and using English language corpora*. Ámsterdam; Atlanta: Rodopi, pp. 25-37.
- JOHNS, T. (1991a). "From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning". En: *ELR Journal: classroom concordancing* (ISSN 0265-847X). 4, pp. 27-45.
- JOHNS, T. (1991b). "Should you be persuaded – Two samples of data-driven learning materials". En: *ELR Journal: English Language Research* (Special issue on classroom concordancing) (ISSN 0265-847X). 4, pp. 1-16.
- KÄDING, J. (1897). *Häufigkeitwörterbuch der deutschen Sprache*. Steglitz: Eigenverlag.
- KENNEDY, G. (1992). "Preferred ways of putting things with implications for language teaching". En: J. Svartvik (ed.). *Directions in corpus linguistics: proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlín; Nueva York: Mouton de Gruyter. (Trends in linguistics. Studies and monographs; 65), pp. 335-373.
- KENNEDY, G. (1998). *An introduction to corpus linguistics*. Londres; Nueva York: Longman.

- KETTEMANN, B. (1997). "Concordancing as input enhancement in ELT". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.). *PALC'97: Practical Applications in Language Corpora*. Lodz: Lodz University Press, pp. 63-73.
- KING, P. (1997). "Parallel corpora for translator training". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.) *PALC'97: Practical Applications in Language Corpora*. Lodz: Lodz University Press, pp. 393-402.
- KJELLMER, G. (1986). "'The lesser man': observations on the role of women in modern English writings". En: J. Aarts; W. Meijs (eds.). *Corpus Linguistics II*. Amsterdam: Rodopi, pp. 163-76.
- KNOWLES, G.; Taylor, L.; Williams, B. (1992). *A corpus of formal British English speech*. Londres: Longman.
- KÜBLER, N. (2003). "Corpora and LSP translation". En: F. Zanettin; S. Bernardini; D. Stewart (eds.). *Corpora in translator education*. Manchester: St. Jerome, pp. 25-42.
- LAVIOSA, S. (2002). *Corpus-based translation studies: theory, findings, applications*. Amsterdam; Nueva York: Rodopi. (Approaches to Translation Studies; 17).
- LEECH, G. (1992). "Corpora and theories of linguistic performance". En: J. Svartvik (ed.). *Directions in corpus linguistics: Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlín: Mouton de Gruyter. (Trends in linguistics. Studies and monographs; 65), pp. 105-122
- LEECH, G. (1997). "Teaching and Language Corpora: a convergence". En: A. Wichmann; et al. (eds.) *Teaching and Language Corpora*. Londres; Nueva York: Longman, pp. 1-23.
- LENKO-SZYMANSKA, A. (1997). "Corpora as a resource for a non-native teacher of writing". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.). *PALC'97: Practical Applications in Language Corpora*. Lodz: Lodz University Press, pp. 74-89.
- LEWANDOWSKA-TOMASZCZYK, B.; Melia, P.J. (eds.). (1997). *PALC'97: Practical Applications in Language Corpora*. Lodz: Lodz University Press.
- LEWANDOWSKA-TOMASZCZYK, B.; Melia, P.J. (eds.). (1999). *PALC'99: Practical Applications in Language Corpora. Papers from the International Conference at the University of Łódź, 15-18 April 1999*. Bern: Peter Lang.
- LINDQUIST, H. (1984). "The use of corpus-based studies in the preparation of handbooks for translators". En: W. Wilss; G. Thome (eds.). *Translation Die Theorie des Übersetzens und ihr Aufschlußwert für die Übersetzungs- und Dolmetschdidaktik (Theory and its Implementation in the Teaching of Translating and Interpreting), Akten des Internationalen Kolloquiums der Association Internationale de Linguistique Appliquée (AILA), Saarbrücken, 25-30 Juli, 1983*. Tübingen: Gunter Narr Verlag, pp. 260-270.
- LÓPEZ RODRÍGUEZ, C.I. (2000). *Tipología textual y cohesión en la traducción biomédica inglés-español: un estudio de corpus*. [Tesis doctoral]. Universidad de Granada, Departamento de Traducción e interpretación. Granada.

- LÓPEZ RODRÍGUEZ, C.I. (2002). "Training translators to learn from news report corpora: the case of Anglo-American cultural references". En: B. Maia; J. Haller; M. Ulrych (eds.). *Training the Language Services Provider for the New Millennium*. Oporto: AstraFlup, pp. 213-222.
- MACWHINNEY, B.; Snow, C. (1990). "The Child Language Data Exchange System". En: *ICAME Journal*, 14, pp. 3-25.
- MAGALHÃES, C. (2001). "Pesquisas textuais/discursivas em tradução: o uso da corpora". En: A. Pagano (org.). *Metodologias de pesquisa em tradução*. Belo Horizonte: FALE-UFMG. (Série Estudos Linguísticos, 3), pp. 93-116.
- MAIA, B. (1997). "Do-it-yourself corpora... with a little bit of help from your friends". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.). *PALC'97: Practical Applications in Language Corpora*. Lodz: Lodz University Press, pp. 403-410.
- MAIA, B. (2000). "Making corpora – a learning process". En: S. Bernardini; F. Zanettin (eds.). *I corpora nella didattica della traduzione: Corpus Use and Learning to Translate*. Bolonia: CLUEB, pp. 47-59.
- MAIA, B.; Haller, J.; Ulrych, M. (eds.). (2002). *Training the language services provider for the new millennium*. Oporto: AstraFlup.
- MAIA, B. (2003). "Some languages are more equal than others: Training translators in terminology and information retrieval using comparable and parallel corpora". En: F. Zanettin; S. Bernardini; D. Stewart (eds.). *Corpora in translator education*. Manchester: St. Jerome, pp. 43-70.
- MÁRQUEZ, L.; Padró, L. (1997). "A flexible POS tagger using an automatically acquired language model". En: *Proceedings of the 35th annual meeting of the Association for Computational Linguistics: Joint EACL/ACL*. Madrid., pp. 238-245.
- MÁRQUEZ ROJAS, M.J. (2005). *El anglicismo terminológico integral en los textos especializados: pautas para su tratamiento automatizado*. [Tesis doctoral]. Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada. Barcelona.
- MARTÍNEZ SÁNCHEZ, P.C. (1999). *La traducción de los nombres de "realia" en textos del sector turístico*. [Tesis doctoral]. Universidad de Valencia, Departamento de Filología Inglesa y Alemana. Valencia.
- MAURANEN, A.; Kujamäki, P. (2004). *Translation universals: Do they exist?* Amsterdam; Nueva York: John Benjamins.
- McENERY, T.; Wilson, A. (1996). *Corpus linguistics*. Edimburgo: Edinburgh University Press.
- MONZÓ, E. (2003). "Corpus-based teaching [en línea]: the use of original and translated texts in the training of legal translators". En: *Accurapid* (ISSN 1536-7207). 7(4). URL: <http://accurapid.com/journal/26edu.htm> [Consulta: 1 de abril de 2007].
- MURISON-BOWIE, S. (1993). *Micro-Concord manual: an introduction to the practices and principles of concordancing in language teaching*. Oxford: Oxford University Press.

- NARITA, M.; Kurokawa, K.; Utsuro, T. (2002). "A web-based English abstract writing tool using a tagged E-J parallel corpus" [recurso digital]. En: *Proceedings of the Third International Conference on Language Resources and Evaluation*, pp. 2115-2119.
- NEVALAINEN, T. (2000). "Gender differences in the evolution of standard English". En: *Journal of English Linguistics* (Special Issue on Historical Corpora) (ISSN 0075-4242). 28(1), pp. 38-59.
- OAKES, M.P. (1998). *Statistics for corpus linguistics*. Edimburgo: Edinburgh University Press. (Edinburgh textbooks in empirical linguistics).
- OLIVER DEL OLMO, S. (2004). *Análisis contrastivo español/inglés de la atenuación retórica en el discurso médico: el artículo de investigación y el caso clínico*. [Tesis doctoral]. Universitat Pompeu Fabra, Departament de Traducció i Filologia. Barcelona.
- OLOHAN, M. (2004). *Introducing corpora in translation studies*. Londres; Nueva York: Routledge.
- O'NEILL, E.T.; et al. (2003). "Trends in the evolution of the public Web [en línea]: 1998-2002". En: *D-Lib Magazine* (ISSN 1082-9873). 9(4). URL: <http://www.dlib.org/dlib/aprilo3/lavoie/o4lavoie.html> [Consulta: 31 de marzo de 2007].
- OSBORNE, J. (2000). "What can students learn from a corpus?: building bridges between data and explanation". En: L. Burnard; T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective*. Bern: Peter Lang, pp. 165-172.
- PAGANO, A. (org.). (2001a). *Metodologias de pesquisa em tradução*. Belo Horizonte: FALE-UFMG. (Série Estudos Linguísticos; 3).
- PAGANO, A. (2001b). "As pesquisas historiográficas em tradução". En: A. Pagano (org.). *Metodologias de pesquisa em tradução*. Belo Horizonte: FALE-UFMG. (Série Estudos Linguísticos; 3), pp. 93-116.
- PALUMBO, G. (2002). "The use of phraseology for training and research in the translation of LSP texts". En: B. Maia; J. Haller; M. Ulrych (eds.). *Training the Language Services Provider for the New Millennium*. Oporto: AstraFlup, pp. 199-211.
- PARODI, G. (ed.). (2007). *Working with Spanish corpora*. Londres; Nueva York: Continuum.
- PARTINGTON, A. (1998). *Patterns and meanings: using corpora for English language research and teaching*. Ámsterdam; Filadelfia: John Benjamins.
- PARTINGTON, A. (2001). "Corpus-based description in teaching and learning". En: G. Aston (ed.). *Learning with corpora*. Bolonia: CLUEB, pp. 63-84.
- PATTON, M.Q. (1987). *How to use qualitative methods in evaluation*. Newbury Park: Sage.
- PEARSON, J. (1996). "Electronic texts and concordances in the translation classroom". En: *Teanga* (Dublín). 16, pp. 85-95.
- PEARSON, J. (1998). *Terms in context*. Ámsterdam; Filadelfia: John Benjamins.

- PEARSON, J. (1999). "Using specialized comparable corpora to evaluate student translations". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.). *PALC'99: Practical Applications in Language Corpora. Papers from the International Conference at the University of Łódź, 15-18 abril 1999*. Bern: Peter Lang, pp. 541-552.
- PEARSON, J. (2000). "Surfing the Internet: teaching students to choose their texts wisely". En: L. Burnard; T. McEnery (eds.). *Rethinking language pedagogy from a corpus perspective: papers from the third international conference on Teaching And Language Corpora*. Hamburgo: Peter Lang, pp. 235-239.
- PEARSON, J. (2003). "Using parallel texts in the translator training environment". En: F. Zanettin; S. Bernardini; D. Stewart (eds.). *Corpora in translator education*. Manchester: St. Jerome, pp. 15-24.
- PIENEMANN, M. (1992). "COALA – a computational system for interlanguage analysis". En: *Second Language Research* (ISSN 0267-6583). 8(1), pp. 59-92.
- PINTO, M.; Cerdón, J.M. (eds.). (1999). *Técnicas documentales aplicadas a la traducción*. Madrid: Editorial Síntesis; Valladolid: Publicaciones de la Universidad de Valladolid.
- QUINN, A. (1993). "An object-oriented design for a corpus utility program". En: J. Aarts, P. de Haan; N. Oostdijk (eds.). *English language corpora: design, analysis and exploitation*. Amsterdam: Rodopi, pp. 215-225.
- QUIRK, R. (1968). *Essays on the English Language: medieval and modern*. Londres: Longman.
- QUIRK, R.; et al. (1985). *A comprehensive grammar of the English language*. Londres; Nueva York: Longman.
- REAL ACADEMIA ESPAÑOLA. *Banco de datos (CORDE)* [base de datos en línea]: *Corpus diacrónico del español*. URL: <http://www.rae.es> [Consulta: 7 de enero de 2008].
- REAL ACADEMIA ESPAÑOLA. *Banco de datos (CREA)* [base de datos en línea]: *Corpus de referencia del español actual*. URL: <http://www.rae.es> [Consulta: 7 de enero de 2008].
- RIBAS BRUGUER, M. (2006). *Alineació de textos jurídics paral·lels (català-castellà): alguns problemes*. [Tesis doctoral]. Universitat Pompeu Fabra, Institut Universitari de Lingüística Aplicada. Barcelona.
- RODRÍGUEZ INÉS, P. (1999). *Investigating ideological issues using the Translationsl English Corpus*. [Trabajo de fin de máster / MSc. Dissertation]. University of Manchester Institute of Science and Technology, Department of Language Engineering. Manchester.
- RODRÍGUEZ INÉS, P. (2000). *Application of corpus methodology and techniques to the study of ideology in translation*. [Trabajo de investigación]. Universitat Autònoma de Barcelona, Departament de Traducció i d'Interpretació. Bellaterra.

- RODRÍGUEZ INÉS, P. (2008). "Evaluating the process and not just the product when using corpora in translator education". En: A. Beeby; P. Rodríguez Inés; P. Sánchez-Gijón (eds.). *Corpus use and translating: corpus use for learning to translate and learning corpus use to translate*. Ámsterdam: John Benjamins.
- RODRÍGUEZ MEDINA, M.J. (1999). *Los anglicismos de frecuencia sintácticos en los manuales de informática traducidos*. [Tesis doctoral]. Universidad de Las Palmas de Gran Canaria, Departamento de Filología Moderna. Las Palmas.
- RUNDELL, M. (1997). "Understatement and indirectness in English: from corpus evidence to classroom practice". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.) *PALC'97: Practical Applications in Language Corpora*. Lodz: Lodz University Press, pp. 90-98.
- SALDANHA, G. (2003). "Studying Gender-Related Linguistic Features in Translated Language". En: J. Santaemilia (ed.). *Género, lenguaje y traducción*. Valencia: Universitat de València, pp. 420-432.
- SALDANHA, G. (2008a). *Using Corpora in Translation Research*. Londres; Nueva York: Continuum. (en prensa).
- SALDANHA, G. (2008b). "Style of translation: the use of source language words in translations by Margaret Jull Costa and Peter Bush". En: A. Kruger (ed.). *Corpus Based Translation Studies*. Manchester: St. Jerome. (en prensa).
- SÁNCHEZ-GIJÓN, P. (2003). *Els documents digitals especialitzats: utilització de la lingüística de corpus com a font de recursos per a la traducció especialitzada*. [Tesis doctoral]. Universitat Autònoma de Barcelona, Departament de Traducció i d'Interpretació. Bellaterra.
- SCOTT, M. (1996). *WordSmith*. Oxford: Oxford University Press.
- SCOTT, M. (2008). *Introduction to WordSmith Tools*. URL: <http://lexically.net/downloads/version5/HTML/index.html> [Consulta: 1 de julio de 2008].
- SEGHIRI DOMÍNGUEZ, M. (2006). *Compilación de un corpus trilingüe de seguros turísticos (español-inglés-italiano)* [recurso electrónico]: aspectos de evaluación, catalogación, diseño y representatividad. [Tesis doctoral]. Universidad de Málaga, Departamento de Traducción e Interpretación. Málaga. URL: <http://www.sci.uma.es/bbldoc/tesisuma/16754888.pdf> [Consulta: 1 de julio de 2008].
- SEMINO, E.; Short, M. (2004). *Corpus stylistics: speech, writing and thought presentation in a corpus of English writing*. Londres: Routledge.
- SOCIEDAD GENERAL DE AUTORES Y EDITORES (SGAE). *Derechos*. Madrid: la Sociedad, 26 de junio 2002. URL: http://www.sgae.es/tipology/faq/list/es/1150_238.html [Consulta: 17 de julio de 2007].
- SINCLAIR, J.M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

- SINCLAIR, J.M. (1995). "Corpus Typology - A framework for classification". En: G. Melckens; B. Warren (eds.). *Studies in Anglistics*. Estocolmo: Almqvist & Wiksell International, pp. 17-33.
- SINCLAIR, J.M. (1996). *Preliminary Recommendations on Corpus Typology*. EAGLES Document EAG-TCWG-CTYP/P. URL: <http://www.ilc.cnr.it/EAGLES96/corpuSTYP/corpuSTYP.html> [Consulta: 1 de julio de 2008].
- SINCLAIR, J.M. (2004). *Trust the Text: Language, Corpus and Discourse*. Londres; Nueva York: Routledge.
- SINCLAIR, J.M.; Renouf, A. (1988). "A lexical syllabus for language learning". En: R. Carter; M. McCarthy (eds.). *Vocabulary and language teaching*. Harlow: Longman, pp. 140-158.
- SKEHAN, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- SNELL-HORNBY, M. (1988). *Translation Studies: an integrated approach*. Ámsterdam; Filadelfia: John Benjamins.
- STAR SERVICIOS LINGÜÍSTICOS S.L. (2003). "Editorial" [en línea]. En: *Starlette*. 6. URL: http://www.star-spain.com/es/actualidad/numero_6/STARlette_6.html [Consulta: 30 de junio de 2008].
- STENSTRÖM, A.B. (1987). "Carry-on signals in English conversation". En: W. Meijs (ed.). *Corpus linguistics and beyond*. Ámsterdam: Rodopi, pp. 87-119.
- STUBBS, M. (1996). *Text and corpus analysis: computer assisted studies of language and institutions*. Oxford; Malden: Blackwell. (Language in Society).
- STUBBS, M. (2001). *Words and phrases: corpus studies of lexical semantics*. Oxford; Malden: Blackwell. (Language in Society).
- SVARTVIK, J. (ed.). (1992). *Directions in corpus linguistics: proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991*. Berlín: Mouton de Gruyter. (Trends in linguistics. Studies and monographs; 65).
- TDR: tesis doctorales en red [en línea]. URL: <http://www.tdr.cesca.es/> [Consulta: 29 de junio de 2008].
- TESEO [base de datos en línea]. [Madrid: Ministerio de Ciencia e Innovación]. URL: <https://www.micinn.es/teseo/login.jsp> [Consulta: 29 de junio de 2008].
- TEUBERT, W. (1996). "Comparable or parallel corpora?". En: *International Journal of Lexicography* (ISSN 0950-3846). 9(3), pp. 238-264.
- THORNDIKE, E.L. (1921). *Teacher's wordbook*. Nueva York: Columbia Teachers College.
- TOGNINI-BONELLI, E. (2001). *Corpus linguistics at work*. Ámsterdam; Filadelfia: John Benjamins.

- TORRUELLA, J.; Llisterri, J. (1999). "Diseño de corpus textuales y orales". En: J.M. Blecua; *et al.* (eds.) *Filología e informática: nuevas tecnologías en los estudios filológicos*. Barcelona: Milenio; Bellaterra: Universitat Autònoma de Barcelona, pp. 45-77.
- TOTTIE, G. (1991). *Negation in English speech and writing: a study in variation*. San Diego: Academic Press.
- TRIBBLE, C. (1997). "Improvising corpora for ELT: quick-and-dirty ways of developing corpora for language teaching". En: B. Lewandowska-Tomaszczyk; P.J. Melia (eds.). *PALC'97: Practical Applications in Language Corpora*. Lodz: Lodz University Press, pp. 106-117.
- VARANTOLA, K. (2000). "Translators, dictionaries and text corpora". En: S. Bernardini; F. Zanettin (eds.). *I corpora nella didattica della traduzione: Corpus Use and Learning to Translate*. Bologna: CLUEB, pp. 117-133.
- VARANTOLA, K. (2003). "Translators and disposable corpora". En: S. Bernardini; D. Stewart; F. Zanettin (eds.). *Corpora in translator education*. Manchester: St. Jerome, pp. 55-70.
- VARGAS SIERRA, M.C. (2004). *Aproximación terminográfica al lenguaje de la piedra natural: propuesta de sistematización para la elaboración de un diccionario traductológico*. [Tesis doctoral]. Universidad de Alicante, Departamento de Filología Inglesa. Alicante.
- WEST, M. (ed.). (1953). *A general service list of English words*. Londres: Longman.
- WICHMANN, A.; *et al.* (eds.) (1997). *Teaching and Language Corpora*. Londres; Nueva York: Longman.
- WIDDOWSON, H.G. (1983). *Learning purpose and language use*. Oxford: Oxford University Press.
- WIDDOWSON, H.G. (1990). *Aspects of language teaching*. Oxford: Oxford University Press.
- WILKINSON, M. (2005a). "Using a specialized corpus to improve translation quality" [en línea]. En: *Accurapid* (ISSN 1536-7207). 9(3). URL: <http://accurapid.com/journal/33corpus.htm> [Consulta: 31 de marzo de 2007].
- WILKINSON, M. (2005b). "Discovering translation equivalents in a tourism corpus" [en línea]. En: *Accurapid* (ISSN 1536-7207). 9(4). URL: <http://accurapid.com/journal/34corpus.htm> [Consulta: 31 de marzo de 2007].
- WILKINSON, M. (2006a). "Compiling corpora for use as translation resources" [en línea]. En: *Accurapid* (ISSN 1536-7207). 10(1). URL: <http://accurapid.com/journal/35corpus.htm> [Consulta: 31 de marzo de 2007].
- WILKINSON, M. (2006b). "Legal aspects of compiling corpora to be used as translation resources" [en línea]. En: *Accurapid* (ISSN 1536-7207). 10(2). URL: <http://accurapid.com/journal/36corpus.htm> [Consulta: 31 de marzo de 2007].

- WILLIAMS WHITE, I.A. (2001). *Aspectos de la estructura de tema y rema en la traducción de los artículos biomédicos del inglés al español: estudio contrastivo basado en la lingüística de corpus*. [Tesis doctoral]. Universidad de León, Departamento de Filología Moderna. León.
- WILLIS, J.D. (1990). *The lexical syllabus*. Londres: Collins.
- WILLIS, J.D.; Willis, J.R. (1989). *Collins Cobuild English course*. Londres: Collins.
- WRIGHT, J. (1898-1905). *The English dialect dictionary*. Oxford: Oxford University Press.
- YANG, H. (1985). "The JDEST computer corpus of texts in English for science and technology". En: *ICAME News* (ISSN 0800-6806). 9, pp. 24-5.
- ZANETTIN, F. (1994). "Parallel words: designing a bilingual database for translation activities". En: A. Wilson; T. McEnery (eds.). *Corpora in language education and research: a selection of papers from TALC'94*. Lancaster: Lancaster University, pp. 99-111. (UCREL Technical Papers; 4).
- ZANETTIN, F. (1998). "Bilingual comparable corpora and the training of translators" [en línea]. En: *Meta* (ISSN 0026-0452). 43(4), pp. 616-630. También disponible en <http://www.erudit.org/revue/meta/1998/v43/n4/004638ar.pdf> , y en [URL: http://www.erudit.org/revue/meta/1998/v43/n4/004638ar.html](http://www.erudit.org/revue/meta/1998/v43/n4/004638ar.html) [Consulta: 30 de junio de 2008].
- ZANETTIN, F. (2001). "Swimming in words: corpora, translation, and language learning". En: G. Aston (ed.). *Learning with corpora*. Bolonia: CLUEB, pp. 177-197.
- ZANETTIN, F. (2002). "DIY corpora: the WWW and the translator" [en línea]. En: B. Maia; J. Haller; M. Ulrych (eds.). *Training the language services provider for the new millennium*. Oporto: AstraFlup, pp. 239-248. También disponible en [URL: http://www.federicozanettin.net/DIYcorpora.htm](http://www.federicozanettin.net/DIYcorpora.htm) [Consulta: 30 de junio de 2008].
- ZANETTIN, F.; Bernardini, S.; Stewart, D. (eds.). (2003). *Corpora in translator education*. Manchester: St. Jerome.
- ZHU, Q. (1989). "A quantitative look at the Guangzhou Petroleum English Corpus". En: *ICAME Journal* (ISSN 0801-5775). 13, pp. 28-38.
- ZORZI, D. (2001). "The pedagogic use of spoken corpora: learning discourse markers in Italian". En: G. Aston (ed.). *Learning with corpora*. Bolonia: CLUEB, pp. 85-107.

TRADUCCIÓN, ESTUDIOS SOBRE LA TRADUCCIÓN, DIDÁCTICA DE LENGUAS Y TRADUCCIÓN Y COMPETENCIA TRADUCTORA

- ABRECHT, R. (1991). *L'évaluation formative, une analyse critique*. Bruselas: De Boeck.
- AGENCIA NACIONAL DE EVALUACIÓN DE LA CALIDAD Y ACREDITACIÓN (ANECA). (2004). *Libro blanco* [en línea]: *título de grado en Traducción e Interpretación*. Coordinador E. Muñoz Raya. URL: http://www.aneca.es/activin/docs/libroblanco_traduc_def.pdf [Consulta: 30 de junio de 2008].
- AGRUPACIÓN DE CENTROS ESPECIALIZADOS EN TRADUCCIÓN –ACT– (2005). *Estudio de situación del mercado español de servicios profesionales de traducción*. [Documento protegido adquirido por el proyecto TRACE, Grupo Tradumática].
- BEEBY, A. (1996a). *Teaching translation from Spanish to English*. Ottawa: University of Ottawa Press.
- BEEBY, A. (1996b). “La traducción inversa”. En: A. Hurtado (ed.), *La enseñanza de la traducción*. Castellón de la Plana: Universitat Jaume I. (Col·lecció Estudis sobre la traducció), pp. 57-78.
- BEEBY, A. (2003). “Genre literacy and contrastive rhetoric in teaching inverse translation”. En: D. Kelly; et al. (eds.) *La direccionalidad en traducción e interpretación: perspectivas teóricas, profesionales y didácticas*. Granada: Atrio, pp. 155-166.
- BELL, R.T. (1991). *Translation and translating*. Londres: Longman.
- BIBLIOGRAFÍA DE INTERPRETACIÓN Y TRADUCCIÓN (BITRA) [base de datos en línea]. Alicante: Departamento de Traducción e Interpretación. Universidad de Alicante. URL: http://cv1.cpd.ua.es/tra_int/usu/buscar.asp?idioma=es [Consulta: 16 de octubre de 2007].
- BREHM CRIPPS, J. (2004). *Targeting the source text: a coursebook in English for translator trainees*. Castelló de la Plana: Publicacions de la Universitat Jaume I.
- CÁMARA AGUILERA, E. (2003). “Traducción especializada y direccionalidad: la traducción inversa (español-inglés) como potencial mercado en los ámbitos de la ciencia y la tecnología”. En: D. Kelly; et al. (eds.). *La direccionalidad en traducción e interpretación: perspectivas teóricas, profesionales y didácticas*. Granada: Atrio, pp. 208-223.
- CAMPBELL, S. (1998). *Translation into the second language*. Londres: Longman.
- CHESTERMAN, A. (1997). *Memes of Translation*. Ámsterdam: John Benjamins.
- DELISLE, J. (1980/1988). *L'analyse du discours comme méthode de traduction*. Université d'Ottawa. (Cahiers de Traductologie; 2). (Traducción: An interpretative approach. University of Ottawa Press, 1988).

- DELISLE, J. (1993). *La traduction raisonnée*. Canadá: Les Presses de l'Université d'Ottawa.
- DOLLERUP, C.; Loddegaard, A. (eds.). (1992). *Teaching translation and interpreting: training talent and experience: Papers from the first language international conference, Elsinore, Denmark, 1991*. Ámsterdam: John Benjamins.
- DUFF, A. (1989). *Translation*. Oxford: Oxford University Press.
- DURIEUX, Ch. (1988). *Fondement didactique de la traduction technique*. París: Didier.
- FERNÁNDEZ, A.; et al. (2002). "El perfil profesional de la psicopedagogía e implicaciones para el futuro". En: Z. Martínez de la Hidalga (coord.) *Definición y perspectivas profesionales de la psicopedagogía: actas de las I Jornadas de Psicopedagogía de la Universidad de Deusto, Bilbao, 18 y 19 de abril de 2002*. Bilbao: Universidad de Deusto, pp. 49-64.
- GAMERO PÉREZ, S. (1996). "La enseñanza de la traducción científico-técnica". En: A. Hurtado (ed.). *La enseñanza de la traducción*. Castellón de la Plana: Publicaciones de la Universitat Jaume I, pp. 241-251.
- GARCÍA ÁLVAREZ, A.M. (2006). "Aciertos y desaciertos en la didáctica de la traducción". *Vasos Comunicantes* (ISSN 1135-7037). 35, pp. 69-76.
- GONÇALVES, J.L. (2003). *O desenvolvimento da competência do tradutor: investigando o processo através de um estudo exploratório-experimental*. [Tesis doctoral]. Universidade Federal de Minas Gerais (UFMG), Faculdade de Letras. Belo Horizonte.
- GONÇALVES, J.L. (2005). "O desenvolvimento da competência do tradutor: em busca de parâmetros cognitivos". En: A. Pagano; C. Magalhães; F. Alves (orgs.). *Competência em tradução: cognição e discurso*. Belo Horizonte: Editora UFMG. (Coleção Humanitas), pp. 59-90.
- GONZÁLEZ DAVIES, M. (coord.). (2003). *Secuencias: tareas para el aprendizaje interactivo de la traducción especializada*. Barcelona: Octaedro-EUB.
- GONZÁLEZ DAVIES, M. (2004). *Multiple voices in the translation classroom*. Ámsterdam: John Benjamins.
- GONZÁLEZ, J.; Wagenaar, R. (eds.). (2003). *Tuning educational structures in Europe: Informe final: fase uno*. Bilbao: Universidad de Deusto. URL: http://www.relint.deusto.es/TUNINGProject/spanish/doc_fase1/Tuning%20Educational.pdf [Consulta: 30 de junio de 2008].
- GONZALO GARCÍA, C.; García Yebra, V. (eds.). (2004). *Manual de documentación y terminología para la traducción especializada*. Madrid: Arco. (Instrumenta Bibliologica).
- GRELLET, F. (1991). *Apprendre à traduire: typologie d'exercices de traduction*. Nancy: Presses universitaires de Nancy. (Colección Outils et méthodes).
- HANSEN, G. (1997). "Success in translation". En: *Perspectives: studies in translatology* (ISSN 0907-676X). 5(2), pp. 201-210.

- HARRIS, B. (1973). "La traductologie, la traduction naturelle, la traduction automatique et la sémantique". En: *Cahiers de linguistique* (ISSN 0259-6199). 3, pp. 133-146.
- HARRIS, B. (1977). "The importance of natural translation". En: *Working Papers on Bilingualism* (ISSN 0319-5171). 12, pp. 96-114.
- HARRIS, B. (1980). "How a three-year-old translates". En: *Patterns of Bilingualism* (Singapur), pp. 370-393.
- HARRIS, B.; Sherwood, B. (1978). "Translating as an innate skill". En: D. Gerver; W. Sinaiko (eds.). *Language, Interpretation and Communication*. Oxford: Plenum Press, pp. 155-170.
- HATIM, B.; Mason, I. (1997). *The translator as communicator*. Londres: Routledge.
- HERVEY, S.; Higgins, I. (1992). *Thinking translation: a course in translation method: French-English*. Londres; Nueva York: Routledge. (Thinking Translation Series).
- HEWSON, L.; Martin, J. (1991). *Redefining translation. The variational approach*. Londres: Routledge.
- HURTADO, A. (1983/1984). "Apprendre à traduire". En: *Reflét*. 7, pp. 32-38. ("Aprender a traducir", *Boletín de la Asociación Profesional Española de Traductores e Intérpretes*, 22, Madrid, 1984).
- HURTADO, A. (1984). "La formation des traducteurs". En: *Reflét*. 8, pp. 32-38.
- HURTADO, A. (1992). "Didactique de la traduction des textes spécialisés". En: *Actes de la 3ème Journée ERLA-GLAT. Lexique spécialisé et didactique des langues*. Brest: UBO-ENST, pp.9-21.
- HURTADO, A. (1993). "Un nuevo enfoque de la didáctica de la traducción. Metodología y diseño curricular". En: *Les langues étrangères dans l'Europe de l'acte Unique*. Bellaterra: ICE, Universitat Autònoma de Barcelona, pp. 239-252.
- HURTADO, A. (1995). "La didáctica de la traducción: evolución y estado actual". En: P. Fernández; J.M. Bravo (eds.). *Perspectivas de la traducción*. Valladolid: Universidad de Valladolid, pp. 49-74.
- HURTADO (ed.). (1996a). *La enseñanza de la traducción*. Castellón: Publicaciones de la Universitat Jaume I. (Col·lecció Estudis sobre la traducció; 3).
- HURTADO, A. (1996b). "La enseñanza de la traducción directa 'general': objetivos de aprendizaje y metodología". En: A. Hurtado (ed.). *La enseñanza de la traducción*. Castellón de la Plana: Universitat Jaume I. (Col·lecció Estudis sobre la traducció; 3), pp. 31-55.
- HURTADO, A. (1999). *Enseñar a traducir. Metodología en la formación de traductores e intérpretes*. Madrid: Edelsa.
- HURTADO, A. (2001). *Traducción y Traductología: Introducción a la Traductología*. Madrid: Cátedra.

- HURTADO, A. (2004). “A aquisição da competência tradutória: aspectos teóricos e didáticos”. En: A. Pagano; C. Magalhães; F. Alves (org.) *Competência em tradução. Cognição e discurso*. Belo Horizonte: Editora UFMG, pp. 19-57. (Colección Humanitas).
- HURTADO, A. (2007). “Competence-based curriculum design for training translators”. En: *The Interpreter and Translator Trainer* (ISSN 1750-399X). 1(2), pp.163-195.
- HURTADO, A. (2008). “La compétence en traduction et la formation par compétences”. En: *TTR* (ISSN 0835-8443).
- JÄÄSKELÄINEN, R. (1987). *What happens in a translation process: think-aloud protocols of translation*. [Trabajo de investigación / Pro gradu thesis]. University of Joensuu. Savonlinna School of Translation Studies. Savonlinna.
- JÄÄSKELÄINEN, R. (1989). “Translating assignment in professional versus non-professional translation: a think-aloud protocol study”. En: C. Séguinot (ed.) *The translation process*. Toronto: H.G. Publications, pp. 87-98.
- JÄÄSKELÄINEN, R.; Tirkkonen-Condit, S. (1991). “Automatised processes in professional vs. non-professional translation: a think-aloud protocol study”. En: S. Tirkkonen-Condit (ed.) *Empirical research in translation and intercultural studies*. Tübingen: Gunter Narr.
- JESÚS, S. (2008). *Relações de equivalência: a tradução sob a perspectiva da lingüística sistêmico funcional*. [Tesis doctoral]. Universidade Federal de Minas Gerais, Faculdade de Letras. Belo Horizonte.
- KELLY, D. (2002). “Un modelo de competencia traductora: bases para el diseño curricular”. En: *Puentes* (ISSN 1695-3738). 1, pp. 9-20. También disponible en URL: <http://www.ugr.es/~greti/puentes/puentes1/02%20Kelly.pdf> [Consulta: 30 de junio de 2008].
- KELLY, D. (2005). *A handbook for translator trainers*. Manchester: St. Jerome.
- KELLY, D.; et al. (eds.) (2003). *La direccionalidad en traducción e interpretación: perspectivas teóricas, profesionales y didácticas*. Granada: Atrio.
- KIRALY, D.C. (1995). *Pathways to translation: pedagogy and process*. Kent: The Kent State University Press. (Translation Studies Series; 3)
- KIRALY, D.C. (2000). *A social constructivist approach to translator education. Empowerment from theory to practice*. Manchester: St. Jerome.
- KUZNIK, A. (2007). *Marco conceptual y metodológico para un estudio sobre las tareas de traducción e interpretación desempeñadas en el ámbito laboral*. [Trabajo de investigación]. Universitat Autònoma de Barcelona, Departament de Traducció i d'Interpretació. Bellaterra.
- LATORRE, A.; Del Rincón, D.; Arnal, J. (1996). *Bases metodológicas de la investigación educativa*. Barcelona: Hurtado Ediciones.
- LÓPEZ GUIX, J.G.; Minnett Wilkinson, J. (1997). *Manual de traducción (ingles/castellano)*. Barcelona: Gedisa.

- LORENZO, P. (1999). "La seguridad del traductor profesional en la traducción a una lengua extranjera". En: G. Hansen (ed.) *Probing the process in translation: methods and results*. Copenhagen: Samfundslitteratur. (Copenhagen Studies in Language; 24).
- LOWE, P. (1987). "Revising the ACTFL/ETS scales for a new purpose: rating skill in translating". En: M.G. Rose (ed.). *Translation excellence: assessment, achievement, maintenance*. Nueva York: SUNY Binghamton Press. (American Translators Association Series; 1), pp. 53-61.
- MARTÍNEZ MELIS, N. (2001). *Évaluation et didactique de la traduction: le cas de la traduction dans la langue étrangère*. [Tesis doctoral]. Universitat Autònoma de Barcelona, Departament de Traducció i d'Interpretació. Bellaterra.
- MARTÍNEZ MELIS, N.; Hurtado, A. (2001). "Assessment in translation studies: research needs". En: *Meta* (ISSN 0026-0452). 46(2), pp. 272-287. También disponible en <http://www.erudit.org/revue/meta/2001/v46/n2/003624ar.pdf> [Consulta: 30 de junio de 2008].
- MAURI, C. (2006). "La construcción de los personajes en 'A hora da estrela' y su traducción al italiano". En: *Interlingüística* (ISSN 1134-8941). 17, pp. 727-736.
- MINELLI, J.; Pagano, A. (2006). "The research article and the science popularization article: a probabilistic functional grammar perspective on direct discourse representation". En: *Discourse Studies* (ISSN 1461-4456). 8(5), pp. 627-646.
- MONTALT, V. (2005). *Manual de traducció científicotècnica*. Barcelona: Eumo. (Biblioteca de Traducció i Interpretació; 11).
- NEUBERT, A. (2000). "Competence in language, in languages, and in translation". En: Ch. Schäffner; B. Adabs (eds.) *Developing translation competence*. Ámsterdam: John Benjamins, pp. 3-18.
- NEUNZIG, W. (2003). "Tecnologías de la información y traducción especializada inversa". En: D. Kelly; et al. (eds.) *La direccionalidad en traducción e interpretación: perspectivas teóricas, profesionales y didácticas*. Granada: Atrio, pp. 189-206.
- NEUNZIG, W.; Grauwinkel, G. (2006). "Coste, beneficio y riesgo en la traducción especializada inversa". En: S. Bravo Utrera; R. García López (eds.) *Estudios de traducción: problemas y perspectivas*. Las Palmas: Servicio de Publicaciones de la Universidad de Las Palmas de Gran Canaria, pp. 543-555.
- NEWMARK, P. (1988). *A Textbook of Translation*. Londres, Prentice may.
- NORD, C. (1988/1991). *Text-analysis in translation. Theory, methodology, and didactic application of a model for translation-oriented text analysis*. Ámsterdam: Rodopi. [Traducción al inglés del original en alemán (1988) *Textanalyse und Übersetzen*. Heidelberg: Groos].
- NORD, C. (1992). "Text analysis in translator training". En: C. Dollerup; A. Loddegaard. (eds.). *Teaching translation and interpreting: training talent and experience: Papers from the first language international conference, Elsinore, Denmark, 1991*. Ámsterdam: John Benjamins, pp. 39-48.

- PACTE (2000). "Acquiring Translation Competence: Hypotheses and Methodological Problems in a Research Project". En: A. Beeby; D. Ensinger; M. Presas (eds.) *Investigating Translation*. Amsterdam: John Benjamins, pp. 99-106.
- PACTE (2003). "Building a Translation Competence Model". En: F. Alves (ed.). *Triangulating Translation: Perspectives in process oriented research*. Amsterdam: John Benjamins, pp. 43-66.
- PACTE (2005). "Investigating Translation Competence: Conceptual and Methodological Issues". En: *Meta (Processus et cheminements en traduction et interpretation = Processes and Pathways in Translation and Interpretation)* (ISSN 0026-0452). 50(2), pp. 609-619. También disponible en URL: <http://www.erudit.org/revue/meta/2005/v50/n2/011004ar.html> [Consulta: 30 de junio de 2008].
- PAGANO, A.; Magalhães, C.; Alves, F. (orgs.). (2005). *Competência em tradução: cognição e discurso*. Belo Horizonte: Editora UFMG. (Coleção Humanitas).
- PRESAS, M. (1996). *Problemes de traducció i competència traductora: bases per a una pedagogia de la traducció*. [Tesis doctoral]. Universitat Autònoma de Barcelona, Departament de Traducció i d'Interpretació. Bellaterra.
- PYM, A. (1992). "Translation error analysis and the interface with language teaching". En: C. Dollerup; A. Loddegaard (eds.). *Teaching translation and interpreting: training talent and experience*. Amsterdam: John Benjamins, pp. 279-288.
- RISKU, H. (1998). *Translatorische Kompetenz. Kognitive Grundlagen des Übersetzens als Expertentätigkeit*. Tübinga: Stauffenburg.
- RODRÍGUEZ, N.; SCHNELL, B. (2003) "El discreto encanto de la documentación: consideraciones en torno al papel de la documentación en la enseñanza de la traducción inversa". En: D. Kelly; et al. (eds.) *La direccionalidad en traducción e interpretación: perspectivas teóricas, profesionales y didácticas*. Granada: Atrio, pp. 177-188.
- ROISS, S. (2001). "El mercado de la traducción inversa en España: un estudio estadístico". En: *Hermeneus* (ISSN 1139-7489). 3, pp. 397-408.
- SÉGUINOT, C. (1991). "A Study of Student Translation Strategies". En: S. Tirkkonen-Condit (ed.). *Empirical Research in Translation and Intercultural Studies*. Tübinga: Gunter Narr.
- SHREVE, M.G. (1997). "Cognition and the evolution of translation competence". En: J.H. Danks; et al. (eds.). *Cognitive Processes in Translation and Interpreting*. Thousand Oaks: Sage.
- TATILON, C. (1986). *Traduire: pour une pédagogie de la traduction*. Toronto, GREF.
- TIRKKONEN-CONDIT, S. (1990). "Professional vs. Non-Professional Translation: A Think-Aloud Protocol Study". En: M.A.K. Halliday; J. Gibbons; H. Nicholas (eds.). *Learning, Keeping and Using Language*. Amsterdam: John Benjamins, pp. 381-394.
- TOURY, G. (1992). "Everything has its price': an alternative to normative conditioning in translator training". En: *Interface: Journal of applied linguistics* (ISSN 0774-8698). 6(2), pp. 66-76.

- TOURY, G. (1995). *Descriptive Translation Studies –and beyond*. Ámsterdam: John Benjamins.
- VEGA EXPÓSITO, M. (1999). “La enseñanza de la traducción científica y técnica”. En: *Perspectives: Studies in Translatology* (ISSN 0907-676X). 7(2), pp. 241-251.
- VIENNE, J. (1994). “Towards a pedagogy of ‘Translation in Situation’”. En: *Perspectives: Studies in Translatology* (ISSN 0907-676X). 2(1), pp. 51-59.
- VINAY, J.P.; Darbelnet, J. (1958). *Stylistique comparée du français et de l’anglais. Méthode de traduction*. París: Didier.
- WILSS, W. (1992). “The future of translator training”. En: *Meta* (ISSN 0026-0452). 37(3), pp. 391-396. También disponible en URL: <http://www.erudit.org/revue/meta/1992/v37/n3/004609ar.pdf> [Consulta: 30 de junio de 2008].
- YÁNIZ, C.; Villardón, L. (2006). *Planificar desde competencias para promover el aprendizaje: el reto de la sociedad del conocimiento para el profesorado universitario*. Bilbao: Universidad de Deusto. (Cuadernos monográficos del ICE).

CONFERENCIAS NO ESCRITAS

- BARLOW, M. (2000). *An introduction to corpus linguistics*. Taller realizado entre el 19 y el 23 de julio de 2000, en la IV Conferencia TALC, Graz.
- CURADO FUENTES, A.; Edwards Rokowski, P. (2002). *Task feedback in the design of a Business English corpus*. Conferencia pronunciada entre el 26 y el 31 de julio de 2002, en la Fifth TALC Conference, Bertinoro.
- DAVIES, M. (2002). *Student use of a 100-million word, fully annotated corpus of Spanish to model language variation and change*. Conferencia pronunciada entre el 26 y el 31 de julio de 2002, en la Fifth TALC Conference, Bertinoro.
- FLETCHER, W.H. (2001). *Concordancing the Web with KWicFinder* [en línea]. Conferencia pronunciada entre el 23 y el 25 de marzo de 2001, en la American Association for Applied Corpus Linguistics, Third North American Symposium on Corpus Linguistics and Language Teaching, Boston. URL: <http://www.kwicfinder.com/FletcherCLLT2001.pdf> [Consulta: 31 de marzo de 2007].
- FRANKENBERG-GARCIA, A. (2002). *Lost in parallel concordances*. Conferencia pronunciada entre el 26 y el 31 de julio de 2002, en la Fifth TALC Conference, Bertinoro.
- GAVIOLI, L. (1996). *Corpora and the concordancer in learning ESP: an experiment in a course of interpreters and translators*. Conferencia pronunciada entre el 30 de septiembre al 2 de octubre de 1996, en el XVIII Congreso de la Associazione Italiana di Anglistica, Génova.
- GOUADEC, D. (1994). *L'assurance qualité en traduction – perspectives professionnelles, implications pédagogiques*. Conferencia plenaria pronunciada en 1994, en I Jornadas Internacionales de Traducción e Interpretación, Universidad de Las Palmas de Gran Canaria.
- PÉREZ-PAREDES, P.; P. Cantos Gómez (2002). *Some lessons students learn: self-discovery and corpora*. Conferencia pronunciada entre el 26 y el 31 de julio de 2002, en la Fifth TALC Conference, Bertinoro.
- RODRÍGUEZ INÉS, P. (2004). Los corpus electrónicos como recurso especialmente apropiado para la enseñanza a distancia de la traducción. Conferencia pronunciada entre el 10 y el 12 de mayo de 2004, en el VI Congreso internacional de traducción “Enseñanza a distancia de traducción y lengua”, Departament de Traducció i d'Interpretació, Universitat Autònoma de Barcelona, Bellaterra.
- SCHÄFFNER, C. (1996). *Parallel texts in translation*, Conferencia pronunciada entre el 9 y el 11 de mayo de 1996, en la International Translation Studies Conference, Dublin City University, Dublín.
- SIMPSON, R.C. (2000). *Cross-disciplinary comparisons in a corpus of spoken academic English*. Conferencia pronunciada entre el 19 y el 23 de julio de 2000, en la IV Conferencia TALC, Graz.

SINCLAIR, J.M. (1999). *How to use corpora in translation work*. Seminario realizado entre el 22 y el 27 de junio de 1999, en el Tuscan Word Centre, Toscana.

STENSTRÖM, A.B. (1984). *Discourse items and pauses*. Conferencia pronunciada en el 1984, en la Fifth ICAME Conference, Windermere.

