



Facultad de Ciencias Humanas y Sociales
Departamento de Traducción y Comunicación

La infrarrepresentación de elementos únicos en textos
traducidos de inglés a español: perífrasis verbales,
demostrativos y sufijos apreciativos en un corpus
comparable y paralelo de novela policíaca

TESIS DOCTORAL

Bárbara Martínez Vilinsky

Dirigida por:
Dr. Josep Marco Borillo

Castellón de la Plana, noviembre de 2015

Agradecimientos

Me gustaría aprovechar estas líneas para dar las gracias a todos aquellos que, de una manera u otra, han contribuido a la elaboración de esta tesis.

En primer lugar, quiero agradecer a mi director, el Dr. Josep Marco Borillo, por todos los momentos de su vida que ha dedicado a leer, corregir y comentar mis trabajos, desde aquellos remotos tiempos de la novela de Mary. Gracias por creer que tenía madera de investigadora y apostar por mí desde el principio hasta el final y, sobre todo, por esas ocasiones en las que sus palabras de ánimo me empujaron a seguir adelante.

Gracias también a la Universitat Jaume I por proporcionarme la posibilidad, a través de sus becas de formación de personal investigador, de dedicarme por completo a mi formación investigadora y docente durante cuatro años. También a todo el personal docente y administrativo del Departamento de Traducción y Comunicación, por haberme permitido formarme con ellos como alumna e integrarme como compañera, y por haberme ayudado siempre en todo lo que he necesitado.

Me gustaría agradecer, asimismo, al personal administrativo y docente del Centre for Translation and Intercultural Studies (School of Arts, Languages and Cultures) de la Universidad de Mánchester y, especialmente, a la profesora Mona Baker, por haberme ofrecido la oportunidad de realizar una estancia de investigación con ellos, acceder a todos sus recursos y asistir a seminarios de investigación y clases del máster en Estudios de Traducción e Interpretación. Mi eterna gratitud a Bernadette y John, por acogernos durante dicha estancia y hacernos sentir siempre como en casa.

A mis compañeros doctorandos durante esos cuatro años de beca, con los que compartí alegrías y desdichas predoctorales y muchos buenos consejos: Dra. Ana Prats, Miguel Alpuente, Dra. Beatriz Cerezo, Dra. Ana Muñoz, Dra. Irene de Higes, Teresa Molés, Dra. Esperanza Valero, Dra. Verónica Pastor, Dra. Silvia Florez y Ana Tamayo. Me alegra saber que, a pesar de que cada uno ha seguido su propio camino, mantenemos el contacto. Un agradecimiento especial para el Dr. José Luis Martí Ferriol por sus valiosos comentarios y sugerencias, sobre todo en cuestiones estadísticas y

cinematográficas, y para la profesora María Dolores Oltra Ripoll, por guiarme durante mis primeras incursiones en la docencia y transmitirme su entusiasmo por la enseñanza.

A mis amigos. A los de cerca, por poder contar con ellos cuando lo necesitaba y, sobre todo, a María, Bere y Bea, por tener tanta fe en mí y estar siempre convencidas de que ya lo tenía todo hecho cuando en realidad quedaba todo por hacer. A los de lejos, por seguir presentes en mi vida a través de mensajes, llamadas y esporádicos reencuentros. A Patri y Diana, porque, a pesar de estar lejos, las siento siempre a mi lado.

A Pablo, porque sin él, no habría podido sobrevivir hasta el final. Gracias por cuidarme y obligarme a desconectar. Y por tu paciencia infinita.

A mis papás, Enrique y Silvina, porque no estaría escribiendo estas líneas si no fuera por ellos. Gracias por enseñarme el valor del esfuerzo y el pensamiento crítico, por estar siempre ahí y por seguir adelante a pesar de todos los obstáculos.

ÍNDICE

Índice de tablas	vi
Índice de imágenes	x
Índice de gráficos	xi
1. Introducción	1
1.1. Antecedentes e hipótesis	1
1.2. Objetivos y metodología	3
1.3. Estructura de la tesis	5
1.4. Interés y justificación del objeto de estudio	7
2. Marco teórico y estado actual de la cuestión	9
2.1. La lingüística de corpus	9
2.1.1. El debate entre la lingüística teórica y la lingüística de corpus	9
2.1.2. Corpus para el estudio de la lengua	11
2.1.3. Lingüística de corpus: disciplina independiente o metodología	12
2.1.4. La dicotomía corpus-based y corpus-driven en lingüística	13
2.2. Los estudios de traducción basados en corpus	15
2.2.1. El origen de los estudios de traducción basados en corpus	15
2.2.2. El estatus de los ETC y la dicotomía corpus-based y corpus-driven en traducción	16
2.2.3. Corpus para el estudio de la traducción	17
2.2.4. Clasificación y denominación de corpus para el estudio de la traducción	18
2.2.5. Los riesgos de una metodología basada en corpus	23
2.3. Los universales de traducción	24
2.3.1. El concepto de universal de traducción	24
2.3.2. La controversia en torno al concepto de universal	27
2.3.3. El concepto de universal como instrumento metodológico	29
2.3.4. Candidatos a universal de traducción	33
2.3.4.1. La explicitación	34
2.3.4.2. La simplificación	36
2.3.4.3. La normalización	38
2.3.4.4. <i>Levelling out</i>	40
2.3.4.5. La interferencia	41
2.3.4.6. Otros candidatos	43
2.4. La hipótesis de los elementos únicos (Unique Items Hypothesis)	45
2.4.1. Estudios que respaldan la UIH	45
2.4.2. La UIH como candidato a universal de traducción	48

2.4.3.	Argumentos en contra de la UIH como candidato a universal de traducción	51
2.4.4.	Definición y criterios de selección de elementos únicos	53
2.4.5.	Otros estudios relacionados	56
3.	Metodología	61
3.1.	Introducción	61
3.2.	Diseño del corpus	65
3.2.1.	FASE I - Construcción	67
3.2.2.	FASE II - Remodelación	69
3.3.	Creación del corpus	74
3.3.1.	Selección de los textos	74
3.3.1.1.	Medio de representación del texto original	78
3.3.1.2.	Medio de representación del corpus	78
3.3.1.3.	Número de lenguas	79
3.3.1.4.	Características de la preparación de los textos	79
3.3.1.5.	Funciones del corpus	79
3.3.1.6.	Características de los textos seleccionados	79
3.3.2.	Digitalización	93
3.3.3.	Alineación	95
3.3.4.	Segmentación	102
3.3.5.	Etiquetado	104
3.3.6.	Indexación	112
3.4.	Análisis del corpus	116
3.4.1.	WordSmith y ParaConc	118
3.4.2.	Corpus Query Processor	121
3.5.	Tratamiento estadístico de los resultados	126
4.	Perífrasis verbales	133
4.1.	Perífrasis verbales: definición y caracterización	133
4.2.	Selección de perífrasis	138
4.3.	Corpus utilizado para el análisis	141
4.4.	Resultados y discusión	145
4.4.1.	GRUPO I	148
4.4.1.1.	Empezar a + infinitivo - Comenzar a + infinitivo	149
4.4.1.2.	Volver a + infinitivo	156
4.4.1.3.	Seguir + gerundio - Continuar + gerundio	161
4.4.1.4.	Acabar por + infinitivo - Terminar por + infinitivo	167
4.4.1.5.	Acabar + gerundio - Terminar + gerundio	170
4.4.1.6.	Recapitulación de resultados: grupo I	174
4.4.2.	GRUPO II	177
4.4.2.1.	Haber de + infinitivo	178

4.4.2.2.	Haber que + infinitivo	183
4.4.2.3.	Acertar a + infinitivo	188
4.4.2.4.	Llegar a + infinitivo - Alcanzar a + infinitivo	193
4.4.2.5.	Andar + gerundio	198
4.4.2.6.	Ir + gerundio	204
4.4.2.7.	Venir + gerundio	209
4.4.2.8.	Venir a + infinitivo	213
4.4.2.9.	Recapitulación de resultados: grupo II	216
4.4.3.	GRUPO III	219
4.4.3.1.	Querer + infinitivo	220
4.4.3.2.	Ponerse a + infinitivo	225
4.4.3.3.	Echar(se) a + infinitivo	230
4.4.3.4.	Romper a + infinitivo	235
4.4.3.5.	Estar a punto de + infinitivo	237
4.4.3.6.	Dejar de + infinitivo - Parar de + infinitivo	241
4.4.3.7.	Soler + infinitivo	247
4.4.3.8.	Llevar + gerundio	253
4.4.3.9.	Recapitulación de resultados: grupo III	257
4.5.	Conclusiones	261
5.	Demostrativos	273
5.1.	Demostrativos: definición y caracterización	273
5.1.1.	Determinantes y pronombres demostrativos	274
5.1.2.	Adverbios demostrativos locativos	280
5.2.	Características del corpus y métodos de análisis	281
5.2.1.	Corpus utilizado para el análisis	281
5.2.2.	Fiabilidad y utilidad del etiquetado	282
5.2.3.	El criterio del acento diacrítico para distinguir entre determinantes y pronombres	290
5.3.	Resultados y discusión	300
5.3.1.	Distribución de formas demostrativas por textos	300
5.3.2.	Comparación de distribuciones de las formas determinantes y pronominales entre subcorpus	304
5.3.2.1.	Este: cribado y clasificación manual	305
5.3.2.2.	Ese: cribado y clasificación manual	307
5.3.2.3.	Aquel: cribado y clasificación manual	308
5.3.2.4.	Tablas agrupadas: determinantes	309
5.3.2.5.	Tablas agrupadas: pronombres	311
5.3.2.6.	Tablas agrupadas: determinantes y pronombres	315
5.3.3.	Análisis del corpus paralelo	317
5.3.3.1.	Determinantes	327
5.3.3.2.	Pronombres	340
5.3.4.	Conclusiones parciales: determinantes y pronombres demostrativos	350
5.3.5.	Distribución de adverbios demostrativos por textos	357
5.3.5.1.	AQUÍ	358
5.3.5.2.	AHÍ	359

5.3.5.3.	ALLÍ	359
5.3.5.4.	ALLÁ	360
5.3.6.	Comparación de distribuciones de adverbios demostrativos locativos entre subcorpus	361
5.3.7.	Análisis del corpus paralelo	363
5.3.7.1.	HERE	364
5.3.7.2.	THERE	367
5.3.7.3.	AQUÍ	369
5.3.7.4.	AHÍ	373
5.3.7.5.	ALLÍ	376
5.3.7.6.	ALLÁ	379
5.3.8.	Conclusiones parciales: adverbios demostrativos.	381
5.4.	Conclusiones generales: demostrativos	384
6.	Sufijos apreciativos	391
6.1.	Sufijos apreciativos: definición y caracterización	391
6.1.1.	Los diminutivos	395
6.1.2.	Los aumentativos	397
6.1.3.	Los peyorativos	398
6.1.4.	Los sufijos apreciativos como elementos únicos	399
6.2.	Corpus utilizado y selección de sufijos	403
6.3.	Resultados y discusión	411
6.3.1.	Diminutivos	411
6.3.1.1.	-EJO / -EJA	411
6.3.1.2.	-ETE / -ETA	414
6.3.1.3.	-ICO / -ICA	417
6.3.1.4.	-ILLO / -ILLA	419
6.3.1.5.	-ÍN / -INA, -INO / -INA, -IÑO / -IÑA	424
6.3.1.6.	-ITO / -ITA	426
6.3.1.7.	-UCO / -UCA	433
6.3.1.8.	-UELO / -UELA	435
6.3.2.	Aumentativos	438
6.3.2.1.	-AZO / -AZA	438
6.3.2.2.	-ÓN / -ONA	442
6.3.2.3.	-OTE / -OTA	446
6.3.2.4.	-UDO / -UDA	449
6.3.3.	Peyorativos	452
6.3.3.1.	-ACHO / -ACHA	452
6.3.3.2.	-ACO / -ACA	454
6.3.3.3.	-AJO / -AJA	455
6.3.3.4.	-ALES	457
6.3.3.5.	-ANGO / -ANGA, -INGO / -INGA	458
6.3.3.6.	-ASTRO / -ASTRA	460
6.3.3.7.	-ATO / -ATA	461
6.3.3.8.	-ORRIO / -ORRIA, -ORRO / -ORRA	463
6.3.3.9.	-UCHO / -UCHA	465

6.3.3.10.	-UTE	467
6.3.3.11.	-UZO / -UZA	468
6.3.4.	Análisis de otras alternativas de traducción	470
6.3.5.	Datos agrupados	476
6.4.	Conclusiones	481
7.	Conclusiones generales	487
7.1.	Compilación del corpus	488
7.2.	Selección de elementos y resultados	489
7.3.	Definición de <i>unique item</i>	491
7.4.	Confluencia de fenómenos	497
7.5.	Recapitulación	498
7.6.	Futuras líneas de investigación	501
8.	Bibliografía	505
8.1.	Fuentes primarias	505
8.2.	Diccionarios	506
8.3.	Bibliografía científica	506
Introduction and Conclusions (as required for the International Doctorate Mention)		529
Apéndice 1: Proceso de búsqueda y selección de las novelas del corpus y reseña biográfica de los autores y traductores incluidos		553
Proceso de búsqueda y selección de las novelas del corpus		553
Reseña biográfica de los autores y traductores		558
Listado de recursos consultados para la compilación y modificación del corpus y la redacción de las reseñas biográficas de los autores y traductores no mencionados en el cuerpo del texto		568
Apéndice 2: Concordancias		ver CD adjunto

ÍNDICE DE TABLAS

Tabla 1 - Número de palabras y porcentaje sobre el total de los textos de ES	90
Tabla 2 - Número de palabras y porcentaje sobre el total de los textos de TRAD	91
Tabla 3 - Número de palabras y porcentaje sobre el total de los textos de EN	92
Tabla 4 - Comparación de frecuencias entre subcorpus ES y TRAD y significación estadística de los resultados (FASE I)	143
Tabla 5 - Comparación de frecuencias entre subcorpus ES y TRAD y significación estadística de los resultados (FASE II)	144
Tabla 6 - Distribución de «empezar a + INF» por texto	151
Tabla 7 - Distribución de «comenzar a + INF» por texto	152
Tabla 8 - Equivalentes de «empezar a + INF» en EN	153
Tabla 9 - Equivalentes de «comenzar a + INF» en EN	155
Tabla 10 - Distribución de «volver a + INF» por texto	157
Tabla 11 - Equivalentes de «volver a + INF» en EN	158
Tabla 12 - Distribución de «seguir + GER» por texto	162
Tabla 13 - Distribución de «continuar + GER» por texto	164
Tabla 14 - Equivalentes de «seguir + GER» en EN	164
Tabla 15 - Equivalentes de «continuar + GER» en EN	166
Tabla 16 - Distribución de «acabar por + INF» por texto	168
Tabla 17 - Distribución de «terminar por + INF» por texto	169
Tabla 18 - Equivalentes de «acabar por + INF» en EN	169
Tabla 19 - Distribución de «acabar + GER» por texto	172
Tabla 20 - Distribución de «terminar + GER» por texto	172
Tabla 21 - Equivalentes de «acabar + GER» en EN	172
Tabla 22 - Equivalentes de «terminar + GER» en EN	173
Tabla 23 - Distribución de «haber de + INF» por texto	180
Tabla 24 - Equivalentes de «haber de + INF» en EN	181
Tabla 25 - Distribución de «haber que + INF» por texto	185
Tabla 26 - Equivalentes de «haber que + INF» en EN	186
Tabla 27 - Distribución de «acertar a + INF» por texto	189
Tabla 28 - Distribución de «llegar a + INF» por texto	195
Tabla 29 - Distribución de «alcanzar a + INF» por texto	195
Tabla 30 - Equivalentes de «alcanzar a + INF» en EN	196
Tabla 31 - Equivalentes de «llegar a + INF» en EN	197
Tabla 32 - Distribución de «andar + GER» por texto	201
Tabla 33 - Equivalentes de «andar + GER» en EN	202
Tabla 34 - Distribución de «ir + GER» por texto	206
Tabla 35 - Equivalentes de «ir + GER» en EN	206
Tabla 36 - Distribución de «venir + GER» por texto	211
Tabla 37 - Equivalentes de «venir + GER» en EN	211
Tabla 38 - Distribución de «venir a + INF» por texto	215
Tabla 39 - Equivalentes de «venir a + INF» en EN	215
Tabla 40 - Distribución de «querer + INF» por texto	222
Tabla 41 - Equivalentes de «querer + INF» en EN	223
Tabla 42 - Distribución de «ponerse a + INF» por texto	227
Tabla 43 - Equivalentes de «ponerse a + INF» en EN	228

Tabla 44 - Distribución de «echar(se) a + INF» por texto	233
Tabla 45 - Equivalentes de «echar(se) a + INF» en EN	233
Tabla 46 - Distribución de «romper a + INF» por texto	236
Tabla 47 - Equivalentes de «romper a + INF» en EN	236
Tabla 48 - Distribución de «estar a punto de + INF» por texto	238
Tabla 49 - Equivalentes de «estar a punto de + INF» en EN	240
Tabla 50 - Distribución de «dejar de + INF» por texto	242
Tabla 51 - Distribución de «parar de + INF» por texto	243
Tabla 52 - Equivalentes de «dejar de + INF» en EN	244
Tabla 53 - Equivalentes de «parar de + INF» en EN	246
Tabla 54 - Distribución de «soler + INF» por texto	250
Tabla 55 - Equivalentes de «soler + INF» en EN	250
Tabla 56 - Distribución de «llevar + GER» por texto	255
Tabla 57 - Equivalentes de «llevar + GER» en EN	256
Tabla 58 - Factores que influyen en la utilización de las distintas perífrasis en TRAD	265
Tabla 59 - Comparación de las frecuencias de aparición de las cinco perífrasis incoativas	268
Tabla 60 - Reorganización de los grupos de perífrasis según los resultados arrojados por el test de LL	270
Tabla 61 - Comparación de frecuencias de los grupos I y IV	271
Tabla 62 - Clasificación de los determinantes y pronombres demostrativos en español	279
Tabla 63 - Clasificación de los determinantes y pronombres demostrativos en inglés	279
Tabla 64 - Resultados de la búsqueda de demostrativos en ES mediante expresiones regulares	283
Tabla 65 - Resultados de la búsqueda de demostrativos en TRAD mediante expresiones regulares	288
Tabla 66 - Distribución de demostrativos por texto	300
Tabla 67 - Distribución de <i>este</i> por texto	301
Tabla 68 - Distribución de <i>ese</i> por texto	302
Tabla 69 - Distribución de <i>aquel</i> por texto	303
Tabla 70 - Comparación de frecuencias del determinante <i>este</i> en ES y en TRAD y significación estadística de las diferencias observadas	305
Tabla 71 - Comparación de frecuencias del pronombre <i>este</i> en ES y en TRAD y significación estadística de las diferencias observadas	306
Tabla 72 - Comparación de frecuencias del demostrativo <i>este</i> en ES y en TRAD y significación estadística de las diferencias observadas	306
Tabla 73 - Comparación de frecuencias del determinante <i>ese</i> en ES y en TRAD y significación estadística de las diferencias observadas	307
Tabla 74 - Comparación de frecuencias del pronombre <i>ese</i> en ES y en TRAD y significación estadística de las diferencias observadas	307
Tabla 75 - Comparación de frecuencias del demostrativo <i>ese</i> en ES y en TRAD y significación estadística de las diferencias observadas	307
Tabla 76 - Comparación de frecuencias del determinante <i>aquel</i> en ES y en TRAD y significación estadística de las diferencias observadas	308
Tabla 77 - Comparación de frecuencias del pronombre <i>aquel</i> en ES y en TRAD y significación estadística de las diferencias observadas	308
Tabla 78 - Comparación de frecuencias del demostrativo <i>aquel</i> en ES y en TRAD y significación estadística de las diferencias observadas	309
Tabla 79 - Comparación de frecuencias de determinantes demostrativos en ES y en TRAD y significación estadística de las diferencias observadas	309

Tabla 80 - Comparación de frecuencias de pronombres demostrativos en ES y en TRAD y significación estadística de las diferencias observadas	312
Tabla 81 - Comparación de frecuencias de demostrativos en ES y en TRAD y significación estadística de las diferencias observadas	315
Tabla 82 - Frecuencias de aparición de los demostrativos	317
Tabla 83 - Distribución de demostrativos de cercanía por texto en EN (<i>this</i> y <i>these</i>) y en TRAD (<i>este</i> y formas flexionadas)	319
Tabla 84 - Distribución de demostrativos de lejanía por texto en EN (<i>that</i> y <i>those</i>) y en TRAD (<i>ese</i> y <i>aquel</i> y formas flexionadas)	319
Tabla 85 - Elementos equivalentes de los demostrativos <i>that</i> y <i>those</i> en traducciones	321
Tabla 86 - Elementos equivalentes de los demostrativos <i>this</i> y <i>these</i> en traducciones	323
Tabla 87 - Elementos activadores de la utilización del determinante <i>este</i> en traducciones	327
Tabla 88 - Elementos activadores de la utilización del determinante <i>ese</i> en traducciones	332
Tabla 89 - Elementos activadores de la utilización del determinante <i>aquel</i> en traducciones	336
Tabla 90 - Elementos activadores de la utilización del pronombre <i>este</i> en traducciones	340
Tabla 91 - Elementos activadores de la utilización del pronombre <i>ese</i> en traducciones	344
Tabla 92 - Elementos activadores de la utilización del pronombre <i>aquel</i> en traducciones	348
Tabla 93 - Comparación de utilización de demostrativos en ES y TRAD	351
Tabla 94 - Utilización de demostrativos en expresiones temporales en ES	356
Tabla 95 - Utilización de demostrativos en expresiones temporales en TRAD	356
Tabla 96 - Distribución de adverbios demostrativos locativos por texto	357
Tabla 97 - Distribución de <i>aquí</i> por texto	358
Tabla 98 - Distribución de <i>ahí</i> por texto	359
Tabla 99 - Distribución de <i>allí</i> por texto	359
Tabla 100 - Distribución de <i>allá</i> por texto	360
Tabla 101 - Comparación de frecuencias de adverbios demostrativos en ES y en TRAD y significación estadística de las diferencias observadas	361
Tabla 102 - Frecuencias de aparición de los adverbios	363
Tabla 103 - Distribución de <i>here</i> y <i>there</i> por texto en EN	364
Tabla 104 - Elementos equivalentes del adverbio locativo <i>here</i> en traducciones	364
Tabla 105 - Elementos equivalentes del adverbio locativo <i>there</i> en traducciones	367
Tabla 106 - Elementos activadores de la utilización del adverbio <i>aquí</i> en traducciones	369
Tabla 107 - Elementos activadores de la utilización del adverbio <i>ahí</i> en traducciones	373
Tabla 108 - Elementos activadores de la utilización del adverbio <i>allí</i> en traducciones	376
Tabla 109 - Elementos activadores de la utilización del adverbio <i>allá</i> en traducciones	379
Tabla 110 - Porcentaje de utilización de los adverbios demostrativos españoles en el texto meta provocada por la presencia en el texto original de sus equivalentes ingleses	382
Tabla 111 - Sufijos diminutivos en la bibliografía	405
Tabla 112 - Sufijos aumentativos en la bibliografía	406
Tabla 113 - Sufijos peyorativos en la bibliografía	407
Tabla 114 - Sufijo <i>-ejo</i> en ES y TRAD	412
Tabla 115 - Distribución del sufijo <i>-ejo</i> por texto	412
Tabla 116 - Elementos activadores del sufijo <i>-ejo</i> en traducciones	413
Tabla 117 - Sufijo <i>-ete</i> en ES y TRAD	416
Tabla 118 - Distribución del sufijo <i>-ete</i> por texto	416
Tabla 119 - Elementos activadores del sufijo <i>-ete</i> en traducciones	417
Tabla 120 - Sufijo <i>-ico</i> en ES y TRAD	418
Tabla 121 - Distribución del sufijo <i>-ico</i> por texto	419
Tabla 122 - Elementos activadores del sufijo <i>-ico</i> en traducciones	419

Tabla 123 - Sufijo <i>-illo</i> en ES y TRAD	421
Tabla 124 - Distribución del sufijo <i>-illo</i> por texto	422
Tabla 125 - Elementos activadores del sufijo <i>-illo</i> en traducciones	423
Tabla 126 - Sufijos <i>-ín, -ino</i> e <i>-iño</i> en ES y TRAD	424
Tabla 127 - Distribución de los sufijos <i>ín, -ino</i> e <i>-iño</i> por texto	425
Tabla 128 - Elementos activadores de los sufijos <i>-ín, -ino</i> e <i>-iño</i> en traducciones	425
Tabla 129 - Sufijo <i>-ito</i> en ES y TRAD	429
Tabla 130 - Distribución del sufijo <i>-ito</i> por texto	430
Tabla 131 - Elementos activadores del sufijo <i>-ito</i> en traducciones	431
Tabla 132 - Sufijo <i>-uco</i> en ES y TRAD	434
Tabla 133 - Distribución del sufijo <i>-uco</i> por texto	434
Tabla 134 - Elementos activadores del sufijo <i>-uco</i> en traducciones	434
Tabla 135 - Sufijo <i>-uelo</i> en ES y TRAD	436
Tabla 136 - Distribución del sufijo <i>-uelo</i> por texto	437
Tabla 137 - Elementos activadores del sufijo <i>-uelo</i> en traducciones	438
Tabla 138 - Sufijo <i>-azo</i> en ES y TRAD	439
Tabla 139 - Derivados de <i>-azo</i> no incluidos	441
Tabla 140 - Distribución del sufijo <i>-azo</i> por texto	441
Tabla 141 - Elementos activadores del sufijo <i>-azo</i> en traducciones	441
Tabla 142 - Sufijo <i>-ón</i> en ES y TRAD	443
Tabla 143 - Derivados de <i>-ón</i> no incluidos	445
Tabla 144 - Distribución del sufijo <i>-ón</i> por texto	445
Tabla 145 - Elementos activadores del sufijo <i>-ón</i> en traducciones	446
Tabla 146 - Sufijo <i>-ote</i> en ES y TRAD	447
Tabla 147 - Distribución del sufijo <i>-ote</i> por texto	448
Tabla 148 - Elementos activadores del sufijo <i>-ote</i> en traducciones	449
Tabla 149 - Sufijo <i>-udo</i> en ES y TRAD	450
Tabla 150 - Distribución del sufijo <i>-udo</i> por texto	450
Tabla 151 - Elementos activadores del sufijo <i>-udo</i> en traducciones	451
Tabla 152 - Sufijo <i>-acho</i> en ES y TRAD	452
Tabla 153 - Distribución del sufijo <i>-acho</i> por texto	453
Tabla 154 - Elementos activadores del sufijo <i>-acho</i> en traducciones	453
Tabla 155 - Sufijo <i>-aco</i> en ES y TRAD	454
Tabla 156 - Distribución del sufijo <i>-aco</i> por texto	454
Tabla 157 - Sufijo <i>-ajo</i> en ES y TRAD	455
Tabla 158 - Distribución del sufijo <i>-ajo</i> por texto	456
Tabla 159 - Elementos activadores del sufijo <i>-ajo</i> en traducciones	456
Tabla 160 - Sufijos <i>-ango, -engo, -ingo, -ongo</i> y <i>-ungo</i> en ES y TRAD	458
Tabla 161 - Distribución de los sufijos <i>-ango, -engo, -ingo, -ongo</i> y <i>-ungo</i> por texto	459
Tabla 162 - Elementos activadores de los sufijos <i>-anga, -engo, -ingo, -ongo</i> y <i>-ungo</i> en traducciones	460
Tabla 163 - Sufijo <i>-astro</i> en ES y TRAD	460
Tabla 164 - Distribución del sufijo <i>-astro</i> por texto	461
Tabla 165 - Elementos activadores del sufijo <i>-astro</i> en traducciones	461
Tabla 166 - Sufijo <i>-ato</i> en ES y TRAD	462
Tabla 167 - Distribución del sufijo <i>-ato</i> por texto	463
Tabla 168 - Elementos activadores del sufijo <i>-ato</i> en traducciones	463
Tabla 169 - Sufijos <i>-orro, -orrio</i> y <i>-orio</i> en ES y TRAD	464
Tabla 170 - Distribución de los sufijos <i>-orrio, -orro</i> y <i>-orio</i> por texto	465
Tabla 171 - Sufijo <i>-ucho</i> en ES y TRAD	466

Tabla 172 - Distribución del sufijo <i>-ucho</i> por texto	466
Tabla 173 - Elementos activadores del sufijo <i>-ucho</i> en traducciones	467
Tabla 174 - Sufijo <i>-ute</i> en ES y TRAD	467
Tabla 175 - Distribución del sufijo <i>-ute</i> por texto	468
Tabla 176 - Sufijo <i>-uzo</i> en ES y TRAD	469
Tabla 177 - Distribución del sufijo <i>-uzo</i> por texto	469
Tabla 178 - Uso de sufijos apreciativos españoles en la traducción de sufijos ingleses	471
Tabla 179 - Adjetivos de tamaño en EN y sus equivalentes en TRAD	474
Tabla 180 - Frecuencias de uso de sufijos apreciativos en ES y TRAD	477
Tabla 181 - Frecuencias de uso normalizadas de sufijos apreciativos en ES por texto	478
Tabla 182 - Frecuencias de uso normalizadas de sufijos apreciativos en TRAD por texto	479
Tabla 183 - Frecuencia y porcentaje de uso de elementos activadores de la utilización de sufijos apreciativos en TRAD	480

ÍNDICE DE IMÁGENES

Imagen 1 - Esquema de un corpus bidireccional	21
Imagen 2 - Estadísticas del subcorpus de originales ES (fase I)	87
Imagen 3 - Estadísticas del subcorpus de traducciones TRAD (fase I)	88
Imagen 4 - Estadísticas del subcorpus de originales ES (fase II)	91
Imagen 5 - Estadísticas del subcorpus de traducciones TRAD (fase II)	91
Imagen 6 - Estadísticas del subcorpus de originales EN (fase II)	92
Imagen 7 - Reconocimiento de caracteres con el programa OmniPage	94
Imagen 8 - Primera fase del procesamiento de los textos: fragmento de un texto de ES en formato TXT	95
Imagen 9 - Asistente para la alineación de Déjà Vu: fragmentos desordenados	98
Imagen 10 - Asistente para la alineación de Déjà Vu: fragmento dividido	99
Imagen 11 - Asistente para la alineación de Déjà Vu: fragmento unido	99
Imagen 12 - Exportación de una memoria de traducción de Déjà Vu	101
Imagen 13 - Segunda fase del procesamiento de los textos: fragmentos de EN y TRAD alineados en un único archivo	102
Imagen 14 - Tercera fase del procesamiento de los textos: fragmentos de EN y TRAD segmentados en dos archivos	104
Imagen 15 - Interfaz para Windows del programa de etiquetado morfológico y lematización TreeTagger	109
Imagen 16 - Cuarta fase del procesamiento de los textos: fragmento de TRAD etiquetado con TreeTagger	110
Imagen 17 - Interfaz del programa WinSCP	113
Imagen 18 - Pantalla de acceso a CWB a través del cliente Putty	114
Imagen 19 - Quinta fase del procesamiento de los textos: fragmento de ES preparado para la indexación	115
Imagen 20 - Ventana de selección de textos de WordSmith Tools	117
Imagen 21 - Ventana de búsqueda de concordancias de WordSmith Tools	118
Imagen 22 - Ventana de resultados de la búsqueda de concordancias con WordSmith Tools	119
Imagen 23 - Textos alineados en ParaConc	119

Imagen 24 – Concordancias paralelas en ParaConc	120
Imagen 25 – Acceso a CQP mediante Putty	121
Imagen 26 - Listado de concordancias de la palabra allá en CQP	123
Imagen 27 - Listado de concordancias de la palabra <i>allá</i> en CQP con las etiquetas visibles	124
Imagen 28 – Calculadora de log-likelihood en línea de Paul Rayson	128
Imagen 29 – Pantalla de resultados de la calculadora de log-likelihood	129

ÍNDICE DE GRÁFICOS

Gráfico 1 - Comparación de frecuencias normalizadas de determinantes demostrativos en ES y TRAD	310
Gráfico 2 - Comparación de porcentajes de uso de determinantes demostrativos en ES y TRAD	311
Gráfico 3 - Comparación de frecuencias normalizadas de pronombres demostrativos en ES y TRAD	312
Gráfico 4 - Comparación de porcentajes de uso de pronombres demostrativos en ES y TRAD	313
Gráfico 5 - Comparación de frecuencias normalizadas de demostrativos en ES y TRAD	316
Gráfico 6 - Comparación de porcentajes de uso de demostrativos en ES y TRAD	317
Gráfico 7 - Porcentajes de uso de los distintos elementos equivalentes de los demostrativos <i>that</i> y <i>those</i> en traducciones	321
Gráfico 8 - Porcentajes de uso de los distintos elementos equivalentes de los demostrativos <i>this</i> y <i>these</i> en traducciones	323
Gráfico 9 - Porcentajes de uso de los distintos elementos activadores del determinante <i>este</i> en traducciones	327
Gráfico 10 - Porcentajes de uso de los distintos elementos activadores del determinante <i>ese</i> en traducciones	333
Gráfico 11 - Porcentajes de uso de los distintos elementos activadores del determinante <i>aquel</i> en traducciones	336
Gráfico 12 - Porcentajes de uso de los distintos elementos activadores del pronombre <i>este</i> en traducciones	340
Gráfico 13 - Porcentajes de uso de los distintos elementos activadores del pronombre <i>ese</i> en traducciones	344
Gráfico 14 - Porcentajes de uso de los distintos elementos activadores del pronombre <i>aquel</i> en traducciones	348
Gráfico 15 - Comparación de frecuencias normalizadas de adverbios demostrativos en ES y TRAD	362
Gráfico 16 - Comparación de porcentajes de uso de adverbios demostrativos en ES y TRAD	362
Gráfico 17 - Porcentajes de uso de los distintos elementos equivalentes del adverbio locativo <i>here</i> en traducciones	365
Gráfico 18 - Porcentajes de uso de los distintos elementos equivalentes del adverbio locativo <i>there</i> en traducciones	367
Gráfico 19 - Porcentajes de uso de los distintos elementos activadores del adverbio <i>aquí</i> en traducciones	370

Gráfico 20 - Porcentajes de uso de los distintos elementos activadores del adverbio <i>ahí</i> en traducciones	373
Gráfico 21 - Porcentajes de uso de los distintos elementos activadores del adverbio <i>allí</i> en traducciones	377
Gráfico 22 - Porcentajes de uso de los distintos elementos activadores del adverbio <i>allá</i> en traducciones	379

1. Introducción

1.1. Antecedentes e hipótesis

¿Son especiales las traducciones? ¿Existe algo en la naturaleza de un texto traducido que lo caracterice como tal? Esta es una pregunta que muchos investigadores se han hecho a lo largo de los años, y las maneras en las que han intentado responderla han ido modificándose con el tiempo. Sin embargo, en las últimas décadas esta cuestión ha adquirido especial relevancia. La rama de los estudios basados en corpus la ha convertido en el eje principal de su tarea investigadora. Ya no se busca comparar originales con traducciones para realizar afirmaciones de tipo evaluativo con respecto a aciertos o errores de traducción. El foco de atención ha cambiado y ahora el objetivo es hallar los aspectos en los que los textos traducidos difieren de aquellos producidos en circunstancias naturales o de manera espontánea —es decir, sin el estímulo de un texto en la lengua de partida— en el mismo idioma. Encontrar aquellas características que distinguen a los textos traducidos y nos permiten identificarlos como tales, intentando estudiarlas sin emitir juicios de valor.

Esta idea, anticipada ya en las décadas de los setenta y ochenta por autores como Frawley (1984) o Toury (1977), se popularizó a mediados de los noventa gracias a una serie de artículos publicados por Mona Baker (1993, 1995, 1996) que han gozado de enorme reconocimiento y han sentado las bases de lo que hoy conocemos como Estudios de Traducción Basados en Corpus. En dichos artículos, Baker defiende la idea de que existen en la traducción, como proceso con objetivos, presiones y contextos de producción específicos, ciertas características que le son intrínsecas y se manifiestan independientemente de factores externos como podrían ser el par de lenguas de trabajo o las condiciones particulares del encargo. A estas características las denomina *universales de traducción* y sugiere, además, que las técnicas y herramientas que ya llevaba tiempo utilizando la lingüística de corpus son las que ofrecen las mejores posibilidades de hallarlas. Por lo tanto, la idea que plantea es la de utilizar la metodología de la lingüística de corpus para estudiar la naturaleza particular del fenómeno de la traducción, que, en su opinión, se manifestará en regularidades

lingüísticas o patrones diferentes de los que podríamos hallar en producciones originales en la misma lengua.

Basándose en las conclusiones de determinados estudios que se habían centrado en la naturaleza de la lengua traducida pero a una escala mucho menor que la que podían ofrecer los análisis con grandes corpus de textos, Baker propuso a la comunidad científica algunos posibles candidatos a universal. Entre ellos se encontraban las tendencias de los textos traducidos a incorporar información añadida para hacer el mensaje más comprensible (explicitación), a codificar la información de manera más sencilla que el original (simplificación), a utilizar estructuras menos creativas o más convencionales en la lengua meta (normalización) y a evitar los extremos, gravitando hacia el término medio (*levelling out*). Desde entonces, numerosos investigadores han dedicado sus esfuerzos a intentar confirmar o refutar, mediante estudios basados en corpus, la existencia de los universales enumerados por Baker. Durante estos años, se han realizado estudios que han aportado pruebas a favor y en contra; algunos universales han recibido más atención, otros han resultado tener una base menos sólida de lo que se había pensado en un principio.

También se han propuesto nuevos candidatos a universales, entre los que se incluye el que constituye el principal objeto de estudio de este trabajo: la infrarrepresentación en textos traducidos de elementos únicos de la lengua meta. La hipótesis de la infrarrepresentación surgió a partir de los trabajos de autoras escandinavas como Mauranen (2000), Eskola (2004) o Tirkkonen-Condit (2002). Esta última realizó, en 2004, un estudio en el que descubrió que existen ciertos elementos que, a pesar de ser típicos del finés, presentan una frecuencia menor de aparición en textos traducidos con respecto a textos escritos originalmente en este idioma. A estos elementos los denominó *unique items*, y los definió como elementos que no cuentan con un equivalente inmediatamente accesible en la lengua de partida y, por consiguiente, no se activan con facilidad en el momento de traducir. Tirkkonen-Condit explicaba este fenómeno argumentando que los elementos de la lengua de llegada que son únicos o típicos de ésta, en el sentido de que no se representan de la misma manera en otros idiomas, no encuentran estímulo en los textos de partida para su activación. Por lo tanto, las traducciones tenderían a manifestar estos elementos con menor frecuencia que los textos originales.

La Unique Items Hypothesis, como la bautizó Tirkkonen-Condit, ha sido respaldada mayoritariamente por estudiosos finlandeses y suecos. La hipótesis que

fundamenta el presente proyecto de investigación se basa en dichas conclusiones, ya que lo que se pretende es averiguar si este fenómeno se observa también en el par de lenguas inglés-español. La hipótesis de partida de este trabajo, por tanto, es la siguiente: «En los textos traducidos del inglés al español se produce un fenómeno de infrarrepresentación de los elementos únicos o típicos de la lengua de llegada cuando no tienen un estímulo lingüístico directo en la lengua de partida».

1.2. Objetivos y metodología

Baker afirmaba que, para comprobar si una característica es o no realmente universal, es preciso realizar estudios empíricos en diversas combinaciones de lenguas y con textos de distintos géneros y períodos de tiempo. Nuestro objetivo es, por tanto, averiguar si este fenómeno de infrarrepresentación de ciertos elementos o estructuras únicas de la lengua meta en textos traducidos con respecto a textos escritos originalmente en dicha lengua, hallada por Tirkkonen-Condit en textos académicos y de ficción traducidos al finés, puede encontrarse también en textos literarios traducidos de inglés a español y, por lo tanto, es válida su defensa como candidato a universal de traducción. Para ello, intentaremos replicar estudios realizados anteriormente en torno a esta hipótesis pero alterando determinados factores como la combinación lingüística o las características del corpus utilizado y, por supuesto, los elementos de análisis, y comprobaremos si los resultados obtenidos confirman la Unique Items Hypothesis.

La metodología utilizada en este trabajo será la propia de los estudios de traducción basados en corpus, es decir, el estudio de características lingüísticas a través de conjuntos de textos compilados en corpus electrónicos de distintos tipos (en este caso, paralelo y comparable) y la ayuda de herramientas automáticas de análisis y extracción de datos. Nos basaremos en la consulta de un subcorpus de español no traducido para determinar las frecuencias normales de los elementos seleccionados, de un subcorpus de español traducido para observar las diferencias de uso y frecuencia entre ambas variedades, y así determinar si se produce o no un fenómeno de infrarrepresentación de estos elementos, y, finalmente, de un subcorpus conformado por los textos originales de dichas traducciones en inglés, para poder contrastar el texto

origen con su traducción e intentar encontrar posibles correlaciones o explicaciones para los fenómenos observados.

Por lo tanto, el primer objetivo práctico de nuestro trabajo será la compilación de un corpus electrónico comparable (en términos de género, función, variedad lingüística y demás criterios que estableceremos con anterioridad) de textos de ficción contemporánea —novelas policíacas publicadas a lo largo de la década de los noventa— en español traducido y no traducido y de un componente paralelo compuesto por los textos de partida de las traducciones del componente comparable.

En segundo lugar, será fundamental hallar, a través de la consulta de diversas gramáticas y estudios contrastivos, un elemento del español que encaje con la definición atribuida por lo general a estos *elementos únicos*, es decir, elementos de naturaleza léxica, fraseológica o gramatical que no cuenten con equivalentes inmediatamente accesibles al traductor en la lengua de partida, sino que sean características singulares del español que otras lenguas deban codificar de maneras distintas, a través de recursos como la paráfrasis.

La compilación del componente comparable facilitará nuestro tercer objetivo práctico, que es el análisis de grandes cantidades de texto en relación con los elementos presuntamente únicos del español. Una vez hayamos averiguado su frecuencia de aparición en el subcorpus de textos no traducidos mediante la extracción semiautomática de concordancias con un programa de análisis textual, será posible comprobar si hay diferencias con respecto a su uso en el subcorpus de traducciones y, de haberlas, si estas son significativas. Esto permitirá determinar si se produce o no un fenómeno de infrarrepresentación de dichos elementos en textos traducidos con respecto a textos originales.

Si la respuesta fuera afirmativa, cabría plantearse si este fenómeno se debe a la influencia de la lengua de origen o es una característica inherente al proceso traductor. En este último paso entraría en escena el tercer subcorpus de textos originales en inglés, que servirá para comparar el texto de partida con las soluciones traductorales y averiguar, mediante el análisis detallado de las concordancias paralelas extraídas del corpus alineado, si la diferencia en las frecuencias observadas o los modos de utilización de los elementos estudiados se debe o no a la falta de elementos paralelos en la lengua de partida. Complementaremos, por tanto, el análisis cuantitativo de frecuencias con el cualitativo, y el estudio sistemático de los patrones de utilización de estos elementos en ambas lenguas nos proporcionará las claves para dar un paso más allá de la simple

descripción de fenómenos observados e intentar aventurar posibles correlaciones o conexiones hipotéticamente causales que podrían explicarlos. Todo ello nos permitirá contestar a las diversas preguntas que nos hemos planteado: ¿existen diferencias entre español traducido y no traducido con respecto a las frecuencias de aparición de determinados elementos únicos de nuestra lengua? ¿Son las diferencias significativas? ¿Proviene de la falta de elementos activadores en el inglés? ¿Son consecuencia de diferencias estructurales entre ambas lenguas o una característica inherente al proceso de traducción?

Como explicaremos en el capítulo dedicado al estado de la cuestión en los Estudios de Traducción Basados en Corpus, autoras como Halverson (2003) o Hareide (2013) han relacionado el fenómeno de la infrarrepresentación con una hipótesis denominada Gravitational Pull Hypothesis, que predice que determinados factores cognitivos ejercerán una influencia mayor o menor en la selección de un elemento lingüístico por parte del traductor. Un objetivo secundario de nuestro trabajo será, si se verifica la infrarrepresentación de alguno de los elementos estudiados, intentar identificar, a partir de las conclusiones extraídas del análisis, y en la medida de las posibilidades de un trabajo como el que presentamos aquí, qué papel desempeñan dichos factores en este fenómeno. También intentaremos averiguar qué características de las incluidas por Hareide en su definición de *unique item* se ven verificadas por nuestros resultados.

1.3. Estructura de la tesis

Para alcanzar estos objetivos será necesario, antes de nada, analizar el origen de los Estudios de Traducción Basados en Corpus y revisar el estado de la cuestión en este ámbito de la disciplina. Por otro lado, habrá que definir y explicar el concepto de universal de traducción, que ha estado siempre rodeado de mucha polémica, y determinar si su utilización es beneficiosa o superflua para la teoría de la traducción. Finalmente, nos centraremos en la descripción de los diferentes estudios llevados a cabo en torno al candidato a universal de la infrarrepresentación de elementos únicos en

traducciones. A estas cuestiones dedicaremos el **capítulo 2**, «Antecedentes y estado de la cuestión».

El segundo paso, la compilación de un corpus comparable y paralelo de textos literarios en la combinación lingüística inglés-español, se tratará en el **capítulo 3** («Metodología»). En él —y en el apartado incluido como **Apéndice 1** al final de la tesis— describiremos minuciosamente el proceso de diseño del corpus, la selección, recopilación y preparación de los textos incluidos en el mismo, los programas y las técnicas utilizadas para llevar a cabo el análisis y las dificultades con las que nos encontramos a lo largo de todo el proceso.

Los **capítulos 4, 5 y 6** se dedicarán al análisis de los elementos únicos seleccionados. Al principio de cada capítulo se presenta una revisión de la bibliografía consultada sobre el elemento en cuestión y la justificación de su selección como elemento único. Después de los resultados de cada uno de los tres análisis, incluimos unas conclusiones preliminares sobre cada elemento en particular.

En el **capítulo 4** analizamos un conjunto de 27 perífrasis verbales de infinitivo y gerundio, elementos que, por su mayor número y variedad en el sistema lingüístico español en contraposición con el inglés, son, a juicio de numerosos autores, una peculiaridad de nuestro idioma. En el **capítulo 5** estudiamos la estructura tripartita del sistema de determinantes y pronombres demostrativos (*este, ese y aquel* y demás formas neutras y flexionadas) y de los adverbios locativos *aquí, ahí y allí* y sus variantes, en contraposición con la división formal dicotómica del inglés (*this y that; here y there*). Por último, en el **capítulo 6** nos ocupamos de una selección de sufijos apreciativos diminutivos, aumentativos y peyorativos, un recurso morfológico para expresar cuantificación o valoración afectiva mucho menos presente en inglés que en español.

En el **capítulo 7** («Conclusiones») reunimos las conclusiones parciales de los tres capítulos de análisis e intentamos contestar a las preguntas que nos hacemos en esta introducción y determinar si es posible validar la Unique Items Hypothesis a partir de los resultados de nuestro estudio.

1.4. Interés y justificación del objeto de estudio

En definitiva, lo que queremos averiguar es si ciertos elementos característicos del español se utilizan en menor medida en textos traducidos que en textos no traducidos en dicha lengua y, de ser así, si esto es debido a la falta de elementos equivalentes en la lengua del texto original, si depende del tipo de elementos del que estemos hablando o de otros factores. El candidato a universal que nos ocupa es particularmente interesante, puesto que no se han realizado demasiados estudios al respecto hasta el momento y aún menos que se hayan centrado en el español. La mayor parte se han realizado en combinación con el finés (un idioma sin parentesco directo con la mayor parte del resto de lenguas europeas) y, aunque se han obtenido algunos resultados positivos también con el francés (Cappelle 2012) y el chino (Xiao y Dai 2014) como lengua meta, el único trabajo que conocemos hasta la fecha que estudia la Unique Items Hypothesis en español como uno de sus objetivos principales (Hareide 2013), no encuentra evidencias a favor de la misma. Con este trabajo esperamos abrir nuevas vías de investigación en este idioma.

El estudio de la posible existencia de características compartidas por todos los textos traducidos es especialmente interesante y relevante por las consecuencias metodológicas y didácticas que puede conllevar su corroboración (o no). A partir de la adquisición de conocimientos cada vez más amplios sobre aquellos rasgos que podrían ser inherentes al proceso de trasvase de un texto de una lengua a otra, seremos más conscientes de lo que hacemos o dejamos de hacer al traducir, de nuestros recursos y nuestras limitaciones, y, por tanto, podremos aprender y enseñar a traducir de manera más eficaz. Asimismo, si descubrimos que la traducción posee características propias y diferentes a las de otras situaciones comunicativas multilingües, podremos seguir construyendo la disciplina a partir de datos reales de uso de la lengua en el contexto específico de esta actividad.

Si, finalmente, la hipótesis de la infrarrepresentación no se viera confirmada, el análisis de un corpus de traducciones como el nuestro puede proporcionar datos muy interesantes con respecto a otros fenómenos, universales o no, presentes en el tipo de textos estudiados, y contribuir al conocimiento de estos y de los elementos lingüísticos seleccionados para el estudio. El análisis de las concordancias paralelas aportará datos

valiosos sobre las distintas equivalencias de traducción de dichos elementos, lo que puede ser de utilidad en estudios contrastivos de inglés y español y en el ámbito de la enseñanza de la traducción.

2. Marco teórico y estado actual de la cuestión

2.1. La lingüística de corpus

2.1.1. *El debate entre la lingüística teórica y la lingüística de corpus*

Recently, the term “corpus” has come to refer specifically to a large collection of electronic texts that have been gathered according to explicit criteria (in order to be used as a representative sample of a particular language or subset of that language). (Bowker 2002: 44)

Antes de la era de los ordenadores, un corpus podía ser simplemente un texto o quizás un conjunto de textos sobre el que se basaba, por ejemplo, un estudio sobre un autor, género o tema concreto. Hoy en día, como se extrae de la definición de Bowker, se presupone el formato electrónico y ya no vale con uno o dos textos: los corpus actuales contienen cientos de miles o millones de palabras. Muy acertadamente dice Leech (1998: xvi) que «corpus linguistics» significa hoy en día «computer corpus linguistics».

Podemos afirmar, sin miedo a equivocarnos, que prácticamente todos los investigadores del ámbito de la lingüística y de la traducción consideran que el desarrollo de la lingüística de corpus —tal y como la conocemos ahora, es decir, el estudio de la lengua basado en el análisis de enormes cantidades de texto mediante herramientas informáticas— ha traído consigo una revolución en múltiples ámbitos y disciplinas. Si bien es cierto que no a todos este cambio les ha parecido positivo, lo que es innegable es que la manera de investigar ha cambiado radicalmente desde que la evolución de la informática, reflejada principalmente en el aumento de la capacidad de almacenamiento de datos de los ordenadores y la sofisticación de las herramientas de procesamiento y análisis textual, han permitido al fin a los investigadores tener acceso a corpus extensos de textos auténticos con ejemplos del uso real de la lengua, y dejar de

dependen exclusivamente de su propia intuición lingüística, de ejemplos de uso prefabricados o descontextualizados o del análisis manual de un número limitado de casos. Así lo explican Biber, Conrad y Reppen (1994: 170): «The analyses in each area show that, as linguists, we often have strongly held intuitions, but those intuitions frequently prove to be incorrect when they are tested empirically against the actual patterns of use in large text corpora».

Dentro del ámbito de la lingüística ha habido quien no ha compartido el mismo entusiasmo por el uso de corpus, puesto que el enfoque de la lingüística tradicional — especialmente de los representantes de la gramática generativa— se basaba, sobre todo, en la época en la que surgieron estos, en la intuición del hablante nativo y no concedía demasiada importancia a los datos de uso real de la lengua (McEnery, Xiao y Tono 2006: 7; Meyer 2002: 1). Gries (2010: 331) no está de acuerdo con esta polarización de posturas: «I find that, for instance, Fillmore’s polarization of arm-chair linguists vs. corpus linguists is too often taken way too seriously: theoretical linguists need (corpus) data, but CL also needs (more) theory». Tanto un método como otro se basa finalmente en determinadas intuiciones o ideas sobre la naturaleza del lenguaje; la diferencia es que la lingüística de corpus no se detiene en este punto, sino que extrae sus conclusiones a partir de los resultados de análisis empíricos (Müller y Waibel 2009).

A pesar de las críticas, lo cierto es que numerosos investigadores han apostado por seguir el nuevo camino que ha marcado la metodología de corpus en los ámbitos de la investigación del lenguaje, los estudios contrastivos y los estudios de traducción, como evidencia la enorme cantidad de trabajos que pueden encontrarse actualmente en compilaciones como las actas de las distintas ediciones del Congreso Internacional de Lingüística de Corpus —disponibles en la página web de AELINCO¹ (Asociación Española de Lingüística de Corpus)—, en los que investigadores de todo el mundo se valen de la lingüística de corpus para analizar diferentes cuestiones de naturaleza muy diversa en ámbitos tan variados como la lexicografía, la didáctica, la extracción terminológica, la estilística, la variación lingüística o la traducción.

También es prueba de ello la cantidad cada vez mayor de corpus generales y especializados que se han ido creando en todo el mundo².

¹ Disponibles en <http://www.aelinco.es/es/actas-publicaciones-relacionadas> [última fecha de consulta: 05/09/2015].

² Muller y Waibel (2009) distinguen entre distintos tipos de corpus: corpus generales (p. ej. el BNC), sincrónicos (LOB y Brown), históricos (ARCHER o A Representative Corpus of Historical English Registers), de aprendizaje (Cambridge Learner Corpus), corpus para el estudio de variedades de una lengua (ICE) y especializados, es decir, compilados para distintos tipos de estudios específicos (Michigan Corpus of Academic Spoken English).

2.1.2. *Corpus para el estudio de la lengua*

Según Zanettin (2012: 7), el Brown Corpus fue uno de los primeros corpus electrónicos. Se creó en la década de los sesenta, en la Universidad de Brown, en Estados Unidos. Contenía un millón de palabras de inglés americano y, junto con el LOB (Lancaster-Oslo/Bergen Corpus, de inglés británico, completado en la década de los setenta entre la Universidad de Lancaster y la de Oslo) y el London Lund (comenzado en la Universidad de Lund en 1975 y constituido por conversaciones espontáneas en inglés británico), formó parte de una propuesta para aunar esfuerzos de investigadores y proyectos interesados en la lingüística computarizada, denominada «International Computer Archive of Modern English» (Johansson 1978)³. Hoy en día, los corpus cuentan ya con varios millones de palabras y pueden encontrarse en todo el mundo y en numerosas combinaciones lingüísticas. El British National Corpus, por ejemplo, cuenta con cien millones de palabras de textos escritos y transcripciones de textos orales⁴, el COCA o Corpus of Contemporary American English (Davies 2009) cuenta con 450 millones de palabras⁵ y el Bank of English, que es un corpus «monitor», es decir, que se actualiza constantemente para reflejar los cambios lingüísticos, actualmente contiene 650 millones de palabras⁶. En el ámbito hispanoamericano disponemos del corpus de referencia del español actual, el CREA⁷, que en 2008 contaba con 160 millones de palabras, y un corpus diacrónico de 250 millones de palabras, el CORDE⁸, que recoge textos en español de todas las épocas y lugares hasta 1975, ambos

Xiao, por su parte, proporciona una tipología diferente y un listado amplísimo de corpus en la siguiente página: <http://cw.routledge.com/textbooks/0415286239/resources/corpa.htm#S1> [consultado el 04/09/2015]. Casi todos los corpus mencionados en esta sección se encuentran clasificados y descritos en este recurso.

³ Según la descripción de Biber (1990: 257): «The Standard Corpus of Present-Day Edited American English, known as the Brown Corpus, contains one million words of written, edited American English published in 1961; the corpus comprises 500 text samples, each 2,000 words long, taken from fifteen different text categories (e.g. press reportage, editorials, academic prose, general fiction). The LOB corpus is a British English parallel to the Brown Corpus. The London-Lund Corpus is based on the spoken texts collected as part of the Survey of English Usage; it comprises eighty-seven text samples, each at least 5,000 words long, representing several major speech situations (e.g. face-to-face and telephone conversations, radio broadcasts, interviews, public speeches)». Siguiendo el modelo del Brown Corpus se han construido otros corpus comparables de distintas variedades de inglés (p. ej. australiano o indio) y también de otras lenguas (chino).

⁴ Disponible en <http://www.natcorp.ox.ac.uk/>

⁵ Disponible en <http://corpus.byu.edu/coca>

⁶ Disponible en <http://www.collins.co.uk/page/The+Collins+Corpus>

⁷ Disponible en <http://www.rae.es/recursos/banco-de-datos/crea>

⁸ Disponible en <http://www.rae.es/recursos/banco-de-datos/corde>

compilados por la Real Academia Española. También existen corpus multilingües: un ejemplo es el English-Norwegian Parallel Corpus (ENPC)⁹ (Johansson 1998: 8), uno de los primeros corpus paralelos, que contiene textos de distintos géneros traducidos y no traducidos tanto en inglés como en noruego y que más tarde se expandió al Oslo Multilingual Corpus¹⁰. Podemos mencionar además el ACTRES Parallel Corpus o P-ACTRES¹¹, con cuatro millones de palabras de textos originales y traducidos en inglés y español. Más adelante veremos otros ejemplos de corpus del ámbito de la traducción.

2.1.3. *Lingüística de corpus: disciplina independiente o metodología*

Sin embargo, no todo el mundo está de acuerdo en considerar a la lingüística de corpus simplemente una metodología:

In terms of what corpus linguistics ‘is’, not only have various definitions been offered, but alternatives have been explicitly addressed and rejected. These include, as we shall see: *corpus linguistics* is a *tool*, a *method*, a *methodology*, a *methodological approach*, a *discipline*, a *theory*, a *theoretical approach*, a *paradigm* (theoretical or methodological), or a combination of these. (Taylor 2008: 180)

Entre quienes defienden la idea que la lingüística de corpus es más que un conjunto de métodos o herramientas están Tognini-Bonelli: «corpus linguistics has, therefore, theoretical status» (2001: 1); Aarts: «corpus linguistics does exist, and it is a discipline harbouring two mutually exclusive methodologies, the corpus-based and the

⁹ Disponible en <http://www.hf.uio.no/ilos/english/services/omc/enpc/>. Según esta página web: «It started out as a research project at the Department of British and American Studies, University of Oslo, in 1994. The corpus was completed in 1997. In the period 1997-2001 the corpus was extended to include more languages (German, Dutch, Portuguese) and the English and the Norwegian original texts were tagged for part of speech. The manual was completed in 1999 and revised in 2002. The focus has been on novels and fairly general non-fictional books. In order to include material by a range of authors and translators, the texts of the corpus are limited to text extracts (chunks of 10,000-15,000 words). The fiction part of the corpus contains 30 original text extracts in each language and their translations, whereas the non-fiction part contains 20 in each direction».

¹⁰ Disponible en <https://www.hf.uio.no/ilos/english/services/omc/>. Según esta página web: «The Oslo Multilingual Corpus is a product of the interdisciplinary research project Languages in Contrast (SPRIK), which is a collaboration between researchers at the Faculty of Humanities, University of Oslo. The OMC contains many sub-corpora that differ in composition with regard to languages and number of texts included. It is mainly the languages Norwegian, English, French, and German that are represented in the sub-corpora, but some of the corpora include Dutch and Portuguese texts. In addition, there are related parallel corpora for English-Swedish and English-Finnish, compiled in Sweden and Finland, which are accessible from the same site».

¹¹ Disponible en <http://actres.unileon.es/inicio.php?elementoID=12>

corpus-driven methodologies» (2002: 14); Mahlberg (2005: 31 y 2006: 370), quien describe su visión de un «corpus theoretical framework» para estudiar elementos como los sustantivos o fenómenos como la cohesión textual; Teubert (2005: 2), que se refiere a ella como «a theoretical approach to the study of language», o Leech (1992: 106), que afirma que no es «just a newly emerging methodology for studying language, but a new research enterprise, and in fact a new philosophical approach to the subject».

2.1.4. *La dicotomía corpus-based y corpus-driven en lingüística*

El debate que acabamos de describir aún parece encontrarse abierto y se encuentra muy relacionado con la diferencia que hacen ciertos autores entre «*corpus-based* y *corpus-driven studies*». Los primeros constituirían el conjunto de estudios que se realizan utilizando corpus y herramientas de análisis informático para probar diferentes teorías e hipótesis formuladas con anterioridad dentro de la disciplina. Los segundos se alinearían con la postura de autores como Tognini-Bonelli (2001: 177) o Sinclair (2003: xv). Estos investigadores consideran que debemos enfrentarnos a los datos que nos proporcionan los corpus con una actitud en cierta forma ingenua y abierta, no condicionada (en la medida de lo posible) por teorías lingüísticas previas. Solo de esta manera podremos aprovechar al máximo las nuevas oportunidades que nos proporcionan los corpus.

While corpus linguistics may make use of the categories of traditional linguistics, it does not take them for granted. It is the discourse itself, and not a language-external taxonomy of linguistic entities, which will have to provide the categories and classifications that are needed to answer a given research question. This is the corpus-driven approach. (Teubert 2005: 4)

Esta visión viene fundamentada por el hecho de que el análisis de textos reales ha permitido a los investigadores en los últimos tiempos darse cuenta de que los fenómenos lingüísticos son muchas veces más complejos de lo que puede pensarse o de lo que puede abarcar una teoría o modelo lingüístico. Por ejemplo, los estudios con corpus de Sinclair sobre la relación entre léxico y gramática revolucionaron la lingüística, demostrando el importantísimo efecto que tienen las colocaciones sobre el

significado de las palabras, y llevaron a la creación de un nuevo modelo de diccionario basado en corpus de textos reales. El análisis de textos auténticos puede contradecir nuestras ideas preconcebidas sobre el lenguaje y es ahí donde ven su mayor aportación los autores que defienden que los estudios de corpus deberían partir de la evidencia proporcionada por estos y no solo basarse en ejemplos extraídos de ellos para respaldar determinadas teorías e hipótesis. Por eso mismo, como explicaremos con más detalle en el apartado oportuno, según esta visión tampoco deberíamos alterar los textos añadiendo etiquetado ni anotaciones de ningún tipo; esto equivaldría a imponer una teoría lingüística a nuestro análisis y, por tanto, sesgar nuestras conclusiones.

The potential of corpus evidence is not exploited fully because, in order not to threaten dramatically some existing theoretical positions, the richness of language use is in many ways sacrificed and is not allowed to shape the descriptive and theoretical statements that should ideally account for it. (Tognini-Bonelli 2001: 11)

Esta postura tiene mucho sentido, pero, en su versión más extrema, peca de idealismo. Gries (2010: 329), por ejemplo, argumenta que es muy difícil que alguien pueda enfrentarse al análisis del lenguaje sin ideas preconcebidas sobre el mismo. Como dice Stubbs (1993: 24): «there is no such thing as pure induction». Por otra parte, según nuestra experiencia, un estudio de corpus puede partir de una hipótesis previa o una idea preconcebida, pero los resultados arrojados por el análisis reflejarán la realidad de los textos y muchas veces desafiarán dichas hipótesis e ideas. Dependerá de la honestidad del investigador el utilizar o descartar los datos discordantes. En este trabajo, por otra parte, las categorías que hemos utilizado para agrupar los resultados extraídos del corpus paralelo son muchas veces creaciones ad-hoc, configuradas de manera deductiva a partir de la evidencia hallada. Como afirma Sinclair (1991: 36): «Without relinquishing our intuitions, of course, we try to find explanations that fit the evidence, rather than adjusting the evidence to fit a pre-set explanation».

Para McEnery, Xiao y Tono (2006: 11), las diferencias entre un enfoque y otro no son tan acusadas. Como ellos, en este trabajo englobaremos ambos enfoques con el término *corpus-based* o *basados en corpus*. Estos mismos autores (2006: 8) y otros como Saldanha (2009: 2), defienden además la idea de que la lingüística de corpus no es una teoría o disciplina autónoma, sino una metodología o, quizás, como lo denomina Corpas Pastor (2008: 49), un nuevo paradigma de investigación, cuyos métodos y

herramientas pueden aplicarse a multitud de campos distintos, pero no conforman en sí mismos una teoría o rama autónoma de la lingüística. Muller y Waibel (2009) lo explican de la siguiente manera: «Corpus linguistics is thus a method to obtain and analyse data quantitatively and qualitatively rather than a theory of language or even a separate branch of linguistics on a par with e.g. sociolinguistics or applied linguistics. The corpus-linguistic approach can be used to describe language features and to test hypotheses formulated in various linguistic frameworks».

2.2. Los estudios de traducción basados en corpus

2.2.1. El origen de los estudios de traducción basados en corpus

Hemos resumido escuetamente algunos de los debates más relevantes en torno a la lingüística de corpus porque, como es lógico, terminaron siendo exportados a los Estudios de Traducción en los años 90, tras el célebre artículo de Mona Baker (1993) en el que la autora proponía la metodología empírica de corpus (refiriéndose especialmente a los corpus monolingües de traducciones y no traducciones) como la manera más adecuada de estudiar el lenguaje de la traducción y hallar las características que le son propias¹².

Como hemos explicado, la lingüística de corpus es una rama de la lingüística general que surgió con el objetivo de fundamentar las teorías acerca del uso del lenguaje mediante datos empíricos auténticos en forma de corpus de lenguaje natural, principalmente monolingües. Los corpus de traducciones ya habían sido utilizados con anterioridad como herramienta de la lingüística contrastiva, pero no para centrarse en el fenómeno de la traducción. Por otro lado, los estudios descriptivos de la traducción se remontan a los trabajos que Toury publica a partir de finales de los años setenta —en

¹² A pesar de que es innegable que fue gracias a Baker que se popularizó esta idea en el ámbito de los Estudios Descriptivos de la Traducción, explica Laviosa (2004: 32) que, por supuesto, no fue ella la primera en pensar que la utilización de corpus podía ser provechosa para la disciplina (aunque las investigaciones anteriores se centraran en su vertiente aplicada): «In 1986 Martin Gellerstam at the University of Lund had in fact already compiled the first monolingual comparable corpus of Swedish novels to study translationese and Hans Lindquist in 1989, again in Lund, had investigated the Swedish renderings of English adverbials with a parallel language data base. Their research intended to use corpora as aids to improve the practice of translation».

los que caracteriza la rama descriptiva de los estudios de traducción y postula como su objetivo principal la búsqueda de patrones y regularidades que podrían dar lugar a leyes probabilísticas de traducción y, finalmente, conformar una teoría general de la traducción sustentada en datos empíricos—, y, por supuesto, al mapa de la disciplina que traza Holmes en 1972 en su artículo «The name and nature of Translation Studies». Estos trabajos parten de la necesidad de dejar de lado el enfoque prescriptivo que valoraba la traducción únicamente en relación con el grado de equivalencia o fidelidad con el texto original, y de empezar a estudiar la traducción *per se* para constituir una disciplina con entidad propia. A partir de la publicación del artículo de Baker se unieron los caminos de la rama descriptiva de la traducción y la lingüística de corpus, conformando así los «estudios de traducción basados en corpus» o, simplemente, «estudios de traducción con corpus (ETC)» (Corpas Pastor 2008).

2.2.2. *El estatus de los ETC y la dicotomía corpus-based y corpus-driven en traducción*

Como en el ámbito de la lingüística, existen también en el de la traducción posturas más cercanas a los enfoques *corpus-based* y otras más cercanas a los *corpus-driven*, y hay divergencias en cuanto al estatus atribuible a los Estudios de Traducción con Corpus. Laviosa (2000: 160), tras mencionar el asombroso desarrollo de los ETC y los numerosos trabajos llevado a cabo en diversos ámbitos teóricos y prácticos de la traducción mediante el análisis automático del lenguaje, concluye: «The richness, variety, and internal theoretical and methodological coherence of these studies justify the claim that corpus-based translation studies cannot be equated simply to a new way of investigating translational phenomena, but also to a new way of making sense of them». House (2011: 206), por el contrario, afirma que los ETC no son una nueva rama de los Estudios de Traducción, sino simplemente «a methodological basis for pursuing translation research», cuya utilidad debería ser el poder combinarse con otras maneras más tradicionales de analizar la traducción para enriquecerlas. Fantinuoli y Zanettin (2015: 1, 8), por su parte, opinan que la lingüística de corpus se ha convertido en «a major paradigm and research methodology in translation theory and practice» y los estudios de traducción basados en corpus, a lo sumo en «a disciplinary sub-category»

que se abre continuamente a nuevos enfoques y marcos teóricos y genera cada vez más bibliografía. En nuestra opinión, dada la tendencia a utilizar las herramientas de la lingüística de corpus para estudiar fenómenos de muy diversa naturaleza y, más recientemente, a combinarlos con otras técnicas de carácter más experimental (como los *think-aloud protocols*), la postura más lógica parece ser considerarla una metodología, productiva sobre todo en la rama descriptiva de la traducción (especialmente en el estudio de características de los textos traducidos y del estilo individual de determinados traductores), y, en la medida de lo posible, complementar los resultados obtenidos a partir de esta con los de otros métodos.

2.2.3. *Corpus para el estudio de la traducción*

A estas polémicas, se añade el hecho de la denominación de los distintos tipos de corpus que se utilizan en nuestro ámbito, que no coincide muchas veces con las de otros campos. Es el caso de la lingüística contrastiva, por ejemplo, cuyos objetivos no son exactamente los mismos que los de los estudios de traducción (Ramón 2002: 404)¹³, pero comparte algunos de sus intereses principales —entre ellos, la descripción, comparación y contraste de lenguas— y por ello utiliza también las herramientas de la lingüística de corpus.

En realidad, como argumenta Olohan (2002: 419), hasta hace poco la lingüística no había prestado demasiada atención a las traducciones, que no habían sido incluidas en los grandes corpus de referencia, quizás porque no se las consideraba muestras fiables del uso real de la lengua. Se temía que contuvieran trazas de *translationese*, que reflejaran un lenguaje teñido por la influencia de la lengua de partida (Johansson 1998), y este temor sigue presente en la actualidad. McEnery, Xiao y Tono (2006: 93) consideran que el uso de corpus de traducciones para estudios contrastivos puede llevar a falsas conclusiones debido al efecto de la lengua origen en la lengua meta. Las traducciones, en su opinión, son una versión no representativa de la lengua real. Sin

¹³ Según esta autora, «CL aims at acquiring knowledge about two languages on a descriptive basis and TS aims at acquiring knowledge about the actual translation process between two languages. Taking into account these fundamental differences in the goals of the two disciplines and the great similarities between them pointed out in this paper, we can conclude that CL acts on a more general basis, while TS takes the contrast between two languages to a much more specialised ground». En realidad, como explica Ramón (2002: 395), antes de conseguir estatus como disciplina autónoma en los años ochenta, la traducción era considerada un caso especial de comparación entre lenguas y, por tanto, una subdivisión de la lingüística contrastiva.

embargo, hay autores que consideran que la lengua traducida es parte del uso lingüístico habitual —de hecho, en determinados géneros importados de otras culturas, a veces ha precedido a la producción espontánea en lengua meta— y como tal, debería estar representada en los corpus de referencia.

Sea como sea, hoy en día existen numerosos corpus que incluyen traducciones. Uno de los primeros fue el Translational English Corpus (p. ej. Baker 2007), compilado por Mona Baker y Sara Laviosa en la Universidad de Manchester y conformado por textos traducidos al inglés a partir de una variedad de lenguas de partida¹⁴. También podemos mencionar el Corpus of Translational Finnish (CTF) (Mauranen 2004: 72) o el ZJU Corpus of Translational Chinese (ZCTC) (Dai y Xiao 2011: 8). Más tarde se conformaron los primeros corpus de traducciones que incluían también un componente paralelo, como el Dutch Parallel Corpus (Macken, De Clercq y Paulussen 2011)¹⁵, de más de 10 millones de palabras de textos traducidos del holandés al inglés y el francés y viceversa, o COMPARA¹⁶, también bidireccional de inglés y portugués (Santos y Frankenberg-García 2007).

En general, casi todos los autores parecen estar de acuerdo en que no hay una única manera de diseñar un corpus ni criterios válidos para todas las circunstancias. Trataremos la importantísima y muy controvertida cuestión de los criterios de compilación de un corpus (y características como el equilibrio, la representatividad, el tamaño, etcétera) en el apartado de metodología en el que describimos el proceso de creación del corpus en el que basamos nuestro trabajo.

2.2.4. Clasificación y denominación de corpus para el estudio de la traducción

Como adelantábamos, es posible encontrar muy diversas clasificaciones de corpus en la literatura. Aunque, en esencia, casi todos los autores describen los mismos tipos de corpus, existen discrepancias en cuanto a la denominación de dichos tipos. Las mayores divergencias, como apuntan McEnery y Xiao (2007: 19), Granger (2003: 19) o Bowker (2002: 46), se dan en lo que respecta a la palabra *paralelo*, que puede referirse a

¹⁴ Disponible en <http://www.llc.manchester.ac.uk/ctis/research/english-corpus/>

¹⁵ Disponible en <http://www.kuleuven-kulak.be/DPC>

¹⁶ Disponible en <http://www.linguateca.pt/COMPARA>

textos originales en diferentes lenguas compilados según los mismos criterios o a un corpus de textos originales y traducciones. A este último, algunos autores lo denominan «de traducción». Para ejemplificar estas diferencias, reproduciremos la clasificación de tres autores: Johansson, Granger y Zanettin.

Johansson (2007) se refiere en primer lugar a los corpus multilingües —que son los utilizados en los estudios lingüísticos comparativos sobre los que trata este autor en su libro— y menciona los siguientes:

- ✓ corpus de traducción (textos originales y sus traducciones a una o más lenguas)
- ✓ corpus comparables (textos originales en dos o más lenguas compilados según criterios coincidentes de género o fecha de publicación, entre otros)

En un artículo anterior (1998: 4), el autor mencionaba también un tercer tipo de corpus, el *learner language corpus*, y dos tipos adicionales, relevantes para los estudios de traducción, pero también como material de control en los estudios contrastivos que utilizan corpus de originales y traducciones:

- ✓ corpus de textos originales y traducidos en una misma lengua
- ✓ corpus de textos traducidos en lenguas distintas

Granger (2003: 20), consciente de la diversidad terminológica imperante, y refiriéndose sobre todo a las diferencias existentes entre la lingüística contrastiva y la traducción, propone su unificación en la siguiente tipología:

Corpus multilingües:

- ✓ de traducción/paralelos, que contienen textos de partida y sus traducciones y pueden ser unidireccionales o bi/multidireccionales.
- ✓ comparables (textos traducidos o no traducidos del mismo género)

Corpus monolingües (todos son comparables):

- ✓ textos originales y traducidos en una misma lengua
- ✓ textos escritos por nativos y por estudiantes de la misma lengua

Una clasificación más extendida en los Estudios de Traducción es la explicada en Zanettin 2000 y 2012:

- ✓ **Comparable monolingüe:** un conjunto de textos traducidos y otro de textos no traducidos (producidos de manera espontánea) en la misma lengua, seleccionados teniendo en cuenta criterios similares.
- ✓ **Comparable bilingüe:** textos producidos de manera espontánea en dos lenguas distintas, agrupados según criterios de similitud textual (los típicos textos paralelos que se utilizan en los ámbitos profesionales o de enseñanza de la traducción).
- ✓ **Paralelo:** (el denominado corpus de traducción por autores como los mencionados más arriba) conjunto de traducciones a una lengua y sus respectivos textos originales en otra lengua.¹⁷

En este trabajo utilizaremos la primera y la última categoría y nos referiremos a ellas simplemente como corpus comparables y paralelos, en línea con la denominación de Baker (1995) para corpus orientados al estudio específico de la traducción. Más adelante hablaremos de lo que opinaba Baker sobre los corpus y su utilidad en nuestra disciplina; especialmente con respecto a los corpus comparables.

Por el momento, expondremos la postura de Zanettin (2013: 28), quien opina que un modelo más útil de corpus sería el que reuniera las características de ambos tipos, paralelo y comparable: un corpus paralelo bidireccional o recíproco, es decir, con traducciones en ambas direcciones. Esta modalidad es la que idealmente permitiría realizar todo tipo de comparaciones, aunque, como veremos más adelante, el autor expresa sus reservas acerca de los problemas que puede acarrear la construcción de este tipo de corpus. Un corpus bidireccional incluiría, entonces, un subcorpus de traducciones a una lengua A y sus respectivos originales en lengua B (componente paralelo). También un componente comparable (a las traducciones a A) de traducciones a la lengua B con sus respectivos textos originales en lengua A. Este último subcorpus debería ser al mismo tiempo comparable a los originales en lengua B. De esta manera, tendríamos un sistema completo que suele representarse de la siguiente manera:

¹⁷ Las dos últimas categorías las agrupa en el trabajo de 2012 bajo la denominación *corpus bilingüe* (paralelo y comparable) y añade el *multilingüe* (cuando hablamos de más de dos lenguas) como tercera categoría.

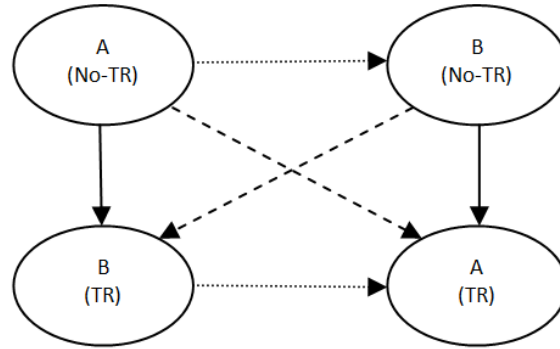


Imagen 1 - Esquema de un corpus bidireccional

Las flechas de trazo continuo indican la relación entre los subcorpus de textos originales en una determinada lengua A o B y sus respectivas traducciones a la otra lengua (componente paralelo). Las flechas formadas por guiones unen subcorpus de textos traducidos y no traducidos en una misma lengua (componente comparable). Las flechas punteadas indican la posibilidad de comparar textos traducidos o no traducidos entre las dos lenguas.

Dos ejemplos de este modelo de corpus son el ENPC y el COMPARA, a los que ya hemos hecho referencia anteriormente. Uno de los principales responsables de la compilación del ENPC fue Johansson. Como hemos explicado cuando describíamos su tipología de corpus, este autor se cuestiona la fiabilidad de los textos traducidos a la hora de hacer generalizaciones sobre el uso habitual del lenguaje en estudios contrastivos, debido a la posibilidad de que presenten trazas de la influencia del texto de partida o características inherentes al proceso de traducción, del tipo de las que habla Baker¹⁸. Para controlar dichos efectos y separar las características específicas de los textos traducidos de las de los textos producidos, por así decirlo, de manera «espontánea», defiende la bidireccionalidad de los corpus de traducciones y la inclusión de un corpus de control comparable, es decir, originales y traducciones en la misma lengua (1998: 6)¹⁹.

¹⁸ Labrador (2007: 31-32) matiza de la siguiente manera esta «desconfianza» hacia los textos traducidos: «This does not mean that translated texts are less entitled to belong to the target language (onwards TL). Curiously enough, sometimes they even contain more TL phrases, structures, and so on particularly characteristic of the TL than the TL texts themselves, what has been called “the Eliza Doolittle phenomenon” (Malmkjær 1998:535). The reason for not choosing explicitly translated texts as the sources of data is that they usually differ, to a greater or lesser extent, from original texts, as they may contain traces of the source language and their creation process is different from that of original writing. And the main purpose of this kind of contrastive research is to observe and present how similar ideas and concepts are naturally expressed in different languages».

¹⁹ La expansión del ENPC de la que hemos hablado anteriormente, mediante traducciones de los textos originales en inglés a más lenguas (noruego, holandés y portugués), refleja la intención de sus creadores de distinguir las características debidas al proceso de traducción de las que surgen de la comparación de pares de lenguas específicos.

Otros autores también abogan por este modelo de corpus mixto, aunque motivados por otros factores. Kenny, por ejemplo (1998: 3), utiliza un corpus paralelo para estudiar la hipótesis de la *sanitization* de colocaciones en traducciones: lo que hace en primer lugar es extraer colocaciones del subcorpus de originales, para luego confirmar su convencionalidad (o no convencionalidad) mediante un corpus de referencia de textos en la misma lengua. Goethals (2007: 93) sostiene que

the specific design criteria of corpus research (bidirectional, large samples and the search for the same category in both ST and TT) make it possible to formulate differentiated hypotheses and to assign them to the discipline that seems best placed to investigate them: Contrastive Linguistics when the quantitative data reveal a difference between two languages, irrespective of their ST and or TT status, Translation Studies when the data reveal a difference between STs and TTs, irrespective of the language pair involved, and Text Linguistics when the data reveal important differences between the samples of the same subcorpus.

Sin embargo, Zanettin (2000: 109; 2012: 49) explica que este modelo ideal pocas veces es realizable, debido a las diferencias existentes entre los universos literarios de las distintas culturas, ya que lo que se traduce y en qué medida se traduce puede diferir mucho de una a otra, y esto puede significar que tengamos que sacrificar la comparabilidad de los subcorpus en aras de la representatividad de las poblaciones de las que provienen, o viceversa. De hecho, su proyecto de corpus CEXI (Translational English Italian Corpus), que no llegó a completarse, fue un claro ejemplo de ello: los investigadores responsables descubrieron que el flujo de traducciones es mucho mayor en una dirección que en la otra y que, además, se traduce un tipo distinto de literatura en cada cultura. Por ejemplo, un tercio de todos los libros vendidos en Italia son traducidos del inglés y, en su mayoría, *best sellers*, mientras que en Inglaterra y Estados Unidos solo de un 2 a un 3 % son traducciones, de las cuales un porcentaje mínimo corresponde al italiano, y, en general, a narrativa de una calidad literaria elevada (Zanettin 2012: 50). Así que, en ocasiones, tendremos que resignarnos a trabajar con el mejor corpus posible en la práctica y renunciar a buscar la configuración de corpus ideal que propugna la teoría (McEnery y Xiao 2007: 27).

2.2.5. *Los riesgos de una metodología basada en corpus*

Una última polémica, relacionada en cierto modo con la que mencionábamos al principio entre la lingüística tradicional y la lingüística de corpus, es la generada por autores que temen que la metodología de corpus empuje a los investigadores a centrarse demasiado en cifras y datos individuales y a dejar de lado cuestiones fundamentales de las que deberían ocuparse los Estudios de Traducción.

Tymoczko (1998) cree que uno de los riesgos es que nos concentremos demasiado en la búsqueda de regularidades y similitudes, cuando la infinitud de diferencias existentes entre las lenguas y la traducción entre ellas pueden enseñarnos mucho también. Por otro lado, advierte de que no debemos caer en la tentación de cuantificar sin motivo y de intentar probar lo evidente. Para ello es importante hacerse las preguntas adecuadas. Por último, es fundamental adoptar un punto de vista diacrónico y no diseñar corpus a partir de la idea de que el universo de la traducción se circunscribe a las normas vigentes actualmente y a un único contexto cultural. En Olohan (2003: 61) se expone la visión plasmada por Mason en un estudio de 2001, en el que sostiene que debemos ser capaces de combinar los resultados numéricos obtenidos a través de corpus, con el análisis de factores contextuales como la influencia ejercida por el género, el discurso y el propósito del texto original o el objetivo último de la traducción sobre las decisiones del traductor.

Laviosa y Baker intentan mitigar estas preocupaciones, apuntando que los investigadores son conscientes de estos riesgos y tienen en cuenta los factores sociolingüísticos (Laviosa 2000: 160-161). Los corpus ya están siendo utilizados para estudiar la singularidad de las traducciones y otros aspectos como la ideología o el estilo del traductor (Laviosa 2004: 47). Por otra parte, tienen claro que el objetivo del estudio de regularidades a través de corpus no debería ser la búsqueda de reglas absolutas (Baker 1998: 4). La ventaja de conocer estas regularidades es, según la autora, que «leur identification permet non seulement de cerner les traits récurrents de la traduction, et par conséquent de comprendre la traduction comme un phénomène en soi, mais elle nous autorise également à comprendre l'exemple original».

Más allá de todas estas cuestiones que dividen a los expertos, lo cierto es que, como en la lingüística de corpus, en traducción también se ha terminado por generalizar

el uso de corpus electrónicos para el estudio de todo tipo de cuestiones del ámbito docente, profesional e investigador, que es el que nos ocupa en este trabajo.

2.3. Los universales de traducción

2.3.1. *El concepto de universal de traducción*

Como habíamos adelantado, en una serie de artículos publicados durante la década de los noventa, Baker (1993, 1995, 1996, 1998) sugirió aplicar los métodos de la lingüística de corpus a la descripción de las características de la lengua traducida, inaugurando los Estudios de Traducción con Corpus. En el primero de dichos artículos, Baker rechaza el punto de vista hasta ese momento mayoritario de los académicos del ámbito de la traducción, que se habían centrado en la comparación de los textos traducidos con los textos originales en un intento de hallar relaciones entre ellos, investigando conceptos como el de equivalencia o correspondencia. Su idea era que las traducciones constituirían un objeto de estudio legítimo en sí mismas, y el objeto de estudio al que debían dedicarse, como es lógico, los Estudios de Traducción. Baker defendía la idea de que todas las traducciones presentan ciertas características lingüísticas sólo por el hecho de ser traducciones, independientemente de cualquier otro factor externo, incluida la lengua de partida. A estas características las denominó *universales de traducción*, y las definió de la siguiente manera: «features which typically occur in translated texts rather than original utterances and which are not the result of interference from specific linguistic systems» (1993: 243).

Este concepto no es nuevo; como apunta la misma Baker en su artículo, debe mucho a autores como Frawley, quien, en un trabajo de 1984, afirmaba que la traducción es en esencia un *tercer código*, derivado de los de la lengua origen y meta pero distinto a ambos, «setting its own standards and structural presuppositions and entailments, though they are necessarily derivative of the matrix information and target parameters» (1984: 169). Desde una perspectiva más prescriptiva, también se ha considerado durante mucho tiempo que las traducciones presentan un lenguaje propio

que se desvía de las normas de la lengua meta: el denominado *translationese*, que algunos autores utilizan en sentido neutro (aclarándolo por lo general para evitar confusiones) pero que, en general, presenta connotaciones negativas. Puurtinen (1998: 2) lo define como «features which distinguish translated texts from original target language texts; for example lexical items or syntactic structures which are not used in their normal target language functions or whose distribution is unusual». Xiao y McEnery (2007: 24) siguen refiriéndose a estas características como «problemáticas», sobre todo en el ámbito de los estudios contrastivos: «As these features are regular and typical of translated English, further research based upon these findings may not only uncover the translational norms or what Frawley (1984) calls the “third code” of translation. It will also help translators and trainee translators to become aware of these problems».

Baker defiende la postura de los teóricos que, a partir de los años setenta, comenzaron a deshacerse de los métodos de análisis introspectivos y el punto de vista prescriptivo, y, en lugar de seguir juzgando las traducciones desde el punto de vista de la equivalencia con un texto traducido o el grado de fidelidad hacia el mismo, o de una manera ideal de traducir, defendieron la necesidad de estudiar traducciones reales y, a partir de ellas, hallar generalidades que permitieran comprender qué es y cómo funciona la traducción. El principal exponente de esta postura es Toury, quien —como él mismo declara en dos artículos de 2004 (Toury 2004a: 24, 2004b: 29)— en 1977 hablaba ya de la posible naturaleza universal de ciertas características del lenguaje traducido. Más tarde desarrolló esta idea postulando dos posibles leyes del proceso de traducción que podrían mostrar una tendencia universal a manifestarse: la ley de estandarización creciente y la ley de interferencia, a las que nos referiremos más adelante:

much energy should be directed toward the crystallization of systematic research methods, including statistical ones, especially if we wish to transcend the study of norms, which are always limited to one societal group at a time, and move on to the formulation of general laws of translational behavior, which would inevitably be probabilistic in nature. (Toury 1995: 69)

Su idea era desarrollar, a través del estudio de normas y leyes, la rama descriptiva de la traducción, fundamental, en su opinión, para alimentar (con observaciones y datos auténticos) las otras dos ramas —las ramas aplicada y teórica del mapa de la disciplina

que trazó Holmes en los años setenta— y conseguir que los Estudios de Traducción se constituyeran como una disciplina verdaderamente autónoma.

No obstante, hasta la publicación del primer artículo de Baker en 1993 solo se habían realizado estudios de alcance muy limitado. Un ejemplo es el de Blum-Kulka, cuyo estudio sobre cohesión y coherencia es uno de los citados por Baker para apoyar su lista de universales, puesto que en él se postula la hipótesis de la explicitación (cohesiva) como característica independiente de los dos sistemas lingüísticos y textuales implicados y, por tanto, inherente al proceso de traducción (Blum-Kulka 1986: 19). Sin embargo, hay que tener en cuenta que Blum-Kulka parece extraer sus conclusiones del análisis de una pequeña colección de traducciones al francés, inglés y hebreo.

La idea de Baker, inspirada por los trabajos en lingüística computacional de Sinclair, era reunir la mayor cantidad de material textual posible, para empezar a fundar las conclusiones de la rama descriptiva de la disciplina en observaciones de suficiente entidad. Su propuesta metodológica consistía en compilar, por un lado, un corpus formado por grandes cantidades de texto traducido a la lengua meta desde distintos idiomas y, por otro, un corpus de textos originales escritos en dicha lengua meta, comparables al corpus de traducciones en términos de cantidad de palabras, género, función o fecha de publicación, entre otros criterios de inclusión, y utilizarlos para buscar ciertas características que, según su teoría, se encontrarían representadas en mayor medida en el corpus de traducciones. Esto permitiría encontrar regularidades y patrones de uso del lenguaje que serían característicos de la lengua traducida y la diferenciarían de la no traducida. El objetivo último de esta empresa sería «not merely to unveil the nature of the 'third code' per se, but most importantly, to understand the specific constraints, pressures, and motivations that influence the act of translating and underlie its unique language» (Laviosa 1998a: 1).

Siguiendo a Toury (1995: 29), que considera que las traducciones son hechos de la cultura meta, y, en cierta manera, intentando reivindicar el estatus de los textos traducidos, menospreciados por otras disciplinas, Baker sugirió realizar esta investigación sin tener en cuenta los textos originales, es decir, mediante el uso exclusivo de corpus comparables.

In other words, we need to explore how text produced in relative freedom from an individual script in another language differs from text produced under the normal

conditions which pertain in translation, where a fully developed and coherent text exists in language A and requires recoding in language B. This shift in form may be facilitated by access to comparable corpora. (Baker 1995: 233)

Este interés por los corpus comparables era comprensible, pues, como explica Bernardini (2011: 2), «parallel corpora, i.e. collections of translated texts aligned to their source texts, appeared to perpetuate a source-oriented, prescriptive view of translated texts as almost by definition inferior to their source texts». Sin embargo, como hemos mencionado, es una idea bastante extendida en la actualidad que la inclusión de los textos originales en un estudio en forma de corpus paralelos proporciona importantes ventajas a la investigación (Johansson 1998; Bernardini 2011), ya que permite analizar las posibles fuentes de interferencia del texto original en la lengua traducida y aislar las características que no provienen del proceso mismo de la traducción, sino que dependen del par de lenguas de trabajo.

2.3.2. La controversia en torno al concepto de universal

Casi al mismo tiempo que se gestaba esta nueva tendencia descriptiva en los Estudios de Traducción, se empezaron a alzar voces que denunciaron la futilidad de la búsqueda de características universales. En el año 1998, en el número de la revista *Meta* que recogía los resultados de los primeros estudios realizados de manera sistemática con métodos basados en corpus electrónicos y herramientas automáticas de análisis, Tymoczko (1998) publicó un artículo que ya reflejaba las numerosas reservas que aún hoy se demuestran hacia este tipo de estudios. En dicho artículo, la autora afirmaba que era inconcebible realizar afirmaciones que afectaran a todas las lenguas y todos los períodos de tiempo. Por otro lado, los resultados de estos estudios empíricos no podrían ser jamás objetivos, ya que las preguntas que han de resolverse mediante la utilización de corpus y los criterios para la compilación de los textos serán siempre formulados por seres humanos, con sus propios intereses, ideas y perspectivas, subjetivas y variables. La autora no cambia de opinión con los años. En un trabajo de 2005 afirma que considera necesario especificar el ámbito de aplicación de las conclusiones de los distintos estudios sobre estas características, dado que no todas pueden ser aplicables a

todo contexto o tipo de traducción: «It follows as well that the field should give up the search for universals» (2005: 1095).

House se vale de argumentos más específicos para defender su opinión de que los universales de traducción ni existen, ni pueden existir. En su opinión, por ejemplo, los universales propuestos no están descritos y operacionalizados claramente y, además, se ha demostrado que la dirección de traducción y el género son variables que ejercen su influencia sobre los resultados de los estudios sobre estos fenómenos.

The unchallenged assumption has been simply that through the technical possibilities corpus methodology has recently afforded translation scholars, universals can be found – in the vein of the empiricist typological approach. However, I see a great difference in the two quests. I want to go on suggesting quite bluntly that the quest for translation universals is in essence futile, i.e. that there are no, and there can be no, translation universals. (House 2008: 11)

Corpas Pastor (2008: 222) realiza un análisis muy completo y complejo de algunos candidatos a universal sobre un corpus de dimensiones considerables y, sin embargo, no halla pruebas que los apoyen al cien por cien y, en ocasiones, los resultados contradicen las hipótesis, por lo que su conclusión es que no cree que existan dichos universales. Del mismo modo, Volansky, Ordan y Wintner (2015), valiéndose en su estudio de análisis estadísticos y una metodología informática bastante sofisticada, concluyen que las declaraciones de universalidad deberían repensarse, puesto que dependen en gran medida del género y el registro de los textos y varían enormemente según la combinación lingüística analizada (de hecho, la característica para la que encuentran más evidencia empírica es la interferencia). Becher, aunque dirige sus argumentos sobre todo contra la hipótesis de la explicitación como característica inherente al proceso traductor, considera que la misma noción de *universal de traducción* es engañosa, pues las características habitualmente consideradas como tales pueden explicarse por otros factores, «general (non-translation-specific) pragmatic features of linguistic communication» (2011: 44). Zanettin (2013: 25) opina también que muchos estudios sobre universales podrían interpretarse más bien como estudios de normas traductoras, limitadas por factores sociales e históricos e influidas por factores relacionados con la lengua, el tipo textual o la dirección de traducción, por lo que sus

conclusiones serían aplicables a contextos concretos, como podría ser la traducción de documentos administrativos al alemán.

2.3.3. *El concepto de universal como instrumento metodológico*

Si bien las razones aportadas por estos autores son válidas, creemos que se refieren a un concepto absoluto de *universal* que no es el que tienen en mente quienes defienden su validez como herramienta de investigación. De hecho, a pesar de todas las dudas que han surgido con respecto a su utilidad o viabilidad como concepto y a las divergencias existentes en cuanto a su denominación, es indudable que el concepto de universal de traducción sigue vigente, si nos fijamos en el gran número de estudios que siguen generándose en torno a esta línea de investigación, y su utilidad como instrumento en la investigación traductológica es innegable²⁰.

Sin negar validez al concepto básico, muchos autores se decantan por otras denominaciones menos «radicales», como leyes (Bernardini y Zanettin 2004), tendencias estadísticas (Eskola 2004), patrones regulares de comportamiento (Zanettin 2012) o características de la traducción (*features of translation*) (Olohan 2002, 2004). Toury, por ejemplo, a pesar de haber postulado posibles *leyes* de traducción e incluso, como hemos mencionado, hablar de la utilidad de la búsqueda de posibles universales de traducción (1995: 111), declara posteriormente en numerosas ocasiones que no comparte la idea de que este tipo de regularidades puedan denominarse universales. En los artículos de 2004 que citábamos antes, las describe como *probabilistic explanations* y prefiere denominarlas *leyes* (2004a: 24). A pesar de no llamarlas universales, señala que su valor radica en las posibilidades que nos ofrecen para explicar y comprender qué es lo que nuestro objeto de estudio (la traducción) comparte con otros objetos, para luego poder definir lo que es único en él.

For me, the whole question of translation universals is not one of existence— ‘in the world’, so to speak – but one of explanatory power. I truly believe this is one of the most powerful tools we have had so far for going beyond the individual and the norm-governed, and therefore I will stick to it; at least for the time being. As a tool,

²⁰ Como afirma Laviosa (2011: 22): «the quest of universals is now being pursued in multilingual settings despite ongoing debate on the tenability of the notion as a valid concept for describing and explaining the specificity of translated language».

that is, even though not necessarily under the title of ‘universals’. (...)The reason why I prefer ‘laws’ is not merely because, unlike ‘universals’, this notion has the possibility of exception built into it (which is important from the probabilistic point of view because no probability is ever 1), but mainly because it should always be possible to explain away [seeming] exceptions to a law with the help of another law, operating on another level. (2004b: 29)

Lapshinova-Koltunski (2015: 95) considera más adecuado utilizar los términos *translation features* o *phenomena*, pues considera que para llamarlas universales deberían analizarse en más pares de lenguas y direcciones de traducción, y para llamarlas leyes o regularidades deberían estudiarse otros factores (cognitivos o relacionados con el estatus de las traducciones), lo que no es posible mediante datos textuales como los que ocupan a la autora (un corpus de originales en inglés, sus traducciones al alemán y textos comparables en alemán también). Puurtinen (1998: 524), en la línea de Toury y Zanettin (aunque sin quitar validez a la denominación *universal*) sostiene:

On the one hand, systematic, large-scale, corpus-based research into features common to all translations could lead to the identification of universals of translation. On the other hand, research on a smaller scale, restricted to translations of a specified text type in a given socio-cultural context, and maybe even from a particular source language, can yield information about translational norms ...

Eskola, por su parte (2004: 83), resuelve la cuestión de si las características que nos ocupan son universales o dependen de normas (a las que define como regularidades de comportamiento que operan en contextos socio-culturales específicos y pueden cambiar con el tiempo) afirmando que no deberíamos distinguir entre estas dos posibilidades, sino hablar de leyes de traducción en sentido más amplio, «as both local and global inherent tendencies and regularities pertaining to translation». Por consiguiente, las leyes universales como la simplificación o la explicitación no son necesariamente absolutas, sino «strong statistical tendencies that can be observed widely (showing what translators on the average tend to do and what they do not tend to do)» (2004: 86).

Como apunta esta autora, y también otras como Mauranen (2007: 35) o Laviosa (2004: 35)²¹, creemos que es evidente el hecho de que no es necesario hablar de los universales en su sentido más estricto como algo absoluto, aplicable en toda ocasión y sin excepciones. Si los consideramos simplemente tendencias generales o universales, como las denomina Malmkjær (2007), fenómenos con altas probabilidades de manifestarse, o *probabilistic laws* (Toury 2004a, 2004b), pueden resultarnos un concepto muy útil para comprender y caracterizar tanto el proceso como los productos de la traducción, sobre todo si no nos limitamos al simple cómputo de frecuencias y tenemos siempre en cuenta la necesidad de realizar un análisis cualitativo en profundidad. Bernardini y Ferraresi (2011: 243), por ejemplo, deciden utilizar la denominación de universal a pesar de su naturaleza controvertida, entendiéndola simplemente como un término para referirse a «linguistic patterns repeatedly observed in translated texts that could point at underlying regularities of behavior (procedural or strategic) whose status (social, cognitive etc.) cannot be settled on the sole basis of corpus evidence».

Becher (2010b: 23), sin embargo, hablando una vez más de la explicitación, apunta que no sería legítimo utilizar la denominación universal de manera más amplia, como defendemos aquí, «in the sense that the tendency to explicitate is not characteristic, but typical of translated discourse», pues contradice el uso habitual del término tal como se lo define en lingüística. De hecho, de la lingüística, y de su búsqueda de características comunes a todas las lenguas, es de donde surge este concepto. En principio, los universales lingüísticos son fenómenos presentes en todas las lenguas. Bernardini y Zanettin (2004: 52) los definen como «cognitively-basic, situationally-unconstrained theoretical constructs which lie at the basis of generative/typological approaches to linguistics». Para comprender en qué sentido una tendencia puede considerarse universal, incluimos una cita algo extensa de Malmkjær (2007: 53):

If, at first sight, it seems a little odd to define something that merely occurs typically as “universal”, we should remember that in the case of linguistic phenomena, there is a long tradition of doing so. On Greenberg’s list (1966) of

²¹ «I think that if one regards the notion of universal not as an absolute necessity which can explain translation strategies in every single circumstance, but one considers it a descriptive construct, an open ended working hypothesis, it will continue to unveil, as it has already done, more and more aspects of the nature of translation products and processes and their complex relationships with culture» (Laviosa 2004: 35).

forty-five universals (twenty morphological, eighteen syntactic and seven word order universals), based on the study of thirty languages, we find both absolute universals such as “All languages have pronominal categories involving at least three persons and two numbers” (Universal 42), and a number of universal tendencies, to which there are exceptions, such as for example “In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes (Norwegian has both genitive orders)” (Universal 2). To qualify as a universal tendency, or, as these are also known, as a non-absolute, or statistical universal, the tendency in question must be demonstrably statistically significant, however (see Song, 2001: 6), so in this tradition a universal can be defined as a property “which must at least be true of the majority of the human languages” (Song, 2001: 8).

Queda claro, por tanto, que no nos referimos a los universales en el sentido que les otorga la corriente teórica generativa impulsada por Chomsky. Como explica Malmkjær, en la literatura sobre universales lingüísticos hay universales absolutos y tendencias universales, para las que existen excepciones. Por ello, a pesar de las limitaciones y las connotaciones en cierto modo negativas del término (que, sin embargo, es el más usado y conocido en el campo de los estudios de traducción basados en corpus), estaría justificada su utilización en este trabajo. Baker misma define sus universales como «features which typically occur in translated text rather than original utterances» (1993: 243); *typically*, dice, no *always*. Y, sin embargo, parece abandonar esta polémica denominación en trabajos posteriores. En un artículo de 2007 habla de «distinctive features of translated text», sin utilizar el término *universal* ni una sola vez.

En 2004, Chesterman opinaba que el término, aunque siempre aproximativo, es válido y útil mientras se reserve para hipótesis de verdadera universalidad, dado que impulsa a los investigadores a buscar patrones y generalizaciones al nivel más alto posible (2004b: 43). Sin embargo, en trabajos posteriores (2010, 2014), declara que «the term “universal” has been an unfortunate and misleading choice in this context, as applied to linguistic features of translations» (2014: 85), «as this meant shifting the meaning of a term borrowed from elsewhere» (2010: 44). Considera que quizás resultaría más productivo buscar patrones a un nivel inferior de generalización y realizar afirmaciones más modestas, referidas a determinadas lenguas, géneros o condiciones laborales, por ejemplo. Estas afirmaciones pueden resultar igual de útiles a la hora de acumular conocimiento sobre el fenómeno de la traducción y, en el futuro, poder

formular hipótesis más generales. Incluso pueden ser de utilidad en la rama aplicada de los Estudios de Traducción, mejorando la calidad de las traducciones a través de la difusión entre estudiantes y traductores de los posibles efectos de estas características en las mismas. Consideramos, por tanto, que sería más apropiado referirse a estas características del lenguaje traducido como tendencias o regularidades, aunque en este estudio utilizaremos también la denominación de universal, sobre todo al referirnos a los trabajos de otros autores, dado que es el más ampliamente utilizado y reconocido por quienes se han dedicado a estudiar este tema.

Como afirman Marco y Guzmán (2007: 157): «Only when evidence from many different quarters piles up should we be confident enough to make claims about the – more or less – general validity of a given feature». Si bien es cierto que jamás tendremos acceso a las traducciones de todos los períodos históricos y de todas las lenguas para verificar la aplicabilidad de estas leyes o regularidades al universo entero de la traducción, esto es así en todas las disciplinas científicas y, como cualquiera de ellas, la traducción también tiene derecho a postular ciertas leyes generales (Halverson 2003: 230; Chesterman 2004b: 33) que, con el tiempo y gracias a la realización de un número cada vez mayor de estudios, podrán ir evolucionando, refinándose o incluso remplazándose por otras. Si esto no fuera así, sería imposible profundizar en nuestro conocimiento de la traducción y perfeccionar los planteamientos teóricos de la disciplina.

2.3.4. *Candidatos a universal de traducción*

Los principales candidatos a universales de traducción que recogió Baker en su artículo de 1993 fueron la explicitación, la simplificación, la preferencia por una «gramaticalidad» convencional (que en 1996 rebautizaría como *normalisation* o *conservatism*), la tendencia a evitar repeticiones, la tendencia a exagerar características de la lengua meta y la distribución atípica de determinadas características en textos traducidos con respecto a originales. Posteriormente (en 1996) añadió a la lista el universal denominado *levelling out*. Otros autores han ido aportando también sus propios candidatos. Chesterman (2004a: 6-7) propuso dividir estos candidatos, según la naturaleza de la relación lingüística que intentan describir, en *S-universals* y *T-*

universals. Los primeros describirían las relaciones que se dan entre los textos traducidos y sus correspondientes originales en la lengua de partida, y tendrían que ver con la manera en la que los traductores procesan el texto de partida. Los segundos describirían las relaciones que se establecen entre los textos traducidos y un conjunto de textos comparables a estos en la misma lengua, por lo que estarían relacionados con la manera en la que los traductores procesan la lengua de llegada. Un mismo universal puede estudiarse de las dos maneras, dependiendo de cuál sea nuestro punto de referencia. Por ejemplo, la explicitación podría estudiarse identificando los elementos que están implícitos en el texto original y observando si se explicitan en el texto meta (S-universal), o comparando los niveles de explicitud en textos traducidos y textos originales en la misma lengua con respecto al elemento de estudio que hayamos seleccionado (T-universal).

2.3.4.1. La explicitación

Es el candidato aparentemente más estudiado, si tomamos como referencia los diversos artículos sobre universales publicados desde los años noventa. Baker la define como «an overall tendency to spell things out rather than leave them implicit in translation» (1996: 180) y afirma que, para hallar evidencias de esta característica, podemos analizar fenómenos como la longitud de los textos, la adición de información o la expresión de conexiones implícitas en el texto original. Como explica Konšalová (2007: 18) —cuyas conclusiones sobre el alemán y el checo apoyan esta hipótesis—, Vinay y Dalbernet estrenaron el término en 1958 para nombrar una técnica de traducción, pero Blum-Kulka (1986) fue la primera en considerar que podía ser una característica inherente al proceso de traducción:

This argument can be stated as “*the explicitation hypothesis*”, which postulates an observed cohesive explicitness from SL to TL texts regardless of the increase traceable to the differences between the two linguistic and textual systems involved. It follows that explicitation is viewed here as inherent in the process of translation. (Blum-Kulka 1986: 19)

Esta definición se amplía con Baker para abarcar más indicadores de explicitación. Gumul (2006: 174) resume en una tabla los diferentes elementos que se habían estudiado como pruebas del fenómeno de la explicitación hasta el momento de la publicación de su trabajo. Entre ellos encontramos la adición de conectores, la especificación léxica o la adición de modificadores y adjetivos calificativos. Muchos de estos estudios han aportado resultados que respaldan la teoría.

Podemos citar los ya clásicos de Olohan y Baker (2000), en el que las autoras descubren que, en textos traducidos, la frecuencia de aparición del pronombre relativo *that* opcional en construcciones con los verbos *say* y *tell* es mayor que en textos originales en inglés²², o el de Øverås (1998), que se centra en el nivel de cohesión léxica presente en traducciones de inglés a noruego y viceversa y obtiene resultados que confirman la hipótesis en ambas direcciones de traducción, aunque con más fuerza en la dirección inglés-noruego²³. Pápai (2004), trabajando en la combinación inglés-húngaro, con un corpus paralelo y otro comparable, identifica dieciséis estrategias de explicitación presentes en mayor medida en el corpus de traducciones que en el de originales y afirma que «The ultimate motivation is the translator's conscious or subconscious effort to meet the target readers' expectations». Frankenberg-García (2009) analiza la explicitación en portugués e inglés desde el punto de vista de la longitud textual y encuentra que las traducciones son, en general, más largas, tanto en una dirección de traducción como en la otra. Marco (2012) también encuentra evidencias de explicitación para evitar ambigüedades y repeticiones en su estudio sobre el pronombre inglés *one* como elemento cohesivo y su traducción al catalán.

La hipótesis ha sido probada por lo general en lenguas europeas, pero recientemente han surgido también estudios que aportan datos positivos en condiciones diferentes a las normalmente estudiadas, como el de Jiménez-Crespo (2011), que, a través de un corpus de varios millones de palabras, confirma que los textos digitales localizados al español presentan un grado mayor de explicitación de determinados elementos sintácticos que los originales del mismo género y lengua, o el de Xiao y Dai (2014), que comparan un corpus de chino traducido del inglés con originales escritos en

²² Una de las críticas que se hacen a este estudio es que, al no hacer referencia a las lenguas de partida de las traducciones del TEC, Olohan y Baker no tienen en cuenta la posibilidad de que en estas el uso del conector sea obligatorio, como es el caso del portugués y el español. Es decir, existe la posibilidad de que los casos hallados en el corpus de traducciones no sean realmente casos de explicitación sino simplemente de interferencia de ciertas lenguas (Saldanha 2008: 22; Pym 2008: 15).

²³ Becher (2010a: 4), sin embargo, se pregunta cómo es posible que una característica supuestamente independiente de las diferencias lingüísticas y textuales de las lenguas implicadas pueda confirmarse más en una dirección que en otra.

chino y obtienen pruebas de explicitación a nivel semántico, gramático y lógico (aunque puntualizan que la explicitación es un fenómeno relativo, pues depende también de las preferencias del traductor y de factores socioculturales).

Sin embargo, esta hipótesis ha recibido importantes críticas. Una de las más relevantes es la de Becher (2010a, 2010b y 2011), que considera que debería abandonarse la búsqueda del tipo de explicitación que se propone en la línea de investigación sobre universales, que es la que surgiría del proceso de traducción sin influencia de otros factores, pues, además de no estar del todo claras ni la definición del concepto ni las maneras de operacionalizar la hipótesis, los trabajos anteriores que revisa en sus artículos se han basado en supuestos casos de explicitación inherente que no son tales, sino que, en su opinión, se deben a diversos factores contrastivos léxicos, gramaticales o pragmáticos²⁴. Por otra parte, varios estudios han obtenido resultados que no respaldan esta hipótesis, ya sea porque la contradicen (Cheong 2006; Baumgarten, Meyer y Özçetin 2008) o porque sus resultados no son concluyentes (Puurtinen 2004; Denturck 2012), lo que parece demostrar que ni siquiera tendencias ampliamente extendidas como esta pueden considerarse absolutas.

2.3.4.2. La simplificación

Descrita por Baker (1996: 181) de manera algo tautológica como «the tendency to simplify the language used in translation», la hipótesis de la simplificación había sido estudiada antes por autores como Vanderauwera (1985), pero Laviosa la puso a prueba por primera vez con un corpus de grandes dimensiones, y halló cuatro «patterns of lexical use» tanto en textos narrativos como periodísticos (1997: 315): «lower proportion of lexical (or content) words versus grammatical words, greater use of the high frequency words of English, greater repetition of the 108 most frequent words (or list heads) used in each subcorpus, use of a smaller number of lemmas in the list heads». Sin embargo, la simplificación no es uniforme para todos los parámetros en todos los

²⁴ Becher se basa en la distinción que realiza Klaudy (2009) entre explicitación obligatoria (debida a diferencias entre las estructuras sintácticas y semánticas de las lenguas), opcional (debida a preferencias estilísticas y estrategias de construcción textual), pragmática (debida a diferencias de conocimiento general y cultural) e inherente a la traducción, que es la que nos ocupa.

tipos de textos. Por ejemplo, la longitud media de las frases es mayor en traducciones de prosa narrativa que en originales del mismo género (Laviosa 1998b).

Estos resultados ambiguos se reproducen en estudios posteriores: Mauranen (2000) descubre frecuencias léxicas y tendencias combinatorias atípicas en traducciones²⁵. Xiao y Dai (2014: 26) solo se atreven a apoyar la hipótesis de manera parcial, pues hallan pruebas de simplificación en su trabajo sobre el chino traducido, pero también de complicación, como denominan al fenómeno contrario. Rabadán, Labrador y Ramón (2009: 314) consideran que «the ‘simplification hypothesis’ is at play in reducing the range of options and thus narrowing the inventory available in translated Spanish», dado que los traductores parecen ceñirse a una variedad limitada de recursos idiomáticos para trasladar la función intensificadora y cuantificadora del inglés (que suele expresarse con largas cadenas de premodificadores) y dejan de lado recursos más idiomáticos en español que sí se utilizan en los originales en esta lengua. Jantunen (2004: 103), no obstante, comenta lo siguiente acerca de un estudio suyo previo de 2001 sobre sinonimia y simplificación:

In that study it was reported, contrary to earlier findings, that the range of synonymous words (amplifiers) is not narrower in translations; in some cases the range of synonymous degree modifiers is even wider in translations. Furthermore, it came out that translators do not tend to favour the most frequent synonym(s) at the expense of the other members of a group of synonyms.

Corpas Pastor, Mitkov, Afzal y Pekar (2008: 7) encuentran pruebas de este fenómeno en textos originales en español y traducidos a esta lengua del campo médico y técnico, pero no para todos los parámetros estudiados (por ejemplo, los textos traducidos presentan una menor densidad y riqueza léxica, pero no oraciones más cortas, ni preferencia por el uso de frases simples en lugar de complejas). En su ambicioso estudio de 2008 que hemos mencionado con anterioridad, Corpas Pastor vuelve a obtener resultados que ofrecen una evidencia aplastante en contra de esta hipótesis (2008: 220).

²⁵ En un trabajo posterior, Mauranen (2005: 79) comenta que sus resultados, a pesar de mostrar una tendencia en contra de la simplificación, podrían ser compatibles con los de Laviosa, dado que este fenómeno parece presentar múltiples niveles, y lo que es simple en uno puede ser complejo en otro: «the individual words which participate in the collocations can themselves be even more frequent in translations than nontranslations — but the variety in collocations suggests that word frequency counts alone cannot capture the whole picture of lexical simplification».

Pym (2008: 317), por su parte, menciona los siguientes fenómenos posiblemente indicativos de una tendencia a la simplificación: la utilización de frases más cortas, de estructuras con formas conjugadas en lugar de impersonales, la resolución de ambigüedades, la eliminación de puntuación atípica o la producción de textos con menor densidad léxica y *type-token ratio*. Algunas de estas características pueden ser representativas también de la tendencia a la explicitación y, de hecho, este autor comenta que los cuatro universales principales propuestos por Baker, en realidad, «appear to be saying much the same thing. (...) they all seem to elaborate Toury's law of standardization, without touching his proposed law of interference» (2008: 318); de ahí los resultados a veces contradictorios y solapados.

Paloposki (2001: 282), en la línea de Becher, opina que «features that are interpreted as inherent in translation may have arisen out of totally contingent and external factors». Explica que quizás fenómenos como la simplificación léxica de Laviosa puedan ser resultado de factores cognitivos (p. ej. la carga cognitiva de tener que trabajar con dos lenguas a la vez) o sociales (p. ej. remuneraciones pobres, poco reconocimiento de la profesión o incluso el concepto de traducción prevalente en determinada época, que puede favorecer los usos más normativos y menos creativos). Como vemos, este universal ha recibido también crítica, y los resultados de gran parte de los estudios indican que tiende a manifestarse de un modo más claro en algunos niveles, como el léxico, que en otros.

2.3.4.3. La normalización

También denominada *standardisation* (Toury 1995), *conservatism* (Kruger y Van Rooy 2012: 38) o *conventionalisation* (Mauranen 2007: 40)²⁶, entre otros, según Baker (1996: 183) se define como «a tendency to exaggerate features of the target language and to conform to its typical patterns». Los estudios que cita Baker en su obra son los de Shlesinger (1991), en el que se menciona que los intérpretes tienden a completar frases inacabadas, gramaticalizar oraciones y eliminar titubeos y reinicios por

²⁶ Algunos autores mencionan el término *sanitization* como otro sinónimo. Si bien ambas hipótesis presentan similitudes, Bowker (2001: 350) distingue entre ambas: «Meanwhile Kenny (1998: 515) defines 'sanitization' as the suspected adaptation of a source text reality to make it more palatable for target audiences. Both normalization and sanitization result in deliberately chosen unconventional lexical or syntactic source text features being watered down in translation so that the target text fits in with the conventions of the target language. What was creative and original in the source text becomes humdrum and typical in the target text».

parte del orador, y los de Vanderauwera (1985), Malmkjær (1997) y May (1997)²⁷, que estudian la normalización de la puntuación experimental en traducciones. Algunos de los estudios más conocidos sobre normalización son los de Kenny (1998 y 2001), que ya hemos mencionado en un par de ocasiones, y Olohan (2004), que halló que la variedad de términos de color utilizada en traducciones era menor que en los originales.

Más recientemente, Dayrell (2008: 50) encuentra que los textos traducidos muestran una preferencia mayor que los no traducidos por los patrones léxicos más habituales en la lengua meta, aunque advierte de que algunas de las diferencias halladas podrían ser específicas del par de lenguas con las que trabaja (inglés y portugués). Bernardini y Ferraresi (2011: 241) descubren que los traductores italianos del ámbito de la programación informática utilizan menos préstamos sin adaptar (cuando hay equivalentes alternativos en italiano) que los escritores de textos no traducidos, y que estos últimos crean más verbos con raíces inglesas, aportan nuevos significados a verbos ya existentes en la lengua y preservan marcas fonológicas extranjeras en mayor medida. «Since translators and writers have very similar profiles (both groups are Perl programmers, and in fact some of the writers are also translators), we conclude that the differences observed might be due to a task effect, i.e., that the very act of translating may induce one to take a more conservative, normalizing attitude» (2011: 242). Teng (2012) estudia el uso de las partículas *wa* y *ga* en japonés traducido del chino y japonés original (en contextos en los que ambas serían perfectamente válidas) y concluye que la opción favorecida por los traductores (*wa* presenta una frecuencia de aparición mayor en traducciones que en originales) es la opción que suele considerarse estándar en la lengua de llegada.

Sin embargo, las conclusiones de Mauranen (2000) sobre frecuencias y colocaciones atípicas expuestas en el apartado anterior contradicen esta idea, y también las de Puurtinen (2003), puesto que esta última halla, por un lado, una menor frecuencia de palabras y expresiones coloquiales en su corpus de traducciones de literatura infantil al finés con respecto a textos originales en esa lengua, pero, por otro, que la variedad de verbos declarativos es más extensa en dichas traducciones. Delaere, De Sutter y Plevoets (2012: 220) encuentran que el fenómeno de la normalización no se da independientemente del tipo de texto y la lengua de partida. A pesar de que observan un mayor grado de estandarización en el neerlandés de Bélgica traducido, sus conclusiones

²⁷ Baker cita estas dos últimas referencias como *forthcoming*, pues aún no se habían publicado cuando escribió su artículo de 1996.

apuntan a que ciertos tipos de texto (ficción, no ficción y artículos periodísticos) contienen más lenguaje estándar que otros, como los textos administrativos.

Una vez más, los resultados no son concluyentes. Pym (2008: 320) opina que no podemos esperar que la ley de la estandarización creciente, como la denomina Toury, esté presente en todas las traducciones, dado que depende de factores socioculturales como el estatus de las culturas o lenguas involucradas en el proceso: «It tends to happen in cases where the translation (or translation in general) is of relatively minor importance (less “central”) within the target culture» o, como lo expresa Baker, «the higher the status of the source text and language, the less the tendency to normalise» (1996: 183). Por otra parte, afirma que ambas tendencias propuestas por Toury, la de estandarización y la de interferencia, que mencionaremos más adelante, son estrategias encaminadas a evitar riesgos y dependen de la presencia o ausencia de recompensas para los traductores que estén dispuestos a asumir dichos riesgos: «Translators will tend to avoid risk by standardizing language and/or channeling interference, if and when there are no rewards for them to do otherwise» (Pym 2008: 326).

2.3.4.4. *Levelling out*

También denominada *convergencia*, es el candidato de Baker que menos atención ha recibido hasta el momento. Según esta autora, se refiere al hecho de que supuestamente hallaremos mayor homogeneidad en un conjunto de textos traducidos que en uno de originales en la misma lengua:

It concerns the tendency of translated text to gravitate towards the centre of a continuum. (...) For example, there is some evidence that the individual texts in an English translation corpus are more like each other in terms of features such as lexical density, type-token ratio and mean sentence length than the individual texts in a comparable corpus of original English. (Baker 1996: 184)

Laviosa (1998b: 8) comenta que, a pesar de haber hallado una heterogeneidad menor en su corpus de artículos periodísticos traducidos con respecto a un conjunto de artículos escritos originalmente en inglés, «the analysis of variance in narrative prose does not confirm the greater homogeneity found in newspaper articles in respect to

lexical density». Rosa (2008) obtiene resultados que apoyan las hipótesis de convergencia y de explicitación para su estudio sobre estrategias narrativas en un corpus de ocho versiones del *Oliver Twist* al portugués, pero, dado que estos resultados aparecen correlacionados con la variable «target readership», no considera que puedan llamarse universales. Corpas Pastor, Mitkov, Afzal y Pekar (2008: 7) tampoco pueden validar esta hipótesis según los resultados del análisis de las características léxicas, estilísticas y sintácticas seleccionadas para su estudio.

No obstante, existe evidencia reciente a favor de esta hipótesis. Lapshinova-Koltunski (2015) trabaja con un corpus de traducciones inglés-alemán realizadas por traductores humanos (con distinto grado de experiencia) y automáticos (basados en distintos sistemas de traducción). Lo que encuentra es que el único universal que realmente puede comprobar mediante los parámetros utilizados en su estudio es el de convergencia, pues para los demás, la evidencia no es concluyente.

2.3.4.5. La interferencia

La ley de interferencia, como la denomina Toury, o *shining through*, como también se la conoce (Teich 2003: 61; Dai y Xiao 2011), se define de la siguiente manera:

In translation, phenomena pertaining to the make-up of the source text tend to be transferred to the target text., whether they manifest themselves in the form of *negative transfer* (i.e., deviations from normal, codified practices of the target system), or in the form of *positive transfer* (i.e. greater likelihood of selecting features which do exist and are used in any case). (Toury 1995: 275)

La interferencia no es un universal propuesto por Baker. De hecho, Pym (2008) observa que esta autora ha dividido la ley de estandarización creciente de Toury en sus cuatro propuestas de universal pero ha dejado de lado la segunda ley de este autor, pues iría en contra de su metodología, que, en principio, no incluye los textos de partida (aunque Baker no es tan categórica en su rechazo de los corpus paralelos, como puede verse en uno de sus artículos del año 2000). En opinión de Pym (2008: 311):

This one-sidedness of the “universal” agenda enhances the justification of comparable corpora but poses serious problems for any kind of causal explanation of the findings. It would thus seem advisable to return to the duality of Toury’s exemplary propositions.

Por otra parte, si la interferencia se define como la transferencia de características de texto original al traducido, lo lógico es pensar que se manifestará de manera diferente, como mínimo, según la combinación lingüística que estemos analizando. No obstante, muchos autores han decidido estudiar su naturaleza para determinar si puede considerarse universal o no.

Mauranen (2004) estudia la distribución de las palabras más frecuentes en textos traducidos al finés desde distintas lenguas de partida y concluye que se parecen más entre ellos que los textos pertenecientes al subcorpus de textos finlandeses originales (lo que apoyaría la hipótesis de la convergencia), pero, a la vez, existen diferencias entre los pares de lenguas: «while English SL texts deviate less from Finnish originals than from other translations, translations from Russian show the reverse tendency. This suggests traces of SL specific interference» (2004: 78), por lo que las características específicas de los textos traducidos no pueden adscribirse a una única causa. Corpas Pastor (2008: 221) también se propone comprobar si este fenómeno se da en su corpus de traducciones y descubre que se confirma en mayor medida que los otros que analiza en la misma obra. Las traducciones presentan interferencia léxica (más cognados en traducciones) y sintáctica (la sintaxis de los textos traducidos se asemeja más a las de los textos no traducidos en inglés), aunque no pronominal (dado que halla más pronombres cero en traducciones). Xiao y Dai (2014) observan que en traducciones del inglés al chino se aprecian diversos ejemplos de interferencia de la lengua de partida, entre ellos, la utilización de puntos en lugar de comas para separar segmentos, una longitud media de palabras más elevada, o el mayor uso de prefijos y sufijos y de oraciones pasivas (aunque aquí también hay diferencias por géneros; por ejemplo, las pasivas son más frecuentes en las historias policíacas escritas en chino originalmente).

Por último, podemos mencionar a Rabadán, Labrador y Ramón (2009), que estudian la interferencia como una característica que es posible apreciar en los textos traducidos, que permite identificarlos como tales y que puede ser utilizada como herramienta para evaluar la calidad de las traducciones.

the overuse of pre-modifying adjectives with a descriptive function in translated Spanish can be considered as symptomatic of interference when the SL is English. This particular feature is actually highly characteristic of Spanish translations from English, making them easily identifiable as such. (2009: 317)

2.3.4.6. Otros candidatos

Entre otras características de los textos traducidos que han observado diversos autores y se han ido proponiendo como hipótesis de estudio podemos mencionar las siguientes:

➤ Presencia de colocaciones o patrones lexicogramaticales atípicos en traducciones: ya hemos mencionado algún estudio que se hace eco de esta tendencia en textos traducidos, como el de Mauranen (2000) sobre expresiones metatextuales. Jantunen (2004) estudia tres modificadores de grado finlandeses de significado similar y sus conclusiones no apoyan la hipótesis por completo, dado que las colocaciones de los textos traducidos difieren de las de los textos escritos originalmente en finés, pero la lengua de partida ejerce influencia sobre estas diferencias. Balaskó (2008) también encuentra diferencias en los patrones de uso de un elemento léxico de significado metatextual en un corpus comparable de húngaro. Dayrell (2007) halla menor diversidad de patrones colocacionales en los textos traducidos de un corpus de traducciones y originales en portugués brasileño. No obstante, sus resultados podrían explicarse según la hipótesis de la simplificación o de la normalización, que en muchos casos confluyen.

➤ Reducción de la repetición: propuesta por Toury en 1991 (según Chesterman 2010: 41) y observada por Mozetič (2014: 104) en la práctica de la traducción literaria en Eslovenia. Jääskeläinen (2004) también la observa en las traducciones de sus estudiantes, a menos que se les conciencie sobre la importancia de la misma como figura de estilo del original.

➤ Sobrerrepresentación de determinados elementos típicos de la lengua meta en textos traducidos: propuesta por Baker como la tendencia a exagerar

características de la lengua meta en el lenguaje traducido, fue observada por Nilsson (2002 y 2004) con respecto a oraciones relativas coordinadas en sueco y a la palabra gramatical *av* y patrones colocacionales relacionados, aunque según los resultados de su estudio, esta sobrerrepresentación se debe en gran parte a la presencia de estructuras similares en la lengua de partida. Ramón y Labrador (2008) encuentran una sobrerrepresentación en español traducido de ciertos adverbios terminados en *-mente* y la relacionan también con la influencia del texto original. En estos trabajos el fenómeno de interferencia parece estar muy relacionado con el de la sobrerrepresentación.

Como puede deducirse de los ejemplos aportados de los distintos universales, estas características no siempre se encuentran aisladas en los textos. De hecho, el fenómeno contrario es el más habitual. Según Laviosa (2011: 16), «The co-occurrence of regularities and instances of counter-examples to prevailing patterns in the investigation of universals» es una característica de su estudio sobre la simplificación y de muchos otros, como el de de Øverås (1998), «(...) which suggests that the universals of translation may present distinctive features but may also overlap with one another to some extent». Además, Baker observa que «the same feature may be expressed in different ways on the surface, and the same surface expression may point to different features or tendencies» (1996: 181), fenómeno que vuelve a apuntar en un artículo posterior (2007: 12). Pym ilustra la confusión que puede acarrear la falta de definición precisa de estas características de la siguiente manera:

implication can also be seen as a form of simplification (cf. Pym 2005) when it reduces the complexity of syntactic structures. For Blum-Kulka and Levenston (1983: 119), lexical simplification was thus “making do with less words” (cf. Laviosa in this volume). Explicitation would then be a kind of simplification that reduces complexity by spreading semantic load over a wider stretches of text. (2008: 318)

2.4. La hipótesis de los elementos únicos (Unique Items Hypothesis)

Como acabamos de comprobar, a lo largo de los años se han ido realizando numerosas investigaciones con el objetivo de validar o refutar estas hipótesis siguiendo las propuestas metodológicas de Baker. Los resultados de buena parte de ellas han resultado contradictorios o no concluyentes, y algunas características han recibido más atención que otras; por ejemplo, la explicitación y la normalización son los candidatos que, hasta el momento, han recibido mayor apoyo empírico. También se han propuesto nuevos candidatos a tendencias universales. En la sección anterior hemos omitido la mención de uno que consideramos especialmente interesante y que hemos decidido convertir en objeto de análisis de nuestro trabajo: la *Unique Items Hypothesis*.

2.4.1. Estudios que respaldan la UIH

Esta hipótesis fue propuesta por Tirkkonen-Condit en 2002, inspirada, como ella misma relata, en un trabajo de Katharina Reiss de 1971 en el que esta autora sugería que las traducciones no explotan al máximo los recursos lingüísticos de la lengua meta. Reiss proponía un test para comprobar el nivel de conocimientos lingüísticos de un traductor: analizar en qué medida aparecen en sus traducciones las palabras más frecuentes de la lengua meta que no existen en la lengua de partida. Ese año, Tirkkonen-Condit publicó un estudio en el que se proponía comprobar si los lectores eran capaces de distinguir una traducción de un texto original en la misma lengua y averiguar qué características lingüísticas compartían los textos señalados como pertenecientes a uno u otro grupo. Los resultados indicaron que la probabilidad de discernir entre traducciones y originales correctamente no era muy elevada (apenas más alta que por azar). Sin embargo, los comentarios de los participantes arrojaron una imagen bastante clara de las características que suelen atribuirse a ambos grupos: los textos que la gente identificaba como originales eran fluidos, naturales e idiomáticos, mientras que las traducciones debían ser los textos con rasgos opuestos.

A pesar de que, como hemos mencionado, los participantes no fueron capaces de distinguir correctamente un texto traducido de un original, Tirkkonen-Condit extrajo la siguiente conclusión:

Unique items seem to be very good triggers of impressions of original writing. The texts which were labelled translations by an overwhelming majority of subjects in fact manifested very few if any Finnish vocabulary items or idioms which can be described as unique items. Conversely, texts which were thought to represent original writing by an overwhelming majority of subjects had a conspicuous density of these items. (2002b: 214)

Estas conclusiones evidenciaron el interés de estudiar estos elementos y averiguar si en efecto presentaban frecuencias inferiores en textos traducidos. La Unique Items Hypothesis fue avalada por datos empíricos en el artículo de Tirkkonen-Condit «Unique items – over or under-represented in translated language?», publicado como parte del libro *Translated Universals: Do They Exist?* (2004) y la hipótesis que fundamenta este trabajo de investigación se basa en las conclusiones de dicho artículo. En él, la autora explica que ha encontrado evidencia de la tendencia inversa a la llamada por Baker *sobrerrepresentación* —«a general tendency to exaggerate features of the target language» (Baker 1993:244)—. Los elementos en los que había centrado su análisis, características únicas del finés, es decir, elementos lingüísticos sin equivalentes directos en la lengua de partida, eran menos frecuentes en la lengua traducida que en el finés de textos originales.

El estudio de Tirkkonen-Condit se llevó a cabo con textos de género académico, por un lado, y de ficción, por otro, traducidos del inglés al finés. Se estudiaron doce verbos modales finlandeses que codifican los dominios léxicos de la suficiencia y la posibilidad, «distinguishing Finnish from many other languages that do not encode these meanings in a verbal structure», como explica Halverson (2003: 237)²⁸ y dos partículas clíticas (*kin* y *hAn*) cuya traducción es extremadamente dependiente del contexto²⁹. La hipótesis resultó confirmada por el análisis comparativo de las

²⁸ Tirkkonen-Condit (2005: 178) nombra, por ejemplo, los verbos *jaksaa* y *maltaa*, que se traducen aproximadamente como *be strong enough* y *be patient enough*, y para los que el inglés no posee equivalentes lexicalizados. En este idioma suelen utilizarse expresiones perifrásticas como *have enough patience* o las traducciones que hemos proporcionado antes.

²⁹ «The translation of the particle *-kin* depends on its pragmatic function, and in different contexts it translates differently, e.g. with the connectors *also*, *but*, *in contrast*, *consequently*, *thus*. The clitic particle *-hAn* is also

frecuencias de aparición de estos elementos únicos, aunque menos rotundamente en el corpus de textos académicos que en el de ficción. Por otro lado, la autora encontró que estos verbos mostraban diferencias en cuanto a comportamiento sintáctico, semántico y colocacional en textos originales y traducidos.

Las conclusiones de este estudio han sido confirmadas por los trabajos de otros investigadores. Por ejemplo, Mauranen (2000) obtiene resultados que apoyan la teoría de los elementos únicos de Tirkkonen-Condit. Mauranen centró su análisis en un conjunto de conectores, verbos y sustantivos finlandeses con funciones metatextuales. Los conectores no arrojaron diferencias importantes de frecuencias entre el subcorpus de traducciones y el de originales finlandeses. Sin embargo, un conector presentó el doble de apariciones en originales que en traducciones: *toisaalta*, que puede corresponderse en cierta medida tanto al *on the one hand* como al *on the other hand* ingleses, pero no tiene un equivalente directo o exacto³⁰ en dicho idioma, además de ser más frecuentemente omitido en traducciones del finés que añadido en traducciones a este idioma. «To put it simply, this suggests that translators tend to translate that which is there in the source text, with perhaps relatively little attention to discourse-level domestication» (Mauranen 2000: 137).

Eskola (2004) encontró que las traducciones de inglés y ruso al finés tienden a sobrerrepresentar características que cuentan con un equivalente directo en la lengua de partida pero, como Tirkkonen-Condit, también halló el fenómeno opuesto, es decir, que las traducciones tienden a infrarrepresentar características lingüísticas únicas y específicas de la lengua meta. Esto la llevó a sugerir que podría tratarse de una ley de naturaleza universal. Dado que las hipótesis sobre tendencias universales solo pueden verificarse a partir de la realización de numerosos estudios llevados a cabo con distintos pares de lenguas y elementos lingüísticos, Eskola respaldó la posible naturaleza universal de la infrarrepresentación, como apunta la misma autora (2004: 96), porque su estudio se centró en estructuras sintácticas (tres construcciones con formas verbales impersonales típicas del finés), a diferencia de los trabajos anteriores, en los que se había analizado elementos léxicos.

multifunctional, and it usually conveys the assumption of shared knowledge along the same lines as the particle *you know* in spoken English» (Tirkkonen-Condit 2004: 178).

³⁰ En este trabajo definimos «equivalente directo/inmediato/exacto» como un equivalente que se asemeja al elemento de la lengua original estructural o formalmente. A pesar de que *on the one hand* y *on the other hand* puedan considerarse en cierto modo equivalentes, Mauranen (2000: 126) sostiene que tiene un significado más vago y general que se solapa solo parcialmente con el de estos elementos.

En un curioso experimento con alumnos de traducción, Kujamäki (2004) creó un texto en finés sobre la conducción en condiciones invernales en el que incluyó algunos elementos considerados únicos del finés (palabras relacionadas con la nieve). Hizo traducir este texto al alemán y al inglés y, posteriormente, lo propuso en clase como encargo de traducción a la lengua nativa de los estudiantes, el finés. Las traducciones de los estudiantes mostraron una clara tendencia a no utilizar los elementos únicos de su lengua que habrían sido los más adecuados en ese contexto³¹, dejándose influir por las estructuras del original.

En 2005, Tirkkonen-Condit volvió a estudiar la partícula *-kin*, esta vez en traducciones al finés de cinco géneros distintos. Los resultados confirmaron una vez más la infrarrepresentación de dicho elemento y, además, que la influencia del género en las frecuencias de uso disminuye en textos traducidos con respecto a los originales; es decir, que hay menos variación entre los textos traducidos, como predice la hipótesis de convergencia. También se han llevado a cabo, aunque en menor medida, estudios basados en otras lenguas no directamente emparentadas con el finés. En un trabajo realizado en 1996 (citado en Mauranen 2007: 42), Gellerstam estudió distintos elementos en novelas escritas en inglés y traducidas al sueco y novelas escritas originalmente en sueco, y llegó a conclusiones muy similares con respecto a ciertas partículas discursivas específicas del sueco. Malmkjær (2007: 56) también menciona el estudio de Jakobsen de 1986 sobre el pronombre *man* y la partícula modal *jo* en originales daneses y traducciones del inglés al danés.

2.4.2. La UIH como candidato a universal de traducción

Los resultados de estos estudios parecen indicar que ciertos elementos lingüísticos típicos de la lengua meta y sin equivalentes evidentes en la lengua origen tienden a utilizarse en menor medida y de manera distinta en las traducciones. Tirkkonen-Condit, como Eskola, propuso estudiar esta infrarrepresentación en textos traducidos de elementos únicos de la lengua meta como candidato a tendencia universal de traducción.

³¹ Un hecho que también se comprobó empíricamente mediante un test posterior en el que se pidió a otro grupo de estudiantes que proporcionaran las tres palabras que mejor encajasen con el contexto de la historia que se les contaba (la misma del texto sobre conducción).

I have called this the unique items hypothesis, and I would like to test the hypothesis from comparative corpora of various languages, if these were available. (...) If it gets validated across languages, there is a good reason to suspect that there is a filtering element in the translation process which directs the translator's mind to those linguistic items in the target language that *do* have linguistic counterparts. This filter blinds the translator so that s/he tends to overlook the unique linguistic items. (Tirkkonen-Condit 2002: 16)

Tirkkonen-Condit argumentaba que este fenómeno podía deberse a factores intrínsecos del proceso de traducción. Al no tener equivalente directo en la lengua de partida, en el sentido de que no se representan de la misma manera en otros idiomas, no existe un estímulo en el texto original que sugiera al traductor la utilización de esos elementos en su traducción. El par de elementos no existe en el «diccionario mental bilingüe» del traductor (Tirkkonen-Condit 2004: 183)³², por lo que no es probable que dichos elementos del finés sean escogidos en las traducciones, aunque probablemente sí aparecerían en textos originales producidos en situaciones normales. Por lo tanto, las traducciones tenderían a manifestar estos elementos con menor frecuencia que los textos originales.

A pesar de la distinción realizada por Malmkjær entre universales absolutos y estadísticos (2007: 53), esta autora considera que los únicos universales que deberían considerarse como tales son los que pueden recibir una explicación basada en factores cognitivos. La opinión de que la causa última de la existencia de características comunes a todos los textos traducidos es de naturaleza cognitiva es compartida también por Chesterman (2004b: 44) y Olohan (2002). El fenómeno que nos ocupa es un buen candidato a universal, por tanto, puesto que, como hemos argumentado, la explicación parece residir en los procesos mentales del traductor:

It seems to me that of the candidates for universal-hood proposed by Baker (1993), listed by Chesterman (2004) and discussed by Mauranen and Kujamäki (2004), very few qualify for the status of *cognitively determined* universals. One that does qualify, though, is identified by Tirkkonen-Condit (2004). (...) The phenomenon receives a cognitive explanation, and similar results have been found for unrelated

³² Otra denominación utilizada por autores como Halverson (2010) o Lanstyák y Heltai (2012) para este concepto es *mental lexicon*.

languages, Swedish and Danish on the one hand and Finnish on the other.
(Malmkjær 2007: 56)³³

Cuando hace unos años nos planteamos comenzar esta investigación, los estudios relacionados con el fenómeno de la infrarrepresentación que se habían llevado a cabo hasta el momento habían sido principalmente con el finés como lengua meta, pero también con el sueco y el danés, lenguas sin parentesco directo con el primero. Esto demuestra que no hablamos de un rasgo específico de un único par de lenguas. Por lo tanto, y dado que Baker (1993:245) sostiene que, para averiguar si una característica es o no universal, es necesario realizar distintos estudios empíricos con diferentes combinaciones lingüísticas, consideramos que sería muy interesante comprobar si un estudio realizado sobre una lengua tan alejada del finés y el sueco como el español arrojaría resultados similares y creímos que estaba más que justificada la necesidad de investigar este candidato a universal en español y descubrir si la infrarrepresentación es una característica que se encuentra en la traducción sin importar los factores específicos a los que se encuentre sometido el proceso traductor. Si las conclusiones de un estudio con el par de lenguas inglés-español apoyaran la hipótesis de Tirkkonen-Condit, se estaría añadiendo un elemento más para comprobar si este fenómeno constituye una verdadera tendencia universal, o al menos con altas probabilidades de manifestarse, lo que puede resultar útil para mejorar nuestra comprensión del proceso traductor y también tener consecuencias en la didáctica de la traducción.

Por consiguiente, podríamos formular la hipótesis de este trabajo en los siguientes términos condicionales o probabilísticos requeridos por toda ley relacional de la traducción, como propugna Toury (1995: 65), que coinciden con el enunciado del presunto universal del que hemos hablado:

En los textos traducidos del inglés al español se produce un fenómeno de infrarrepresentación de los elementos únicos o típicos de la lengua de llegada cuando no tienen un estímulo lingüístico directo en la lengua de partida.

³³ Malmkjær termina su artículo con la siguiente afirmación: «Many – possibly most – other candidates for universal status would be better accounted for by the norm concept [cuyas causas son de tipo sociolingüístico], which therefore remains to do its job relatively undisturbed within Descriptive Translation studies».

2.4.3. *Argumentos en contra de la UIH como candidato a universal de traducción*

Sin embargo, esta hipótesis también ha recibido críticas. House (2008: 12-14) afirma que en el seno de su proyecto *Covert Translation* han descubierto que la utilización de elementos como las partículas modales en textos científicos divulgativos ha aumentado en los últimos veinticinco años en las traducciones al alemán debido probablemente a la influencia del inglés en el género, y es mayor que la presente en originales comparables. Considera que esto invalida la universalidad de la hipótesis de la infrarrepresentación, que, en su opinión, es un fenómeno normal e incluso quizás necesario, pero dependiente del par de lenguas en cuestión y, por tanto, no universal. Según House, la Unique Items Hypothesis no es más que

a reformulation of interference from the source language, often as a direct outcome of insuperable grammatical differences between the source and target language systems. For example, when the source language simply does not encode a feature like a certain tense marking, it will be nearly impossible for the translator to reach a target language-conform frequency and distribution of this feature in translation. (2008: 14)

Está claro que, por un lado, la existencia de estudios que respaldan la hipótesis de la sobrerrepresentación de elementos típicos de la lengua meta en traducciones reduce el alcance de cualquier reivindicación de universalidad con respecto a la hipótesis de la infrarrepresentación. Por otro lado, al depender la selección de los elementos únicos del contraste entre las lenguas analizadas, es casi imposible que la infrarrepresentación de dichos elementos pueda ser un fenómeno generalizable más allá de la combinación lingüística en cuestión. Pero creemos que lo que los distintos autores citados consideran en cierta manera universal, o más bien, generalizable, es el hecho, bastante llamativo en nuestra opinión, de que el traductor no llegue a acceder a estos elementos en su lexicón mental, a pensar en ellos; que no los utilice simplemente porque el texto de partida no se los sugiere. Esto es más comprensible cuando hablamos de elementos léxicos o conceptos inexistentes en la cultura de partida que, por tanto, es difícil que pudieran aparecer en un texto traducido desde esa cultura o lengua. Pero cuando hablamos de elementos como los que estudiamos en este trabajo, que existen en ambos polos y simplemente se manifiestan de manera distinta, también deberían

desempeñar un papel importante las normas de la lengua meta y el conocimiento del traductor sobre los medios más idóneos para expresar estos significados (de hecho, como veremos, muchas veces sucede justamente esto y no se cumplen las predicciones de la Unique Items Hypothesis). Por este motivo no creemos que la infrarrepresentación de la que hablamos sea un fenómeno tan evidente como parece dar a entender House y consideramos que la Unique Items Hypothesis es, de hecho, una hipótesis general que podría llegar a demostrarse en todas las combinaciones lingüísticas.

Zanettin (2013: 25), como House, considera que esta hipótesis puede verse simplemente como «an indirect form of SL interference, which affects the TL by subtraction, that is not in terms of direct transfer or ‘shining-through’, but by causing under-use of TL features». No es el único. Pym (2008: 320) menciona que Laviosa considera la Unique Items Hypothesis como un caso particular de interferencia **negativa**. Lanstyák y Heltai (2012: 106), por su parte, la ven como un subtipo de la ley de interferencia pero, en su caso, como «a side-effect of **positive** transfer: features or items that do not appear in the ST and cannot be transferred, will not appear in the TT either»³⁴.

Malmkjær (2007: 57) también opina que este fenómeno depende de la relación entre las lenguas implicadas, pero aporta una nota positiva al sugerir que esta interferencia de la que hablamos puede venir determinada no únicamente por el polo de partida, como es lo habitual, sino también por el de llegada: «Differences between translations into *L* and texts originally in *L* are determined as much, if not wholly, by *L*'s unique features, rather than features of the language of the original for the translation». A fin de cuentas, la comparación se realiza finalmente entre los textos traducidos y los producidos de manera «espontánea» en la misma lengua, no los originales en la lengua de partida: la UIH sigue siendo un *T-universal*, si usamos la terminología de Chesterman. Por esta razón, es posible estudiarlo únicamente a través de corpus comparables (y de hecho, se hace así en la mayor parte de los estudios): «The trick, of course, is that the comparative linguistics has to be done prior to the corpus work, in the selection of the “unique” items to be tested» (Pym 2008: 320).

³⁴ Recordemos que la interferencia negativa implicaba desviaciones de las prácticas normales del sistema meta, y la positiva, mayor probabilidad de seleccionar características que existen y se utilizan en ella (Toury 1995: 275).

2.4.4. *Definición y criterios de selección de elementos únicos*

Con respecto a los criterios de selección de los *unique items*, en primer lugar es preciso aclarar que los elementos de los que hablamos no tienen por qué ser intraducibles ni carecer de equivalentes. Basta con que no tengan un equivalente directo o fácilmente identificable porque, por ejemplo, entre ellos y los elementos correspondientes en el texto de partida no existe identidad formal.

Every language has linguistic elements that are unique in the sense that they lack straightforward linguistic counterparts in other languages. These elements may be lexical, phrasal, syntactic or textual, and they need not be in any sense untranslatable; they are simply not similarly manifested (e.g. lexicalized) in other languages. (Tirkkonen-Condit 2004: 177)

En este sentido, estamos de acuerdo con la opinión de Chesterman (2007: 5) acerca de que el concepto de *uniqueness* debe ser adoptado en un sentido relativo, es decir, que la unicidad no es una propiedad absoluta sino que solo puede entenderse al relacionar o contrastar dos o más elementos. Cuanto menos similares sean dos elementos equivalentes en dos idiomas, mayor será el grado de unicidad que podrá asignarse al elemento meta estudiado. ¿Y a qué nos referimos cuando decimos similar? Nos referimos a similitud formal o estructural, como la utilización de una categoría lingüística similar para expresar un mismo significado. Se podría decir que, una vez hemos definido en qué categoría lingüística clasificar un determinado elemento del español, podemos determinar si en inglés se utiliza otro tipo de categoría lingüística para expresar lo mismo. Por ejemplo, no habría equivalencia a nivel formal o gramatical cuando, para expresar un mismo significado, el español utiliza una construcción impersonal como «haber que + infinitivo» mientras que el inglés normalmente requiere un sujeto que realice la acción expresada por el verbo principal y, por lo tanto, prefiere las formas personales (*one has to + infinitivo* o *we must + infinitivo*)³⁵. La existencia en lengua meta de otras alternativas para estos significados o conceptos más formalmente similares al elemento de la lengua de partida probablemente incrementará el grado de unicidad del elemento en cuestión.

³⁵ En este caso, también existe alguna estructura impersonal equivalente, pero formalmente muy diferente (p. ej. «to be necessary to» o construcciones de pasiva con *it* como sujeto, como «It must be added that...»).

Cappelle (2012: 180-181) extiende la definición de Chesterman (2007: 8) de elemento único —un elemento «for which the translation equivalent only maintains unit correspondence at some higher level or levels, not at given lower levels»— para poder incluir los verbos de movimiento³⁶ en la categoría de elementos únicos del inglés comparado con el francés:

the more grammatical environments there are in which the target-language item can be found in its language without there being a (for all intents and purposes) comparable grammatical environment for its closest translation equivalent in the source language, the ‘more dissimilar’ that element is from its translation equivalent, and hence the ‘more unique’ it is. This reformulation shifts the notion of ‘uniqueness’ from the item itself to the contexts in which it is found: an item is more unique for a given language as its grammatical environments are more unique for that language.

Por otro lado, a pesar de lo que pudiera dar a entender la denominación *único*, estas características o elementos no tienen por qué ser exclusivos de la lengua en cuestión. Pueden darse en otros idiomas, siempre y cuando no se trate de uno de los que conformen la combinación lingüística que estamos analizando. Es decir, son elementos característicos de la lengua meta que no se manifiestan de la misma manera en la lengua de partida, pero es posible que sí lo hagan en otras lenguas. Entonces, según la definición que proporcionan los diccionarios de *único* («solo y sin otro de su especie»³⁷), esta denominación no sería del todo adecuada para los elementos que tratamos. Una posible alternativa sería utilizar el adjetivo *típico* («característico o representativo de un tipo»³⁸; «peculiar de la persona o cosa de que se trata»³⁹), ya que hablamos de formas de expresión que son propias y características de la lengua meta, tengamos o no en cuenta la existencia de los mismos en lenguas distintas. Ya hemos visto que Tirkkonen-Condit (2002b) demostró, mediante un experimento con textos auténticos, que la presencia de este tipo de elementos en un texto tiende a hacer pensar

³⁶ Se refiere específicamente a los *manner-of-motion verbs*, que lexicalizan la manera en que se realiza el movimiento (i.e. *crawl* o *roll*).

³⁷ Definición de *único* proporcionada por el *Diccionario de la lengua española* de la RAE (2001) [consulta: 10 de julio de 2010].

³⁸ Definición de *típico* proporcionada por el *Diccionario de la lengua española* de la RAE (2001) [consulta: 10 de julio de 2010].

³⁹ Definición de *típico* proporcionada por el *Diccionario de uso del español de María Moliner* (2001).

que se trata de un original y no de una traducción (tanto si lo es realmente como si no), ya que los lectores los sienten propios del lenguaje espontáneo y natural en su idioma.

Resumiendo lo que hemos dicho hasta el momento, cuando hablamos de *unique item*, nos referimos a elementos lingüísticos característicos de la lengua meta sin equivalentes directos o estructuras paralelas en la lengua de partida, es decir, que no se representan de la misma manera en uno o más idiomas y, por tanto, es fácil que presenten una frecuencia de aparición inferior en textos traducidos de dichos idiomas; es decir, que estén infrarrepresentados en traducciones⁴⁰. Por supuesto, no es nuestra intención utilizar el término *infrarrepresentación* en sentido normativo. Consideramos que un elemento se encuentra infrarrepresentado cuando no aparece en textos traducidos con la misma frecuencia con la que lo haría en textos originales en la misma lengua. Pero los datos sobre dichas frecuencias «habituales» en estos últimos textos se obtienen de manera empírica, a través del análisis de un corpus de textos escritos originalmente en la lengua de llegada, y desde este punto de vista descriptivo podemos afirmar que un elemento se encuentra infrarrepresentado, por lo que la palabra no implica de ningún modo afirmaciones prescriptivas o normativas acerca de la mayor o menor normalidad de determinadas frecuencias. Leech (1998: XIX y XX), por ejemplo, habla de utilizar los términos *underuse* y *overuse*, en el contexto del aprendizaje de lenguas, sin juicios de valor, «descriptively, as a convenient shorthand for “significantly more / less frequent use (...)”». Lo mismo defiende Johansson (2007), que proporciona algunos ejemplos de *underuse* de determinados elementos sin equivalentes directos en inglés en traducciones al noruego:

It should be stressed that overuse and underuse are meant as descriptive terms and do not necessarily imply that there is anything wrong with translated texts (...)
Translations are an important channel of linguistic and cultural influence, and they may be a source of enrichment of the target language. (2007: 33)

⁴⁰ Las primeras sugerencias que proporcionó Tirkkonen-Condit acerca de qué tipo de elementos podían estudiarse fueron las siguientes: «Unique items in the above sense may appear in lexis, phraseology, syntax, textual organization, collocational patterns, or pragmatic habits, in other words at any linguistic level. Word order in Finnish is used, among other things, to express illocutionary function, whereas English tends to express illocution by means of intonation, whereas German uses a wide range of discourse particles to express illocutionary function. Any experienced translator or language teacher can produce a list of words that are tricky to translate into foreign languages» (2002: 216-217).

2.4.5. Otros estudios relacionados

Además de los ejemplos que ya hemos proporcionado para apoyar los resultados de Tirkkonen-Condit, a lo largo de los años se han publicado otros estudios relacionados con esta hipótesis. Resumiremos algunas de estas aportaciones a continuación.

Un resultado contradictorio lo proporciona la misma Tirkkonen-Condit, en colaboración con Mäkisalo, en 2007. En un corpus de subtítulos finlandeses de millones de palabras, los autores encontraron que los subtituladores explotaban los recursos de la lengua meta considerados únicos, a veces incluso con mayor frecuencia que la hallada en finés original. Esto sucede porque la necesidad de concisión exigida por el formato específico de producción de estos textos empuja a los profesionales a valerse de los recursos más útiles en este sentido, y se utilizan aunque no haya ningún estímulo en el original que pueda sugerirlos. Por lo tanto, existe al menos un contexto de traducción o tipo textual en el que la hipótesis no se ha validado. Sin embargo, ha seguido estudiándose.

Ya hemos mencionado a Cappelle (2012), que estudia los verbos de manera de movimiento del inglés y su presencia en textos traducidos a esa lengua desde el francés y el alemán. Estos verbos se encuentran efectivamente infrarrepresentados en traducciones al inglés del francés, idioma considerado *verb-framed*, dado que codifica el movimiento y su trayectoria en el verbo mismo (y el modo a través de otros recursos como gerundios o frases preposicionales). No ocurre lo mismo en las traducciones del alemán, una lengua *satellite-framed*⁴¹, es decir, que codifica el modo del movimiento de la misma manera que el inglés, en la raíz verbal (y la trayectoria, a través de preposiciones o adverbios, los «satélites»). Cappelle argumenta que sus resultados apoyan la Unique Items Hypothesis, pero no su naturaleza universal, dado que la causa de estas diferencias radica en las diferencias tipológicas existentes entre dos lenguas específicas.

Xiao y Dai (2014), en el mismo estudio en el que ponen a prueba otros de los candidatos de los que hemos estado hablando, encuentran bastante evidencia a favor de la Unique Items Hypothesis. Entre los elementos que presentan una frecuencia menor en traducciones están determinadas frases idiomáticas, marcas de pausa, construcciones

⁴¹ En español también se las denomina «lenguas de marco verbal» y «lenguas de marco satélite».

con 把 (ba), marcadores aspectuales, partículas modales y las formas arcaicas del auxiliar 的 (de), es decir 之 (zhi), y de la forma pasiva 被 (bei), es decir, 为...所 (wei...suo). «It is clear that as far as translated texts in English-to-Chinese translation are concerned, the more characteristic a linguistic feature is of Chinese, the better it can serve to differentiate between native and translational Chinese» (2014: 26). Opinan que no se le ha prestado la debida atención a esta hipótesis, probablemente por haberse originado en un contexto europeo, en el que las lenguas están muy emparentadas entre sí.

Con el objetivo de hallar maneras de evaluar la calidad de las traducciones valiéndose del concepto de universal de traducción, Rabadán, Labrador y Ramón (2009) hallan apoyo empírico para nuestra hipótesis. Comparan las distintas funciones del *simple past* inglés con los del pretérito perfecto simple e imperfecto españoles y, partiendo de estos datos de uso en español original, descubren lo siguiente con respecto a la utilización del imperfecto en traducciones al español:

The results indicate that there is a raw underuse of the imperfect as a translation of the meaning ‘absolute past’. No evidence has been found of this meaning in translated Spanish, which seems to prove the usefulness of the ‘unique grammatical features hypothesis’ discussed earlier as a tool to provide empirical data in order to produce an informed quality assessment report. (2009: 321)

Finalmente, Hareide (2013) estudió la Unique Items Hypothesis, pero no de manera aislada, sino en combinación con la hipótesis de la sobrerrepresentación (en teoría, completamente opuesta a la primera), dado que su objetivo último era probar la hipótesis de Halverson (2003) denominada *Gravitational Pull Hypothesis*. Halverson opina, como Malmkjær, que los universales de traducción tienen una base cognitiva — como lo describe Chesterman (2010: 43): «something in the mind of translators that affects the way they process texts simultaneously in two languages»—. La autora afirma que pueden explicarse a través de la gramática cognitiva de Langacker, dado que surgen «from the existence of asymmetries in the cognitive organization of semantic information» (2003: 197). Los conceptos teóricos que subyacen a la hipótesis de Halverson y que esta explica en profundidad en su artículo de 2003 son complejos, por lo que intentaremos resumirlos de la manera más clara posible.

Las unidades lingüísticas⁴², que constituyen el punto de acceso a nuestras redes semánticas, presentan dos polos: el fonológico y el semántico. En este último es en donde se organiza el conocimiento enciclopédico (fenómenos como forma, color, función, etcétera) en dominios y matrices con distintos grados de centralidad. Este polo, a pesar de encontrarse ligado a una unidad fonológica, puede abarcar una amplia variedad de sentidos relacionados y, por tanto, constituir una categoría compleja estructurada en torno a prototipos, cuyos miembros pueden ser más o menos centrales⁴³, y representada en el nivel más alto de abstracción posible por una red esquemática (*high-level schema*). Estas redes se conectan a su vez con otras que representan significados relacionados o solapados, construyendo nuestra estructura semántica.

Las palabras en las dos lenguas de una persona bilingüe pueden compartir todos, algunos o ningún nodo en el nivel semántico. Es decir, que puede haber configuraciones semánticas que estén ligadas a una representación fonológica en una sola de las dos lenguas (p. ej. conceptos específicos de una de las culturas y quizás también nuestros *unique items*) o redes en las que dos nodos fonológicos compartan todos los semánticos y por tanto no haya problemas a la hora de relacionar los términos en ambas lenguas. Casi todas las redes son asimétricas; es decir, ciertos elementos son más o menos centrales que otros. Cuando activamos una red — accedemos a ella mediante una unidad lingüística del texto original, al intentar traducirla, por ejemplo—, las dos estructuras más prominentes, los prototipos de la categoría y los *high-level schema*, ejercerán cada uno una fuerza diferente que constituirá la fuerza gravitacional (*gravitational pull*), como la denomina Halverson, que hará que seleccionemos estos nodos más prominentes y, por tanto, su correspondiente elemento en lengua meta.

Given a schematic network with a prototype or schema, the linguistic item linked to that prototype or schema will be overrepresented in translated, as opposed to non-translated text. For instance, for lexical items with a prototypical sense, this sense will be overrepresented in translated text as compared to non-translated text. Or, for categories of words with related senses, the most prototypical or schematic word will be overrepresented in translated, as opposed to non-translated text. (Halverson 2009: 91)

⁴² Halverson explica que en la lingüística de Langacker, estructuras gramaticales y léxicas no son cosas distintas, sino que difieren únicamente en grado, por lo que pueden ser explicadas por las mismas estructuras y procesos cognitivos.

⁴³ Lo mismo sucede con el polo fonológico: las distintas realizaciones de una unidad fonológica representan miembros más o menos centrales de una categoría.

En redes donde no exista un prototipo o esquema claro, no habrá sobrerrepresentación, sino una diversidad de soluciones de traducción. El estatus de prototipo depende de la frecuencia de uso y garantiza que dicho prototipo se seleccione más frecuentemente que elementos más periféricos. En principio, las simplificaciones provendrían de la fuerza ejercida por el prototipo de una determinada categoría, y las generalizaciones, de la fuerza ejercida por los esquemas de alto nivel, aunque Halverson advierte de que otros factores del cotexto o contexto podrían anular estas fuerzas. En la siguiente cita, Halverson comenta la hipótesis de Tirkkonen-Condit:

in translation situations where there is no conceptual overlap between the nodes associated with a given TL structure and any specific SL structure, there is underrepresentation of that TL structure relative to non-translated text. Her case is, of course, the extreme, as there is no conceptual overlap at all. To put this in another way, the networks accessed by various SL structures are only very distantly and weakly connected, if at all, to the networks for the TL structures in question here. Thus there is no cognitive salience in immediately accessed networks acting as a force in the translation of such items. (2003: 223)

Entonces, Halverson acepta la existencia de estos dos fenómenos opuestos: la sobrerrepresentación y la infrarrepresentación de un determinado elemento en textos traducidos, dependiendo de la estructura de la red a la que se acceda en cada caso. Hareide se propone averiguar cuáles son las circunstancias específicas que provocan estos dos tipos de relaciones opuestas entre las lenguas y para ello pone a prueba la Gravitational Pull Hypothesis de Halverson mediante la comparación del uso del gerundio español y del elemento prototípico «estar + gerundio»⁴⁴ en textos traducidos del inglés (que posee equivalentes para estas estructuras) y del noruego (que no las posee) y textos escritos originalmente en español. Los tres factores identificados por Halverson que pueden dar lugar a una infrarrepresentación o a una sobrerrepresentación son los siguientes, según Hareide (2013: 35-40):

⁴⁴ Hareide (2013: 37) basa el estatus de prototipicidad de «estar + gerundio» en la descripción de dicha estructura incluida en la gramática de la RAE (2009: 2186).

- ✓ Patrones de prototipicidad en la lengua meta: si un elemento de la lengua meta es prototípico o se utiliza con frecuencia, ejercerá una fuerza gravitatoria que dará lugar a su sobrerrepresentación.
- ✓ Estructuras conceptuales o la representación del concepto en la lengua origen: un elemento prominente (*salient*) o prototípico en la lengua de partida puede ejercer también su influencia y originar una sobrerrepresentación.
- ✓ Diversos patrones de conexión en el lexicón mental del traductor: un grado de relación, es decir de distancia o de solapamiento entre conceptos relacionados de las dos lenguas en el lexicón mental del traductor puede provocar patrones de sobre o infrarrepresentación, dependiendo de la presencia o ausencia de enlaces directos entre ciertas configuraciones de red y elementos léxicos en las lenguas del bilingüe. En los casos en los que no se perciben solapamientos o enlaces directos, no hay elementos de la lengua origen que activen el elemento en cuestión, lo que puede originar infrarrepresentación.

Los resultados del estudio apoyan la hipótesis de la sobrerrepresentación del gerundio y de la estructura «estar + gerundio» en traducciones en todos los casos, incluso, sorprendentemente, en el par noruego-español. Esto lleva a Hareide a afirmar (2013: 255) que la fuerza gravitatoria ejercida por esta estructura prototípica del español es más fuerte que la ejercida por la falta de estructuras correspondientes en la lengua de partida. Hareide se pregunta también qué condiciones ha de cumplir un elemento único para contrarrestar esa fuerza y validar la hipótesis de Tirkkonen-Condit: ¿Presentar una frecuencia baja, no ser prototípico, no poseer una correspondencia prominente en la lengua de partida, o conexión con ningún concepto relacionado en el diccionario bilingüe del traductor...? Hareide opina que el concepto de elemento único y la Unique Items Hypothesis deben refinarse y plantea su propia formulación de la hipótesis:

A unique item is a target language feature that has no direct counterpart in the source language, and has no salient or prototypical source language structure that may impact its choice in the target language. It is characterized by being infrequent and non-prototypical. (2013: 276)

3. Metodología

3.1. Introducción

Esta investigación se enmarca principalmente dentro de lo que Chesterman (2003) denomina modelo teórico comparativo, dado que parte del análisis de ciertas características lingüísticas de un conjunto de textos traducidos y textos comparables escritos originalmente en lengua meta, para determinar la naturaleza de la relación de similitud o diferencia que existe entre ambos tipos de texto. La metodología utilizada para ello es la propia de los estudios de traducción basados en corpus, es decir, el estudio de las características de la lengua a través de conjuntos de textos compilados en corpus electrónicos de distintos tipos y la ayuda de herramientas automáticas de análisis y extracción de datos. Como corresponde a este tipo de estudios, la investigación que se ha llevado a cabo es eminentemente descriptiva y empírica, pues partimos de una hipótesis de tipo descriptivo —es decir, una afirmación sobre la supuesta generalidad de ciertas características observadas en un determinado fenómeno— y nos basamos en la búsqueda y el análisis de datos con el objetivo de encontrar patrones y regularidades y, a partir de ellos, intentar realizar generalizaciones en mayor o menor medida. El análisis está orientado en principio hacia el texto meta pero sin dejar de lado la aportación de los textos originales, y es principalmente cuantitativo pero combinado con aspectos cualitativos.

No obstante, no nos hemos limitado a extraer resultados numéricos referidos a porcentajes y frecuencias, sino que, además, nos hemos detenido en el estudio detallado de las concordancias, tanto del componente comparable del corpus como del paralelo, para no centrarnos únicamente en el aspecto descriptivo sino también tratar de identificar —aunque sea de manera muy tentativa— relaciones y posibles causas e intentar llegar a conclusiones acerca de tendencias o generalidades observadas. Recordemos que la Unique Items Hypothesis lleva incorporada en sí misma el elemento causal, dado que se asume que este fenómeno es provocado por la falta de estímulo en los originales. El análisis cualitativo es necesario, en primer lugar, para comprobar que

las cifras aportadas por las herramientas automáticas de extracción de datos son correctas, es decir, para asegurarnos de que se tienen en cuenta únicamente los elementos en los que nos interesa centrarnos, realizando un cribado manual de los datos mediante el análisis de las concordancias y el contexto⁴⁵. También es esencial a la hora de estudiar si, en los casos en los que los elementos analizados aparecen en el texto traducido, lo hacen de la misma manera que en el texto escrito originalmente en español; es decir, si se manifiestan de modo similar a como lo harían en el lenguaje espontáneo o si, por el contrario, presentan diferencias sintácticas, semánticas o colocacionales. Por último, el proceso que requiere de un análisis cualitativo más detallado es el de la búsqueda de elementos activadores en las concordancias paralelas. En este caso, el programa nos muestra las frases alineadas, pero depende del investigador extraer conclusiones acerca de posibles equivalencias.

El primer paso de este trabajo es la compilación de un corpus comparable de textos formado, por un lado, por obras escritas originalmente en español y, por otro, por textos traducidos de inglés a español. Este componente comparable nos permitirá comparar las frecuencias y formas de uso de los elementos analizados en español original (o escrito de manera «espontánea») y español traducido, y determinar similitudes y diferencias de comportamiento. Por otra parte, hemos considerado necesario complementar este corpus comparable con un componente paralelo conformado por los textos de partida en inglés de las traducciones seleccionadas con anterioridad, para así controlar los efectos de los textos originales en la lengua traducida. Otra ventaja de esta metodología es que nos permitirá recopilar un inventario de equivalentes para los elementos que analizaremos, que pueden servir como recursos de traducción a estudiantes y profesionales (Ramón y Labrador 2008: 280, 289; Labrador 2007: 30). El proceso de diseño, creación y modificación del corpus se explicará con detalle en los apartados 1.2 y 1.3.

Una vez compilado el corpus, el segundo paso es hallar determinados elementos del español que puedan considerarse únicos o típicos de esta lengua, al menos si la comparamos con la lengua de partida, que en nuestro caso es el inglés. Para identificarlos es necesario revisar diccionarios, gramáticas monolingües y bilingües y estudios contrastivos (Chesterman 2007). A veces nos encontramos con la afirmación

⁴⁵ Por ejemplo, en el capítulo de las perífrasis, este cribado sería necesario para distinguir los casos en los que *suelo* es sustantivo de aquellos en los que funciona como auxiliar, que son los que nos interesan, pero también para distinguir, dentro de estos últimos, los casos en los que los distintos auxiliares constituyen verdadera perífrasis.

de algún experto acerca del especial desarrollo de determinada característica en español en contraposición con otras lenguas y ese es nuestro punto de partida para estudiar dicho elemento y comprobar si esta afirmación se corresponde con la realidad o es respaldada por otros autores. Las fuentes de las que extrajimos esta información se detallarán en los capítulos respectivos. Los elementos lingüísticos del español que finalmente serán objeto de análisis en nuestro trabajo son los siguientes:

- una selección de 27 perífrasis verbales
- el conjunto de los determinantes, pronombres y adverbios demostrativos⁴⁶
- una selección de 27 sufijos diminutivos, aumentativos y peyorativos

Cada uno de los capítulos del análisis comienza con una introducción teórica que intenta describir el elemento objeto de análisis. Hemos intentado ofrecer una visión panorámica de lo que de estos elementos se ha escrito en lingüística, traducción o estudios contrastivos, pero es evidente que no hemos podido profundizar en exceso y quizás hayamos podido pecar de superficialidad en el análisis, pues a cada uno de estos elementos por separado podría dedicarse una o más tesis doctorales. En el capítulo correspondiente a cada uno de estos elementos, explicamos, además, los motivos por los cuales hemos decidido que pueden considerarse elementos únicos o típicos del español y el interés que puede tener su estudio.

El análisis de los elementos elegidos constará —en líneas generales, pues explicaremos en profundidad el análisis realizado dentro de cada uno de los capítulos correspondientes— de los siguientes pasos:

- Extracción y cribado de concordancias⁴⁷ para obtener todos los casos válidos de utilización del elemento seleccionado (y descartar los que no nos interesan) y determinación de su frecuencia de aparición en el subcorpus de textos escritos originalmente en español.

⁴⁶ Da la casualidad de que, si contamos por separado las formas determinantes y pronominales o las acentuadas y las no acentuadas, como hace Labrador (2011), también suman 27.

⁴⁷ Las concordancias extraídas del subcorpus de originales españoles (ES) estarán, por lo general, marcadas con el color burdeos, mientras que las del subcorpus de textos traducidos al español (TRAD) aparecerán en color azul y las del subcorpus de textos originales ingleses (EN), en verde.

- Extracción y cribado de concordancias y determinación de la frecuencia de aparición de dicho elementos en el subcorpus de textos traducidos del inglés al español.
- Comparación de ambas frecuencias y realización de tests de significación estadística para averiguar si estas frecuencias difieren o no y, en caso afirmativo, si las diferencias son significativas.
- Análisis de ejemplos de uso en contexto para estudiar la manera en la que dichos elementos aparecen en la lengua traducida y averiguar si se manifiestan de distinta forma que en la lengua espontánea.
- Análisis del corpus paralelo para determinar la presencia o ausencia de elementos activadores en los textos originales e intentar averiguar el grado de influencia que pueden haber tenido sobre las decisiones traductoras.

La importancia de nuestro trabajo radica en el hecho de que, si los resultados obtenidos concordaran con los de estudios anteriores sobre el fenómeno de la infrarrepresentación en textos traducidos realizados por otros investigadores, en condiciones diferentes y con distintos elementos de estudio y herramientas de análisis, podríamos añadir un elemento más a la defensa de la naturaleza universal (en el sentido laxo probabilístico que hemos definido en la introducción teórica) de dicho fenómeno. Es evidente que nuestro estudio se basa en un corpus de dimensiones reducidas y tiene en cuenta una única combinación lingüística, por lo que estas conclusiones no podrán generalizarse ni extrapolarse a otros contextos mientras no se lleven a cabo estudios de mayor alcance. Lo único que pretendemos es averiguar, a través de los resultados de un trabajo empírico, si en la combinación lingüística inglés-español también se produce un fenómeno de infrarrepresentación en textos traducidos de los elementos únicos o típicos de la lengua de llegada cuando no tienen un estímulo lingüístico directo en la lengua de partida. Si no fuera así, este estudio contribuirá de todos modos a analizar las relaciones existentes entre textos traducidos y no traducidos y entre textos originales y traducciones y profundizar en nuestro conocimiento de qué es y cómo funciona la traducción.

Para conseguirlo, hemos seguido las líneas de investigación marcadas por los primeros trabajos de traducción basados en corpus que se dedicaron al estudio de los universales de traducción pero, a la vez, hemos tenido en cuenta las críticas y sugerencias de mejora que han ido surgiendo a lo largo de los últimos tiempos, sobre todo en cuanto a aspectos metodológicos. Por ello hemos incorporado cuestiones como el análisis de los textos originales, que, en nuestra opinión, ofrece la posibilidad de realizar una valoración más acertada de las causas que pueden llegar a explicar — siempre de manera tentativa— los fenómenos observados en la lengua traducida, o la aplicación de técnicas estadísticas para la validación de los resultados, una tendencia cada vez más extendida en ciertos ámbitos de los Estudios de Traducción, que quizás buscan consolidar su estatus como una disciplina lo más «científica» posible (Gile 1998).

3.2. Diseño del corpus

Corpus-based research in descriptive translation studies critically depends on the availability of suitable tools and resources (...). (Fantinuoli y Zanettin 2015: 1)

Si nos basamos en las sugerencias metodológicas recogidas en Baker (1995) para la búsqueda de universales de traducción, el primer paso necesario para la consecución de los objetivos prácticos de este trabajo es la compilación de un corpus comparable. En nuestro caso se trata de un corpus de textos literarios contemporáneos compuesto por un subcorpus de textos escritos originalmente en español y otro subcorpus de textos traducidos del inglés al español. Nos referiremos en adelante al primero como ES y al segundo como TRAD. Estos dos subcorpus constituyen el componente comparable de nuestro corpus. Por los motivos que ya hemos adelantado, a este componente comparable se añade un tercer subcorpus conformado por los textos originales en inglés que han dado lugar a las novelas traducidas al español de nuestro corpus comparable. Junto a TRAD, constituye el componente paralelo de nuestro corpus y lo denominaremos EN.

La creación de un corpus de estas características vino motivada, en primer lugar, porque esta tesis doctoral tiene su punto de partida en el trabajo de investigación

realizado para el itinerario de investigación del Máster en Tecnologías de la Traducción y Localización de la Universitat Jaume I. En aquel momento nos interesaba adentrarnos en las cuestiones relativas a la creación de corpus electrónicos para el estudio de la traducción y fue por este motivo por el que decidimos compilar nuestro propio corpus desde cero en lugar de utilizar uno disponible. Por otra parte, no conocíamos de la existencia de ningún corpus fácilmente accesible y disponible de manera gratuita que ofreciera un componente comparable de literatura contemporánea traducida al español y que permitiera la selección de un género y unas fechas de publicación específicas para su comparación⁴⁸.

En un primer momento pensamos en utilizar, como explica Rabadán (2005-2008: 313 y 2007: 487, 490), una selección del CREA que pudiera servir como componente comparable a nuestro subcorpus de traducciones, en vez de construir un corpus comparable completo desde cero. La herramienta de búsqueda del CREA permite filtrar los resultados por área temática, área geográfica, medio y fecha de publicación, por lo que sería factible trabajar con un subcorpus de características lo más comparables posibles a las del corpus paralelo compilado para este trabajo. Aunque existen muchos más criterios de comparabilidad que hay que tener en cuenta, los básicos —ámbito temático, fecha de publicación, tamaño y variedad lingüística de ambos subcorpus— podrían ser similares. La desventaja es que no podríamos disponer de un corpus etiquetado si lo necesitáramos y que algunas pruebas de búsqueda de expresiones nos convencieron de que la interfaz de consulta era bastante menos flexible que la de los programas de concordancias. Finalmente, nuestra preferencia por trabajar con un género narrativo en concreto nos llevó a tomar la decisión de crear el corpus entero.

Laviosa enumera, acertadamente, algunos de los problemas de los estudios de traducción basados en corpus, entre los que se encuentran la sobrerrepresentación de lenguas romances —cuestión que no podemos hacer nada aquí por resolver, pues trabajamos con las lenguas que dominamos— y el problema de la comparabilidad entre corpus de textos originales y traducidos en el mismo idioma, que trataremos más adelante. Sin embargo, se muestra confiada en el hecho de poder establecer si determinados usos lingüísticos (ella se refiere a los *core patterns of lexical use* que hemos comentado en la introducción teórica) dependen o no de factores específicos que

⁴⁸ Borja (2007: 252-264) ofrece un listado muy interesante de corpus disponibles en español en las fechas de su publicación. Algunos ya no pueden encontrarse en las direcciones facilitadas pero otros enlaces siguen activos.

pueden controlarse, como es el caso de las lenguas implicadas, el campo temático de los textos o la modalidad: «The procedure one could adopt for this purpose would consist of creating and comparing ad hoc subcorpora in which each of the latter features is controlled in turn» (Laviosa 1998b: 9).

Eso es lo que pretendemos hacer aquí. No podemos plantearnos compilar un corpus excesivamente grande ni incluir numerosas lenguas de partida, pero sí podemos intentar controlar el mayor número de variables posible para alcanzar un grado de comparabilidad elevado entre subcorpus e intentar asegurarnos de que nuestras conclusiones sean válidas para nuestros datos y no dependan de factores que no hemos tenido en cuenta, y, además, explicitar dichas variables para que nuestro análisis pueda ser replicable en circunstancias similares o diferentes.

Como hemos adelantado, el corpus pasó por dos fases, una de construcción y otra de remodelación.

3.2.1. FASE I - Construcción

Como indican Baker (1995: 235) y Laviosa (2003: 114) en sus respectivos trabajos sobre universales de traducción, compilar un corpus comparable de originales en una lengua y traducciones a esa misma lengua es un método válido para hallar patrones de uso que pueden ser típicos de la lengua traducida en contraposición a la lengua no traducida⁴⁹. Muchos autores han seguido las sugerencias metodológicas de Baker con el objetivo de descubrir si la lengua traducida y la no traducida difieren con respecto a la utilización de diferentes elementos o patrones léxicos o gramaticales, generalmente con el objetivo de hallar pruebas que apoyen la existencia de las características del proceso traductor de las que ya hemos hablado anteriormente, como la normalización o la explicitación.

Chesterman (2004a: 8) clasifica la infrarrepresentación de elementos únicos como un posible *T-universal*, es decir, una característica de la relación entre traducciones y no traducciones escritas en la lengua meta. Por este motivo, en un principio consideramos el componente comparable —es decir, los subcorpus ES y

⁴⁹ Siempre y cuando, advierte Laviosa (2002: 23-34) siguiendo a Baker, se den las siguientes condiciones: «elaboration of explicit criteria and procedures for the selection, acquisition and annotation of the texts to be included in the corpus» y «precise definition of the linguistic features which are considered concrete manifestations of the universals of translation (...) in order to render these global and abstract constructs operational and verifiable».

TRAD— como el corpus más adecuado a los objetivos que nos habíamos planteado. De hecho, el primer estudio que realizamos como parte del programa de máster y cuyos resultados incluimos como parte del capítulo de las perífrasis, se basó exclusivamente en los subcorpus de textos originales españoles y traducciones al español, sin tener en cuenta los textos originales en inglés.

McEnery y Xiao (2007: 21) afirman: «the components of a comparable corpus overcome translationese by populating the sampling frames with L1 texts from different languages». Si bien es cierto que lo que Baker (1993: 245) propuso en un principio fue precisamente la elaboración de un corpus que recogiera traducciones de diversas lenguas con el fin de minimizar el factor contrastivo —dado que las características inherentes al proceso traductor deberían mostrarse independientemente de la lengua de la que parta la traducción (Laviosa 1997: 296)—, no creímos que esto fuera lo más adecuado para el trabajo que íbamos a empezar en aquel entonces. Muchos autores que se han manifestado a partir de la publicación del artículo de Baker consideran que si deseamos separar las características inherentes al proceso de traducción de las ocasionadas por la influencia de la lengua de partida, es necesario contar con un corpus paralelo en el que poder consultar las obras originales. Y esto no es posible a menos que el investigador conozca las lenguas de partida; de ahí nuestra decisión de restringir el componente traducido del corpus comparable al inglés: consideramos que teníamos que dejar abierta la posibilidad de añadir los textos paralelos en una fase posterior del análisis, lo que permitiría comprobar si las características observadas en las traducciones eran o no dependientes del par de lenguas analizado.

Después de una extensa consulta de bibliografía dedicada al género policíaco en España y en el ámbito anglosajón y a sus autores más representativos⁵⁰, decidimos incluir en nuestro corpus las siguientes novelas:

Subcorpus de español original (ES - fase I):

- *Cuentas pendientes* (Juan Madrid, 1995)
- *El premio* (Manuel Vázquez Montalbán, 1996)
- *El hombre de la navaja* (Andreu Martín, 1992)
- *Gálvez y el cambio del cambio* (Jorge Martínez Reverte, 1995)
- *Ritos de muerte* (Alicia Giménez Bartlett, 1996)

⁵⁰ Explicamos este proceso en el apéndice 1, dedicado a la descripción de las distintas tareas llevadas a cabo para la selección de los textos del corpus y a la reseña bibliográfica de los autores y traductores incluidos en el mismo.

Subcorpus de traducciones al español (TRAD – fase I):

- *Criminal Conversation / Conversación criminal* (Evan Hunter, 1994)

Trad: Carlos Milla Soler (1995)

- *Devil in a Blue Dress / El demonio vestido de azul* (Walter Mosley, 1990)

Trad: Rosa Corgatelli (1994)

- *L.A. Confidential / Los Ángeles confidencial* (James Ellroy, 1990)

Trad: Carlos Gardini (1991)

- *Ripley Under Water / Ripley en peligro* (Patricia Highsmith, 1991)

Trad: Isabel Núñez (1992)

- *Kissing the Gunner's Daughter / Un beso para mi asesino* (Ruth Rendell, 1992)

Trad: Carme Camps (1993)

3.2.2. FASE II - Remodelación

En efecto, los resultados de ese primer trabajo sobre perífrasis, aunque válidos en sí mismos y muy interesantes, evidenciaron la necesidad de complementar el corpus con un componente paralelo de textos originales en inglés. Sin este componente, nos resultaba imposible averiguar si las diferencias observadas eran causadas por diferencias estructurales entre el inglés y el español o la influencia de los textos originales o si eran una peculiaridad del proceso de traducción.

Como hemos explicado con anterioridad, Baker había apostado por cambiar el foco de los Estudios de Traducción de la comparación de originales y traducciones a la comparación de textos traducidos y no traducidos en la misma lengua (1995: 233). Y, si bien en ese momento su sugerencia de valerse de corpus comparables para el estudio de las características intrínsecas del proceso traductor fue revolucionaria y ayudó a que la disciplina experimentara una rápida evolución, los investigadores comenzaron a darse cuenta de que los análisis volvían a quedarse cojos si se descartaba por completo el estudio de originales⁵¹ y, en consecuencia, volvieron a ellos, aunque esta vez, como método complementario y no único (McEnery y Xiao 2007: 26; Macken, De Clercq y

⁵¹ Por supuesto, Baker no pretendía descartarlos del todo. Según Saldanha (2011: 238) la metodología de Baker (2000) para el estudio del estilo del traductor parte del análisis de los textos meta pero propone acudir a los originales posteriormente, para comprobar el efecto de las distintas variables que pueden ejercer su influencia en el texto traducido.

Paulussen 2011: 376). Entre los ejemplos de este tipo de estudios más comprensivos encontramos el de Per-Ola Nilsson (2002), en el que el autor pretende estudiar la sobrerrepresentación de un patrón lingüístico del sueco en traducciones a ese idioma, partiendo de un enfoque orientado a los textos meta —utiliza un corpus comparable para extraer las diferentes frecuencias de utilización— pero complementándolo, posteriormente, con los textos originales alineados para distinguir los casos en los que esta sobrerrepresentación es efectivamente una característica del proceso traductor de los casos en los que se debe a las diferencias inherentes a los distintos sistemas lingüísticos, y hallar así, explicaciones de los fenómenos observados en la comparación intralingüística. Es la misma metodología que adopta Laviosa (2007) y que pretendemos seguir en este trabajo.

Como sostiene Bernardini (2011: 3), «a methodology combining different types of corpora is not only advisable, but critical if one is to avoid methodological pitfalls». A través de dos estudios de caso, esta autora demuestra claramente las ventajas de un análisis cuantitativo llevado a cabo en un corpus comparable (que favorece un enfoque más sistemático, riguroso y objetivo; además de que permite la aplicación de tests de significación estadística a los resultados de la comparación entre variedades lingüísticas), pero también sus límites, dado que permite formular hipótesis sobre tendencias pero no explicaciones a los fenómenos observados. El análisis manual exhaustivo del componente paralelo es indispensable, según Bernardini, para asegurarnos de que las diferencias se deben al proceso de traducción y no a otras variables y para añadir poder explicativo a nuestras observaciones sobre las decisiones del traductor, además de para conectar la teoría con la práctica y la docencia de la traducción (ámbitos a los que, en su opinión, los corpus paralelos son más exportables que los comparables).

Operatively, I would suggest that a tripartite corpus structure (source texts in language A, target texts in language B and comparable/reference originals in language B) is the minimal setup to start to shed light on the distinctive features of translated language. (Bernardini 2011: 12)

No obstante, ya habíamos apuntado antes que la configuración ideal de un corpus para el estudio de la traducción sería la bidireccional, lo que habría supuesto añadir un cuarto subcorpus de traducciones al inglés para averiguar cómo afecta a los

resultados la dirección de traducción. Sin embargo, en teoría esto habría sido complicado manteniendo la configuración original de nuestro corpus. Si realizamos una rápida consulta al *Index Translationum*⁵², vemos que hay un desequilibrio muy grande entre las traducciones en una dirección y en otra. Tomando como referencia la década de los noventa y sin especificar el país de la traducción, encontramos que en la dirección inglés-español hay 40 331 obras catalogadas, mientras que en la dirección inversa, español-inglés, solo hay 3 814. De hecho, el único de los autores españoles incluidos en nuestro corpus que encontramos traducido al inglés en la base de datos es Vázquez-Montalbán, que, además, cuenta con traducciones al catalán, finés, húngaro, búlgaro, serbio, rumano, francés, italiano, japonés, alemán, islandés, danés, neerlandés, noruego, portugués, griego, sueco, polaco, croata, farsi y coreano. Pese a ello, *El premio* no se encuentra traducido al inglés⁵³. Andreu Martín no cuenta con ninguna traducción al inglés, aunque sí varias al alemán, el vasco, el francés, el gallego, el neerlandés, el estonio, el portugués, el italiano, el lituano, el serbio y el catalán (algunas de estas últimas, a juzgar por la falta del nombre del traductor, hechas por él mismo)⁵⁴. Juan Madrid tampoco está traducido al inglés; solo al alemán, francés, eslovaco, neerlandés, italiano, catalán, vasco y danés. Martínez-Reverte, al danés, alemán y griego, y Giménez Bartlett, al francés, alemán, griego, polaco, búlgaro y checo.

Sin embargo, hay que tener en cuenta que esta base de datos no está actualizada (actualmente, según se indica, se encuentran procesando los datos recogidos a partir de 2009). De hecho, mediante una búsqueda en la tienda *Amazon.com* sobre la autora Giménez Bartlett (motivada por la lectura de un artículo en el que se mencionaba su gran éxito en Estados Unidos⁵⁵), hemos encontrado varias traducciones de sus obras al inglés: dos publicadas antes de 2008. De los autores restantes no hemos encontrado ninguna. Sea como sea, y aunque hablaremos más adelante de la composición de nuestro

⁵² El *Index Translationum* es una base de datos de obras traducidas y publicadas en todo el mundo creada por la UNESCO en 1932. Puede consultarse en http://portal.unesco.org/culture/es/ev.php-URL_ID=7810&URL_DO=DO_TOPIC&URL_SECTION=201.html [última fecha de consulta: 7 de septiembre de 2015].

⁵³ Sí encontramos *Asesinato en el comité central*, *El pianista*, *Los mares del Sur*, *La Soledad del Manager*, *Galíndez*, *Barcelonas*, *El quinteto de Buenos Aires*, *El Hombre de mi vida*, *Tatuaje*, *El laberinto griego* y *El delantero centro fue asesinado al atardecer*.

⁵⁴ En un artículo sobre *El hombre de la navaja* y su reescritura al catalán cinco años después, Arnscheidt (2002) explica que Martín, hijo de padre castellanohablante y madre catalanoparlante, comenzó su carrera escribiendo en castellano porque no se sentía seguro de su dominio literario del catalán, y no fue hasta años más tarde, y en colaboración con correctores, que comenzó a traducirlas y a darse cuenta de que, en el proceso, podía enriquecerlas adaptándolas al contexto de la nueva lengua (tanto en una dirección como en otra).

⁵⁵ Alicia Giménez Bartlett gana el Premio Pepe Carvalho del 2015. (29 de octubre de 2014). [en línea] Recuperado el 27 de septiembre de 2015 de *Todoliteratura.es* <http://www.todoliteratura.es/noticia/7577/novela-negra/alicia-gimenez-bartlett-gana-el-premio-pepe-carvalho-del-2015.html>

corpus, podemos adelantar que el género que hemos seleccionado es especialmente dado a traducir del inglés pero no a exportar desde otras lenguas. Karen Seago (2014) comenta al respecto:

Detective or crime fiction has been an extremely popular form from its inception and translation has been a central feature of disseminating what has been seen as an originally Anglo-American genre. Despite the success of Nordic Noir and Italian crime bestsellers, translation into English remains at below 5% (2.5% in the UK, 3% in the USA; literature 4.5%) (Zgadzaj and Roberts 2013), while the bulk of translated crime is Anglo-American fiction: translated books make up 25%–40% of European publications.

Nuestra selección de textos habría sido extremadamente complicada si hubiéramos diseñado un corpus bidireccional, pues ¿qué autores se traducen al inglés? Generalmente, los representantes de lo que Zanettin (2012: 50) describe como «“difficult literary fiction” (Schiller 1993: 28), published in translation by small, specialized publishers». El género policíaco no se caracteriza por ser considerado especialmente *highbrow*. Esta visión ha ido cambiando poco a poco en el mundo anglosajón (Priestman 2003: 1), pero, como comenta Janerka (2010: 84), «Resulta curioso el notable desprecio que demuestra el mundo académico español hacia la novela policíaca». El hecho de que veamos una mayor cantidad de traducciones de obras de Vázquez Montalbán seguramente se debe a que este autor se ha forjado también una carrera en otros géneros literarios considerados por algunos de «mayor calidad literaria» o, simplemente, más consagrados (Berriatúa 2006).

Además de añadir el subcorpus de originales en inglés (EN) por los motivos ya expuestos, en esta segunda fase decidimos realizar ciertos cambios en la configuración de ES y TRAD (de este último en mayor medida). Una práctica que, a pesar de nuestras reservas iniciales al respecto, no es extraña en la lingüística de corpus. Tognini-Bonelli (2001: 58), refiriéndose a la cuestión de la representatividad de los corpus, afirma que, una vez el corpus ya ha sido compilado y obtenemos los primeros datos, puede suceder que su evaluación nos haga cuestionar los parámetros según los que se ha compilado el corpus o la tipología de los textos incluidos. Encontramos la misma advertencia en Atkins, Clear y Ostler (1991):

Since it is theoretically suspect to aim at achieving a perfectly ‘balanced’ corpus, it is practical to adopt a method of successive approximations. First, the corpus builder attempts to create a representative corpus. Then this corpus is used and analysed and its strengths and weaknesses identified and reported. In the light of this experience and feedback the corpus is enhanced by the addition or deletion of material and the cycle is repeated continually.

Como resultado de estas modificaciones, el corpus de la fase II, que es el que utilizaremos en los tres capítulos de análisis, quedó conformado como se presenta a continuación (los cambios se encuentran marcados con asteriscos y explicaremos las razones que nos llevaron a realizarlos a lo largo del siguiente apartado):

Componente comparable de español original (ES – fase II):

- *Cuentas pendientes* (Juan Madrid, 1995)
- *El premio* (Manuel Vázquez Montalbán, 1996)
- *El hombre de la navaja* (Andreu Martín, 1992)
- *Gálvez y el cambio del cambio* (Jorge Martínez Reverte, 1995)
- *Ritos de muerte* (Alicia Giménez Bartlett, 1996)
- **La danza de los reptiles* (Soledad Núñez, 1994)

Componente paralelo de inglés original (EN) e inglés traducido al español (TRAD – fase II):

- **Rum Punch / Cóctel explosivo* (Elmore Leonard, 1992)
Trad: Enrique de Hériz Ramón (1994)
- **Cruel and Unusual / Cruel y extraño* (Patricia Cornwell, 1993)
Trad: Jorge Luis (Jordi) Mustieles Rebullida (1994)
- *Ripley Under Water / Ripley en peligro* (Patricia Highsmith, 1991)
Trad: Isabel Núñez (1992)
- **White Jazz / Jazz blanco* (James Ellroy, 1992)
Trad: Hernán Sabaté Vargas (1993)
- **L is for Lawless / L de ley (o fuera de ella)* (Susan Grafton, 1995)
Trad: Antonio-Prometeo Moya (1996)

3.3. Creación del corpus

A continuación, describiremos los pasos que seguimos para la selección, la compilación y el preprocesamiento de los textos del corpus utilizado en este trabajo. Los subapartados en los que dividiremos la explicación serán los siguientes⁵⁶:

3.3.1. Selección de los textos

3.3.2. Digitalización

3.3.3. Alineación

3.3.4. Segmentación

3.3.5. Etiquetado

3.3.6. Indexación

3.3.1. Selección de los textos

El primer paso en la creación de un corpus de textos electrónicos es la selección de los textos que conformarán los diferentes subcorpus o componentes. Dado que la validez de los resultados de la investigación depende de la representatividad del corpus en general y, en particular, de los criterios de inclusión de los textos, es imprescindible realizar, en primer lugar, una meticulosa selección de criterios para decidir qué textos deben incluirse y cuáles no y, en caso de que el corpus sea comparable, definir los parámetros de comparabilidad más adecuados.

Si repasamos los criterios que cada autor considera fundamentales para garantizar la creación de un corpus representativo y equilibrado, comprobaremos que estos difieren enormemente en número y definición, porque hay tantos diseños de corpus como objetivos diferentes en el mundo de la práctica y la investigación de la lingüística y la traducción. Borja (2007: 263) resume esta cuestión de la siguiente manera:

⁵⁶ Somos conscientes de que la extensión de este capítulo es superior a lo habitual, pero hemos querido realizar una descripción minuciosa de los pasos que hemos dado porque creemos que puede resultar provechoso para orientar a quienes, como nosotros, empiecen la tarea de crear un corpus desde cero.

There is a vast variety of corpus designs and as a result when discussing the design of corpora for translators we need to define the profile and objectives of the corpus users with great precision. Nevertheless, as there is no single translator profile or a single profile of Translation Studies approach, it is impossible to talk about a single ideal corpus design for translators, but rather of specific designs for specific translation or research purposes.

Lo primero que debemos tener en cuenta es que no es lo mismo compilar un corpus general, cuyo objetivo es incluir la mayor cantidad posible de tipos de texto de manera equilibrada para representar el estado de una lengua⁵⁷, que un corpus especializado, con el que solo se pretende hacer generalizaciones con respecto a un tipo específico de textos (McEnery y Xiao 2007: 21). Atkins, Clear y Ostler (1992) advierten de que el equilibrio de textos perfecto es una meta difícil de alcanzar, sobre todo antes de la construcción y aprovechamiento del mismo ya que, como hemos apuntado antes, las fortalezas y debilidades del diseño se apreciarán una vez que el corpus sea utilizado para los fines para los que se ideó. Por otra parte nos dicen (1992: 6) que, en su experiencia, todo corpus es fuente de información e inspiración, aunque su diseño no sea perfecto: «It would be shortsighted indeed 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'».

La representatividad es una de las primeras características que hallamos por lo general en las definiciones de lo que es un corpus. No obstante, estamos de acuerdo con Saldanha (2009: 1) cuando afirma: «The problem with making representativeness the defining characteristic of a corpus is that it is very difficult to evaluate and it will always depend on what the corpus is used for (...) and on the specific linguistic features under study». No solo de evaluar —un problema que también menciona Tognini-Bonelli (2001: 57)—, sino de conseguir, dado que existen limitaciones prácticas a la hora de obtener los textos más adecuados, como cuestiones de disponibilidad, derechos de autor o falta de presupuesto (Zanettin 2012: 46). Saldanha defiende la definición más flexible de Bowker y Pearson (2002: 9), que se centra en los criterios de compilación como característica definitoria de un corpus, en lugar de en la representatividad: «A corpus

⁵⁷ Es decir, «los documentos que lo constituyen deben guardar entre sí, y en todo momento, las mismas proporciones que existen en el conjunto del estado o nivel de lenguas representados» (Alvar Ezquerro, Blanco Rodríguez y Pérez Lagos 1994: 10).

can be described as a large collection of authentic texts that have been gathered in electronic form according to a specific set of criteria».

Para ejemplificar la variedad de opiniones que mencionábamos con respecto a criterios de compilación, podemos citar a Sinclair (1991: 20), por ejemplo, que considera que entre los que deberían tenerse en cuenta para crear un corpus general deberían encontrarse los seis siguientes: si la obra es de ficción o no, si es un libro, una revista o periódico, si el registro es formal o informal, y la edad, sexo y origen de los autores. Bowker y Pearson (2002: 45) describen nueve criterios para la compilación de corpus de lenguaje especializado: tamaño del corpus, selección de textos enteros o fragmentos, número de textos, medio de los textos originales (p. ej. escrito u oral), campo temático (un criterio difícil de evaluar, pues los límites entre distintos campos a veces no están claros, ni el nivel de generalidad para determinar la exclusión de determinados textos), tipo textual (comunicación entre expertos o de experto a no experto, por ejemplo), autoría, lengua (incluida la cuestión de si trabajamos con originales o traducciones y si el autor es nativo o no) y fecha de publicación.

En general, los criterios que acabamos de mencionar son —o están basados en— los establecidos por lingüistas para corpus monolingües (Baker 1995: 230). Laviosa, centrándose en los corpus comparables, afirma que lo importante es el nivel de comparabilidad entre el componente de textos traducidos y el de no traducidos «which should be similar in as many respects as possible in an attempt to ensure that any linguistic differences found between them can be reliably attributed to their different status as translation vs. non-translation, rather than to confounding variables» (1997: 290). Esta autora se basa en la tipología de Atkins, Clear y Ostler (1991), modificándola un poco y ampliándola en ciertos apartados para adaptarla al contexto de nuestra disciplina (hemos subrayado en ella los criterios que definirían nuestro corpus):

Level I

Corpus types:

- FULL-TEXT, SAMPLE, MIXED (FULL-TEXT AND SAMPLE), MONITOR
- SYNCHRONIC, DIACHRONIC
- GENERAL, TERMINOLOGICAL
- MONOLINGUAL, BILINGUAL, MULTILINGUAL
- LANGUAGE(S) OF CORPUS
- WRITTEN, SPOKEN, MIXED (WRITTEN AND SPOKEN)

Level II

Monolingual Corpus Types: - SINGLE, COMPARABLE

Level IIIa

Single Corpus Types: - TRANSLATIONAL, NON-TRANSLATIONAL

Level IIIb

Comparable Corpus Types: - TRANSLATION-DEPENDENT, NON-TRANSLATION-DEPENDENT,
INDEPENDENT

Level IV

Translational Corpus Types: - MONO-SOURCE-LANGUAGE, BI-SOURCE-LANGUAGE, MULTI-SOURCE-LANGUAGE
- MONO-TRANSLATING-MODE, BI-TRANSLATING-MODE, MULTI-TRANSLATING-MODE
- MONO-TRANSLATION-METHOD, BI-TRANSLATION-METHOD, MULTI-TRANSLATION-METHOD
- INTO-MOTHER-TONGUE, INTO-FOREIGN-LANGUAGE, INTO-LANGUAGE-OF-HABITUAL-USE, MIXED-TARGET-LANGUAGE-STATUS
- PROFESSIONAL, STUDENT
- PUBLISHED, UNPUBLISHED

Además de estos parámetros, a la hora de diseñar el English Comparable Corpus (ECC, uno de cuyos subcorpus es el TEC) se tuvieron en cuenta también cuestiones como la distribución de autores/traductores según el sexo y de autores/traductores individuales o equipos y características de la audiencia meta como la edad o el género.

De todas formas, aunque sería deseable que los corpus comparables se conformaran siguiendo esta u otras largas listas de parámetros para garantizar el mayor grado posible de comparabilidad entre sus componentes, como todo, esta es una cuestión relativa. Por ejemplo, Philip (2009) opina que para algunos estudios, dos corpus monolingües de referencia como el BNC y el CREA —aunque ella habla del Bank of English y el CORIS (Corpus di Italiano Scritto)—, pueden ser en cierto modo comparables, independientemente de que coincidan o no en tamaño y en composición: «Dissimilar composition is not proof of incomparability: in accepting that languages are anisomorphic, it should come as no surprise that different languages give more priority to some text types, and less to others» (Philip 2009: 61-62). Recordemos también el caso de CEXI: a veces, la búsqueda de representatividad en un corpus puede ir en detrimento de la comparabilidad de sus componentes y viceversa.

Dice Tognini-Bonelli (2001: 62), con respecto a cuestiones que afectan a la creación de corpus, como la representatividad, los criterios de selección o la autenticidad de los textos, que estas seguirán siendo objeto de debate durante mucho tiempo; «the important point is that workers in the field should be as explicit as

possible, particularly about what they have done in assembling their corpus, and if possible some discussion about why they have taken these decisions»⁵⁸. A continuación, nos valdremos de las seis categorías que enumera Austermühl (2001: 125-128) para diferenciar entre tipos de corpus⁵⁹, complementándolas con aspectos de la tipología de Laviosa que hemos transcrito, para describir nuestro corpus y los criterios que hemos seguido para su creación. Las categorías son las siguientes (los subapartados *b*, *d* y *e* son adición nuestra):

3.3.1.1. Medio de representación del texto original

3.3.1.2. Medio de representación del corpus

3.3.1.3. Número de lenguas

3.3.1.4. Características de la preparación de los textos

3.3.1.5. Funciones del corpus

3.3.1.6. Características de los textos seleccionados

- a) criterio cronológico
- b) género
- c) criterios geográficos y sociolingüísticos
- d) sexo de los autores y traductores
- e) tamaño

3.3.1.1. Medio de representación del texto original

Los textos a partir de los cuales hemos construido el corpus son novelas completas y se encuentran en formato impreso en papel.

3.3.1.2. Medio de representación del corpus

El corpus se ha creado en formato electrónico, en forma de archivos TXT.

⁵⁸ Opinión compartida con Sinclair (1991: 13): «until we know a lot more about the effects of our design strategies we must rely in publishing a list of exactly what is in a corpus».

⁵⁹ Este autor también opina que «The different types of corpora are as numerous as their functions».

3.3.1.3. Número de lenguas

Se trata de un corpus bilingüe⁶⁰. El componente comparable monolingüe incluye un subcorpus de textos escritos originalmente en español y otro de textos traducidos al español. El componente paralelo está formado por las obras originales en inglés que han dado lugar a dichas traducciones.

3.3.1.4. Características de la preparación de los textos

Para la FASE I se dejaron los textos en formato TXT puro, sin realizar etiquetado o marcado de ningún tipo. Para la FASE II, se etiquetaron todos los textos a nivel morfológico.

3.3.1.5. Funciones del corpus

La función del componente comparable de este corpus es la de reflejar y representar el español contemporáneo traducido y el producido en circunstancias naturales en un género concreto, y servir como herramienta de investigación para comprobar nuestras hipótesis acerca de las características propias de los textos traducidos. El componente paralelo del corpus, es decir, el subcorpus de textos originales ingleses (EN) a partir de los cuales se realizaron las traducciones al español del subcorpus TRAD, nos permitirá superar las limitaciones inherentes al análisis del corpus comparable e ir un paso más allá, buscando en los textos originales o en la lengua de partida explicaciones a los fenómenos observados en los textos escritos en español.

3.3.1.6. Características de los textos seleccionados

Esta es la categoría más importante y en la que nos extenderemos más.

⁶⁰ Según McEnery y Xiao (2007: 19), también podría denominarse *multilingüe*, un término que a veces se reserva para corpus compuestos por tres lenguas o más, aunque en sentido general se utiliza también cuando está involucrada más de una única lengua.

a) Criterio cronológico

El nuestro es un corpus sincrónico, pues se circunscribe a un período de tiempo determinado, y contemporáneo, ya que intenta reflejar el estado actual —o lo más cercano posible al actual— del uso del lenguaje. Hemos seleccionado a autores que escribieron y fueron traducidos en torno a la misma época, y de manera contemporánea en ambos subcorpus. Todas las obras incluidas se escribieron o tradujeron y, en definitiva, se publicaron durante el período comprendido entre los años 1990 y 1996.

b) Género

Los textos incluidos en el corpus son obras literarias de ficción narrativa para adultos. Se decidió utilizar este tipo de textos por diversos motivos: por un lado, debido a nuestro interés personal y subjetivo por el área de la traducción literaria y, por otro, porque creemos que la prosa narrativa publicada ofrece mejores probabilidades de encontrar un lenguaje más elaborado, recursos más creativos y soluciones de traducción más meditadas⁶¹ que otro tipo de textos de naturaleza más especializada (Olohan 2004: 191, resumiendo a Tymoczko 1999), y según algunos autores, hasta puede resultar menos permeable a la influencia de la lengua original (Ramón y Labrador 2008: 289). Por otra parte, a pesar de que, por supuesto, poseen sus características particulares relacionadas con la utilización profusa de diversos recursos literarios, en ellos es posible encontrar todo tipo de textos, temas, estilos, registros y dialectos (Marco, Verdegal y Hurtado 1999: 168).

Ya hemos comentado que uno de los criterios de comparabilidad que consideramos más importantes a la hora de crear nuestro corpus comparable es el de la selección de un género textual específico y bien identificable. La delimitación de los distintos géneros no es una cuestión banal y conlleva sus dificultades, pero creemos que puede aportar una coherencia a los dos subcorpus que sería difícil de conseguir de otra manera —en términos, por ejemplo, de campo temático, usos lingüísticos, tipología de personajes o cantidad de diálogo— y que puede ayudar a minimizar el impacto de las diferencias que podrían darse de lo contrario en los distintos subconjuntos de textos

⁶¹ Aunque no siempre sucede en la realidad, se supone que los traductores literarios no tienen, por norma general, que tener la traducción lista «para ayer», como sucede en otros ámbitos profesionales de la traducción, sino que pactan una fecha de entrega acorde con las características del encargo.

sobre los fenómenos que pretendemos estudiar. Por poner un ejemplo muy evidente, no es lo mismo estudiar el uso del imperativo en artículos periodísticos que en recetas de cocina⁶².

García Izquierdo (2005) afirma que el género es «una interfaz entre el texto de partida y el texto de llegada y representa el recipiente en el que convergen todos los aspectos lingüísticos y culturales de las comunidades en contacto», pero advierte del hecho de que «el género no es estático. Es una categoría dinámica e híbrida, y de ahí la dificultad de establecer los límites entre géneros, en especial en los ámbitos de especialidad». Aun así, Lee (2001: 37) revisa las nociones *genre*, *register*, *text type*, *domain*, *sublanguage* y *style* y decide decantarse por la clasificación de los textos de un corpus en géneros, puesto que considera que es el nivel de categorización textual más práctico y útil desde el punto de vista teórico y pedagógico: «language teachers and researchers need to know exactly what kind of language they are examining or describing. Furthermore, most of the time we want to deal with a specific genre or a manageable set of genres, so that we can define the scope of any generalisations we make». Hareide y Hofland (2011: 85) adoptan una postura pragmática y definen el género de la siguiente manera (siguiendo a Swales 1999: 42): «how texts are perceived, categorized and used by members of a community». Estos investigadores utilizan, por tanto, el catálogo de una librería en línea sueca para categorizar los textos de su corpus sobre la base de categorías externas, no basadas en su propia intuición.

Por estos motivos, además de optar por los textos narrativos en sentido amplio, quisimos seleccionar un género bien identificable, presente en todos los catálogos de bibliotecas y librerías de ambas tradiciones literarias e incluso en ciertos corpus muy conocidos como el LOB (Lee 2001: 51) o el ENPC (Johansson, Ebeling y Oksefjell 1999/2002: 5)⁶³, y así nos decantamos por la novela policiaca o *crime fiction*, también denominada por distintos autores *mystery story* o *detective novel* (Janerka 2010: 80). Utilizamos estos términos en su acepción más general, puesto que este género engloba numerosos subgéneros (todos relacionados, por supuesto, con la investigación de un crimen; generalmente un asesinato):

⁶² Baker (2004: 171) afirma al respecto: «Even if we accept the decision to include biography in the TEC corpus, there is still the question of the comparability of fictional texts in BNC and TEC: ‘fiction’ is far too broad a category to ensure a reliable basis of comparison». Bernardini y Zanettin (2004: 59), por su parte, comentan lo siguiente sobre la importancia del género: «Any feature characterizing a corpus of translations may be the result not just of the process of translation, but of the genres of the texts and of the influence of the source language or languages. The importance of genre and target audience for non-fiction texts has been shown by Mauranen (2000, 2002), who compares translated and original Finnish non-fiction texts from the academic and popular domains».

⁶³ Denominado en el LOB *mystery and detective fiction* y en el ENPC, *detective fiction*.

Crime fiction is an umbrella term adopted to refer to a multitude of sub-genres, from the early detective stories of ratiocination (Leblanc's Arsène Lupin or Doyle's Sherlock Holmes) to the clue puzzles of the Golden Age (Christie's Poirot or Miss Marple series, Sayers' Lord Peter Wimsey or the American Ellery Queen); the private investigators of Hammett's or Chandler's hard-boiled or Paretsky's feminist hard-boiled; the professional investigative teams in the police procedural to the thriller (action, forensic, pathology, etc.). (Seago 2014)

En el ámbito hispanohablante, con una tradición menos amplia en este género (Janerka 2010: 44), se suele distinguir entre dos corrientes principales: la novela policíaca clásica o de enigma y la negra. Hoy en día, la más extendida y consumida parece ser esta última (Conte 1984; Janerka 2010: 82-83), y en nuestro corpus se refleja esa realidad, ya que predominan los autores de esta modalidad. La novela negra es considerada a veces como un subgénero, y otras, como la continuación o evolución de la clásica. Sin entrar en un debate terminológico, para una diferenciación sencilla entre ambos términos podemos recurrir al reconocido escritor español Andreu Martín —uno de los autores que incluimos en el subcorpus de originales en español— quien las describe de la siguiente manera en una entrevista:

La policíaca digamos que es el embrión del género y su característica principal es el hecho de que hay un enigma. Se plantea un juego al lector que debe ir resolviendo. La negra evolucionó de la policial. Sin embargo, la negra tiende más al análisis social, prescindiendo del elemento 'juego' y por eso, en lugar de complementarse, se produce una dicotomía: O juego o crítica. (Calderón 2004)

Una ventaja añadida de este género es que suele incluir un componente de diálogo bastante importante (De Haan 1996: 26), como estrategia para involucrar al lector en la investigación y resolución de los casos. Según el escritor y guionista estadounidense Syd Field (1994: 184), el diálogo, entre otras cosas, hace avanzar la historia, revela datos y acontecimientos, ayuda a conocer mejor a los personajes, sus relaciones y sus estados emocionales, comenta la acción y relaciona escenas. A veces, simplemente ayuda a captar la atención del lector añadiendo un componente de realismo y espontaneidad. Las partes dialogadas dotan a nuestros textos de variedad y, además, nos proporcionan material que intenta imitar el habla oral y espontánea, que es donde

creemos que puede haber más posibilidades de hallar algunos de los elementos estudiados en este trabajo (nos referimos a las perífrasis verbales y a los sufijos apreciativos). De hecho, según los resultados de un trabajo anterior en el que analizamos nuestro corpus con respecto a esta característica (Martínez Vilinsky 2014), en el subcorpus de originales en español «Dialogue, as a narrative technique, can be said to be widely used in this genre: the approximate percentage of dialogue found in our sample is above 50%. Two authors come close to 70% of dialogue, two are above 50% and two do not reach 40%», y, por otra parte, «the proportion of dialogue in the subcorpus of translations into Spanish is approximately 7.4% higher than in that made up of texts originally written in Spanish».

Para no interrumpir en exceso la continuidad de este capítulo, incluimos la descripción detallada del proceso de búsqueda y selección de las novelas y la reseña biográfica de los autores y traductores del corpus al final del trabajo, en el apéndice 1.

c) Criterio geográfico y sociolingüístico

Dado que es imposible que un corpus como el nuestro, con un número tan reducido de obras, abarque todas las variedades del español, hemos decidido limitar su alcance al ámbito del español de España, que es donde llevamos a cabo este trabajo. Los originales en español han sido escritos por autores españoles y, tanto estos como las traducciones, han sido publicados en España por editoriales españolas, de lo que cabría esperar tres cosas: en primer lugar, que el hecho de que los originales hayan sido publicados signifique que han pasado por un proceso de edición, lo que garantizaría un mínimo grado de aceptación en el sistema literario y, posiblemente, mayor cantidad de lectores que si no fuera así; en segundo lugar, que el hecho de que las traducciones hayan sido hechas por traductores profesionales y publicadas garantice también ese grado de aceptabilidad, además de su estatus como traducciones en la cultura de llegada (Johansson 1998: 11; Laviosa 1997: 296); y, finalmente, que la variedad lingüística utilizada en los textos fuera la del destinatario meta principal de estas editoriales, es decir, adultos cuya lengua es el español europeo.

Sin embargo, como ya hemos explicado, hubo que corregir el diseño del corpus con posterioridad a la realización de los primeros análisis. El motivo fue que durante el estudio de las perífrasis llevado a cabo con el corpus en su fase I, nos percatamos de que en tres de las novelas se utilizaban determinadas perífrasis —como «andar + gerundio»,

«llevar + gerundio», «romper a + infinitivo» y «comenzar + infinitivo»— con mayor frecuencia que en otras. Dos de estas novelas resultaron ser las traducidas por traductores de origen hispanoamericano; en concreto, dos traductores argentinos muy prolíficos: Rosa Corgatelli en el caso de la novela de Mosley y Carlos Gardini en la de Ellroy. Aunque suponemos que la aceptación por parte de los editores de su publicación en España⁶⁴ significa que la manera de traducir de Corgatelli y Gardini minimiza estas diferencias, y aunque estas puedan deberse a otros factores, dado que la tercera novela con frecuencias particulares es la de Hunter, traducida por Milla Soler, este descubrimiento nos hizo pensar que, a la hora de estudiar determinados fenómenos lingüísticos, quizás era mejor ceñirse en lo posible a una única variedad del español, con el objetivo de eliminar la variable geográfica y garantizar el criterio de comparabilidad entre subcorpus⁶⁵. Procedimos entonces a buscar otras traducciones de novelas policíacas estadounidenses que hubiesen sido realizadas por traductores españoles, sin descartar la posibilidad de realizar en el futuro un análisis comparativo entre variedades del español, ahora que somos conscientes de estas posibles diferencias.

En cuanto a la variedad lingüística de las novelas originales en inglés, en la primera fase no habíamos tenido en cuenta el origen de los escritores, puesto que solo teníamos en mente el estudio del español traducido y no traducido, pero, al pasar de la idea de un corpus únicamente comparable a la adición del componente paralelo y, sobre todo, al observar las diferencias mencionadas en ES, pensamos que quizás era preferible unificar también la variedad de inglés utilizada. En este caso seleccionamos la estadounidense por ser la más frecuente en el corpus original y, en consecuencia, eliminamos de la lista a la única autora inglesa, Ruth Rendell.

Todas las novelas son de autores y traductores diferentes, para evitar que el corpus represente el idiolecto personal de un autor determinado. Aun así, existe la posibilidad de que, al ser textos pertenecientes al género policíaco, el lenguaje utilizado represente a un grupo social determinado, con las características particulares de sus protagonistas prototípicos —policías y detectives profesionales son los que pueblan actualmente este tipo de novelas (Abasolo 2002)—.

⁶⁴ Posteriormente comprobamos que la editorial que publicó *El demonio vestido de azul* en España (Anagrama), según los datos que figuran en la base de datos del ISBN de este país, no fue la primera, puesto que en la del ISBN de Argentina está la primera edición de la traducción realizada por Corgatelli, con fecha de 1992 (Emecé Editores).

⁶⁵ Claro está que esta meta es difícil de alcanzar cuando en un mismo país conviven tantas maneras diferentes de hablar el español. Sin embargo, en algún lugar hay que poner el límite, y en este caso, hemos decidido hacerlo en las fronteras del estado español.

Un último apunte con respecto a la variedad lingüística representada es el hecho de que los escritores españoles son originarios de distintas partes de España pero, seguramente debido al mapa editorial de este país, la mayor parte de los traductores proceden de Cataluña o han tenido algún tipo de contacto con esta región. Son traductores profesionales experimentados y prolíficos —algunos, incluso, escritores también—. Las editoriales probablemente prefirieron escoger profesionales de confianza para la traducción de las novelas de estos autores estadounidenses conocidos en España o muy exitosos en su país de origen. Además, como dice Seago (2014: 8): «Crime fiction is a hugely creative genre making very great demands on authors and translators alike, demanding literary and aesthetic skills and the ability to deploy these within the constraints of genre norms».

d) Sexo de los autores y traductores

En cuanto al criterio del sexo de autores y traductores, mencionado también como una variable que podría influir en los resultados por Laviosa y otros autores, hay más mujeres en el subcorpus de traducciones, lo que refleja el hecho de que el fenómeno de las autoras de novela policiaca está claramente más extendido en el mundo anglosajón; de hecho, muchas páginas web y libros especializados en este género, librerías electrónicas y recopilaciones bibliográficas cuentan ya con un apartado para las «deadly women» o «damas del misterio», como a veces se las denomina. En el subcorpus de español original, por el contrario, nos resultó difícil hallar en la segunda fase del corpus, cuando debimos ampliar el número de novelas, otra mujer que hubiese escrito novela policiaca en los años que buscábamos. Para la sexta novela barajamos los nombres de autores de reconocido prestigio como Francisco González Ledesma o Eduardo Mendoza. Aunque, de todas formas, este último no cuenta con novelas publicadas entre las fechas necesarias, en principio, la elección de otro autor de sexo⁶⁶ masculino estaría justificada por el criterio de representatividad. Así como en la novela policiaca anglosajona las autoras están muy presentes (y así sucede en nuestro corpus de traducciones), en España los nombres que se encuentran dentro de este género suelen ser hombres, con excepciones como las de Giménez Bartlett, o autoras cuya producción

⁶⁶ Hareide y Hofland (2011: 87) recuerdan la importancia de consignar información acerca del género de autores y traductores al compilar un corpus, puesto que «Several studies confirm that men and women make different use of the language (Lakoff 1975; Holmes 1995). Less is known about the importance of the translator's gender, however, it is assumed to be a potentially interesting subject of study».

policíaca es demasiado antigua (p. ej. Lourdes Ortiz⁶⁷) o reciente (p. ej. Mercedes Castro⁶⁸) para encajar con los criterios de compilación de nuestro corpus.

Sin embargo, dado que nuestro objetivo, además de la representatividad, era la comparabilidad, y en TRAD hay tres autoras frente a una única en ES, quisimos añadir una más. Encontramos a Soledad Núñez, una escritora vasca no demasiado conocida que contaba con dos novelas policiacas de una misma serie en su haber. Aun cuando hubiera quien por falta de fama o galardones no considerara a esta autora a la altura de los demás componentes del corpus, su inclusión estaría más que justificada, pues, como dice Sinclair (1991: 17): «One of the principle (sic) uses of a corpus is to identify what is central and typical in the language. [...] If we are to approach a realistic view of the way in which language is used, we must record the usage of the mass of ordinary writers, and not the stray genius or the astute journalist».

Desafortunadamente, no fue posible controlar la variable del sexo en el caso de las traducciones, puesto que en nuestras novelas solo hay una traductora: Isabel Núñez (Highsmith). En la fase I había dos más, la traductora de Mosley (Rosa Corgatelli) y la de Ruth Rendell (Carme Camps), que fueron reemplazadas por hombres al cambiar las novelas del subcorpus. Es difícil conseguir el equilibrio en cuestiones de sexo, puesto que, como es evidente, los hombres no traducen solo a hombres ni las mujeres a mujeres.

e) Tamaño

Con el fin de que el corpus fuera comparable en cuanto a tamaño, en un primer momento se seleccionaron cinco novelas de autores españoles y cinco novelas de autores anglosajones traducidas al español. Las novelas se digitalizaron completas. Sin embargo, tras la digitalización, se comprobó que el número de palabras difería bastante de un subcorpus a otro, por lo que las cifras resultantes del análisis posterior debieron ser normalizadas.

⁶⁷ Lourdes Ortiz (2003). [en línea] Recuperado el 27 de septiembre de 2015 de *escritoras.com* <http://escritoras.com/escritoras/Lourdes-Ortiz>

⁶⁸ Azancot, N. (20 de febrero de 2009). Póquer de damas [en línea]. Recuperado el 27 de julio de 2009 de *elcultural.com* http://www.elcultural.com/articulo_imp.aspx?id=24789

FASE I

Subcorpus de español original (ES): 324 431 palabras

Subcorpus de español traducido (TRAD): 520 410 palabras

Total: 844 841 palabras

La diferencia en el número de palabras radica en el hecho de que, en este género, las novelas estadounidenses suelen ser bastante más extensas que las españolas. En las imágenes 2 y 3 podemos ver las capturas de pantalla de la información estadística que proporciona el programa WordSmith⁶⁹ (mediante la función «estadísticas generales» en WordList) para cada subcorpus. En nuestro corpus, la media de palabras de las novelas españolas es de aproximadamente 65 000, mientras que las traducciones están en torno a las 104 000⁷⁰.

N	1	2	3	4	5	6
Text File	OVERALL	GIMNEZ-1.TXT	MADRID-1.TXT	MARTN_-1.TXT	MARTNE-1.TXT	VZQUEZ-1.TXT
Bytes	1.932.441	519.315	231.899	325.918	324.234	531.075
Tokens	324.431	84.825	41.309	56.075	54.306	87.916
Types	29.040	12.322	5.968	9.966	8.463	13.266
Type/Token Ratio	8,95	14,53	14,45	17,77	15,58	15,09
Standardised Type/Token	47,64	50,78	44,39	45,87	47,12	47,61
Ave. Word Length	4,56	4,68	4,26	4,50	4,61	4,69
Sentences	13.622	3.603	2.010	2.047	3.110	2.852
Sent.length	23,63	23,53	19,88	27,20	17,36	30,66
sd. Sent. Length	23,14	20,11	15,67	31,81	12,23	29,56
Paragraphs	174	2	96	21	30	25
Para. length	1.665,93	42.412,50	430,30	1.030,57	1.810,20	3.511,60
sd. Para. length	6.956,40	59.971,85	432,86	4.686,73	2.222,84	5.055,16
Headings	0	0	0	0	0	0
Heading length						
sd. Heading length						
1-letter words	17.642	3.831	2.495	3.435	2.872	5.009
2-letter words	83.493	21.124	11.361	14.626	13.828	22.554
3-letter words	47.591	12.268	5.966	8.272	8.667	12.418
4-letter words	30.155	8.396	4.278	5.587	4.491	7.403
5-letter words	37.593	10.390	5.217	6.282	6.106	9.598
6-letter words	29.637	7.897	3.928	4.888	4.839	8.085
7-letter words	26.046	6.754	3.450	4.269	4.396	7.177
8-letter words	19.320	5.133	2.008	3.216	3.400	5.563
9-letter words	13.294	3.652	1.296	2.252	2.292	3.802
10-letter words	9.390	2.376	767	1.590	1.581	3.076
11-letter words	4.899	1.358	314	784	913	1.530
12-letter words	2.665	819	132	440	467	807
13-letter words	1.442	445	58	228	239	472
14(+)-letter words	700	206	20	112	138	224

Imagen 2 - Estadísticas del subcorpus de originales ES (fase I)

⁶⁹ Uno de los programas de análisis léxico más conocidos, creado por Mike Scott en 1996. En este trabajo hemos utilizado la versión 3 (1999), distribuida por Oxford University Press. Según Wilkinson (2011): «*WordSmith 4.0* and *WordSmith 5.0* tidied up the interface considerably, making it more intuitive and user-friendly. Moreover these later versions seem to be more stable than their predecessor».

⁷⁰ En la fase II incluimos los originales y comprobamos que esta diferencia no tiene demasiado que ver con el proceso de traducción, sino que es un reflejo bastante fiel de la mayor longitud de las novelas originales. En este caso, las cifras aproximadas son las siguientes: 61 000 palabras de media en ES, 97 500 en TRAD y 93 500 en los originales de EN. La diferencia de 4 500 palabras entre originales y traducciones se debe al incremento que se produce habitualmente en el número de palabras del texto meta cuando traducimos de una lengua como el inglés al español, aunque en algunas traducciones es más notable que en otras (por ejemplo, el texto de Grafton cuenta con más palabras en inglés que el de Ellroy y el de Highsmith, mientras que las traducciones de los textos de estos últimos superan a la del de Grafton).

N	1	2	3	4	5	6
N	1	2	3	4	5	6
Text File	OVERALL RENDEL-1.TXT MOSLEY-1.TXT HUNTER-1.TXT HIGHSM-1.TXT ELLROY-1.TXT					
Bytes	3.131.833	729.609	319.338	689.580	577.592	815.714
Tokens	520.410	122.360	54.532	115.073	97.607	130.838
Types	30.493	11.981	6.767	12.449	9.597	14.877
Type/Token Ratio	5,86	9,79	12,41	10,82	9,83	11,37
Standardised Type/Token	46,31	45,63	44,52	45,77	45,00	49,16
Ave. Word Length	4,55	4,53	4,31	4,52	4,49	4,79
Sentences	25.497	6.076	2.416	4.251	5.660	7.094
Sent.length	20,28	20,14	22,52	27,03	17,24	18,00
sd. Sent. Length	16,90	15,66	16,84	24,35	12,69	13,96
Paragraphs	723	8	10	39	27	639
Para. length	719,79	15.295,00	5.453,20	2.950,59	3.615,07	204,75
sd. Para. length	5.414,48	43.189,28	16.720,53	8.565,97	1.537,02	600,45
Headings	0	0	0	0	0	0
Heading length						
sd. Heading length						
1-letter words	28.796	6.536	3.256	6.292	5.898	6.814
2-letter words	123.991	29.262	13.508	28.855	22.822	29.544
3-letter words	74.785	17.743	7.748	16.718	15.098	17.478
4-letter words	52.591	13.082	6.263	10.692	9.553	13.001
5-letter words	66.959	16.111	7.748	14.464	12.494	16.142
6-letter words	52.252	12.298	5.764	11.758	9.699	12.733
7-letter words	46.782	10.669	4.522	10.430	9.207	11.954
8-letter words	29.875	6.689	2.691	6.392	4.863	9.240
9-letter words	19.489	4.276	1.495	4.051	4.017	5.650
10-letter words	11.941	2.698	782	2.493	1.977	3.991
11-letter words	6.622	1.509	449	1.532	1.036	2.096
12-letter words	3.356	751	181	729	492	1.203
13-letter words	1.679	415	68	349	275	572
14(+)-letter words	767	198	28	188	96	257

Imagen 3 - Estadísticas del subcorpus de traducciones TRAD (fase I)

Preferimos incluir textos completos en lugar de fragmentos, a pesar de que esta última técnica es muy común, sobre todo en corpus generales que requieren mucha variedad de textos o que van a estar disponibles al público general. Si incluimos textos completos, dice Sinclair (1991: 19), no tendremos que preocuparnos por la técnica de muestreo utilizada ni por las diferencias que existen entre las distintas partes de un texto —por ejemplo, en determinados artículos científicos, la mayor carga informativa se encuentra al principio, donde se definen los términos más importantes que se utilizarán en el resto del texto—. La desventaja de contar con textos completos es la posibilidad de abarcar menos diversidad de textos y de que un tema concreto o el estilo particular de algún autor pueda destacar. De todas formas, si se ampliara el corpus, siempre podría elegirse utilizar solo fragmentos del material que ya tengamos a nuestra disposición, mientras que sería más difícil tomar la decisión inversa.

En cuanto al tamaño ideal de un corpus, tampoco hay acuerdo entre los autores: «the term ‘corpus’ can be seen in the literature referring sometimes to a couple of short stories stored in electronic form and sometimes to the whole world wide web» (Saldanha 2009: 1). Hay quien dice que debería contener millones de palabras (Sinclair 1991: 18-19); otros argumentan que, para determinados estudios, puede obtenerse más información relevante de un corpus pequeño pero bien diseñado que de uno grande que no esté adaptado a las necesidades de nuestro proyecto (Bowker y Pearson 2002: 45). Aun así, Saldanha (2009: 2) opina que si el corpus es lo suficientemente grande, el

tamaño puede pesar más que los beneficios de seleccionar cuidadosamente los criterios de compilación. Olohan (2004: 46), por su parte, afirma que, como todas las demás cuestiones referidas al diseño de corpus, el tamaño también tiene que guardar relación con el propósito para el que se construye, y siguiendo a Kennedy (1998: 66-70), comenta que no siempre un corpus más grande será mejor, pues hay límites con respecto a la cantidad de información que puede analizar un investigador, por lo que para palabras de elevada frecuencia o estructuras gramaticales puede resultar útil un corpus pequeño. En nuestro caso, por supuesto, somos conscientes de que el tamaño reducido de nuestro corpus limitará las generalizaciones que podamos hacer sobre los fenómenos observados durante el análisis.

En la fase II, el cambio de novelas del subcorpus de traducciones ocasionó una disminución del número de palabras en TRAD. En esta ocasión, elegimos eliminar de la lista la novela de Evan Hunter, dado que era la más larga (en torno a 116 000 palabras), e intentamos que las nuevas fueran algo más cortas. Además, se añadió una novela más al subcorpus ES para intentar equilibrar los cómputos de palabras totales (a pesar de que esto significó desequilibrar el número de novelas de los subcorpus) y aumentar así el grado de comparabilidad de los subcorpus. No obstante, las cifras siguen siendo inferiores en ES, por lo que no es posible omitir el paso de la normalización de los resultados del análisis. Al respecto de esta cuestión, podemos citar a Baker (2004: 171):

It is in the nature of any type of comparison, any attempt to look for similarities and differences, that what is being compared can never be totally balanced in every respect. It is also particularly a feature of full text corpora, and particularly those of literary texts, that they cannot be balanced even internally: literary texts vary tremendously in their lengths, and if we are to include full texts in our corpus (which is desirable for many reasons), we have to accept that the individual texts will be seriously imbalanced in terms of size.

FASE II

Subcorpus de español original (ES): 367 091 palabras

Subcorpus de español traducido (TRAD): 488 653 palabras

Subcorpus de inglés original (EN): 458 403 palabras

Total: 1 314 057 palabras

En las imágenes 4, 5 y 6 vemos los datos recogidos por WordSmith acerca de la composición final de nuestros tres subcorpus de la fase II, que es la versión sobre la que basaremos casi todos los análisis de este trabajo. Observamos una diferencia en el cómputo de palabras entre los datos de dichas imágenes de WordSmith y los que aparecen en el recuadro de la página anterior y las tablas 1, 2 y 3, que son los que finalmente tendremos en cuenta para porcentajes, normalizaciones y tests de significación estadística. Estos últimos son los que hemos obtenido con la función «contar palabras» de Microsoft Word para cada archivo por separado. Decidimos realizar el cómputo de esta manera porque, en esta segunda fase de compilación, observamos que los diferentes programas utilizados (Word, Corpus Workbench y WordSmith) nos proporcionaban diferente número total de palabras, a veces con un margen de hasta 100 000 palabras en el caso de Corpus Workbench (los cálculos de Word y WordSmith se aproximan más). Esto sucede probablemente porque diferentes programas de concordancias y procesamiento de textos pueden adoptar distintos criterios a la hora de definir lo que es una palabra (por ejemplo, si *don't* cuenta como uno o como dos tokens).

A partir de los datos recogidos en estas imágenes, podríamos afirmar que, a primera vista, no se cumple lo predicho por Baker (1996) y Laviosa-Braithwaite (1997) en referencia a la hipótesis de la simplificación, con respecto a la variedad del vocabulario (vemos que el *type/token ratio* normalizado⁷¹ es apenas algo mayor en los textos originales en español que en las traducciones), pero sí a la longitud media de las oraciones, que es cinco puntos menor en los textos traducidos, a pesar de aumentar tres con respecto a los originales en inglés.

Texto ES	N.º de palabras	Porcentaje del total de palabras del subcorpus
Vázquez Montalbán	87 801	23,9 %
Giménez Bartlett	84 866	23,1 %
Martín	56 196	15,3 %
Martínez Reverte	54 336	14,8 %
Núñez	42 464	11,6 %
Madrid	41 428	11,3 %
Total ES	367 091	100%

Tabla 1 - Número de palabras y porcentaje sobre el total de los textos de ES

⁷¹ Olohan (2003: 87 nota 3) define así los conceptos de *type* y *token*: «Tokens are the number of words in a text, including repetitions of words. Types are the number of different words in a text, i.e. excluding repetitions». Zanettin (2012: 15) explica que esta proporción es un indicador de variedad léxica (cuantas más formas se utilicen con respecto al total de palabras, mayor será la variedad de vocabulario utilizada en un texto). Esta medida se normaliza para poder comparar la proporción entre textos de distintas longitudes (para conseguirlo, WordSmith lo computa cada 1 000 palabras).

N	1	2	3	4	5	6	7
N	1	2	3	4	5	6	7
Text File	OVERALL	GIMNEZ-1.TXT	MADRID-1.TXT	MARTIN-1.TXT	MARTNE-1.TXT	NEZ_SO-1.TXT	VZQUEZ-1.TXT
Bytes	2.191.915	519.317	235.656	335.129	324.235	247.250	530.328
Tokens	366.721	84.825	41.309	56.075	54.306	42.414	87.792
Types	31.069	12.321	5.974	9.966	8.463	7.680	13.249
Type/Token Ratio	8,47	14,53	14,46	17,77	15,58	18,11	15,09
Standardised Type/Token	47,85	50,78	44,38	45,87	47,12	49,43	47,61
Av e. Word Length	4,54	4,68	4,26	4,50	4,61	4,46	4,69
Sentences	17.459	3.603	2.658	2.047	3.110	3.203	2.838
Sent.length	20,90	23,53	15,52	27,04	17,36	13,18	30,78
sd. Sent. Length	21,08	20,11	10,38	31,57	12,23	10,34	29,59
Paragraphs	203	2	118	11	30	23	19
Para. length	1.733,83	42.412,50	350,08	4.677,91	1.810,20	1.422,17	4.597,95
sd. Para. length	6.402,98	59.971,85	447,05	2.659,58	2.222,84	2.511,70	5.373,45
Headings	0	0	0	0	0	0	0
Heading length							
sd. Heading length							
1-letter words	19.924	3.831	2.493	3.435	2.872	2.289	5.004
2-letter words	94.367	21.124	11.364	14.626	13.828	10.909	22.516
3-letter words	53.585	12.268	5.966	8.272	8.667	6.016	12.396
4-letter words	34.695	8.396	4.275	5.587	4.491	4.552	7.394
5-letter words	42.880	10.390	5.219	6.282	6.106	5.293	9.590
6-letter words	33.617	7.897	3.928	4.888	4.839	3.998	8.067
7-letter words	29.650	6.753	3.450	4.269	4.396	3.612	7.170
8-letter words	21.675	5.134	2.011	3.216	3.400	2.356	5.558
9-letter words	14.763	3.652	1.296	2.252	2.292	1.473	3.798
10-letter words	10.291	2.376	764	1.590	1.581	911	3.069
11-letter words	5.393	1.358	313	784	913	495	1.530
12-letter words	2.944	819	132	440	467	280	806
13-letter words	1.576	445	59	228	239	133	472
14(+)-letter words	752	206	20	112	138	52	224

Imagen 4 - Estadísticas del subcorpus de originales ES (fase II)

Texto TRAD	N.º de palabras	Porcentaje del total de palabras del subcorpus
Cornwell	109 871	22,5 %
Ellroy	109 757	22,5 %
Highsmith	98 295	20,1 %
Grafton	92 850	19 %
Leonard	77 790	15,9 %
Total TRAD	488 563	100 %

Tabla 2 - Número de palabras y porcentaje sobre el total de los textos de TRAD

N	1	2	3	4	5	6
N	1	2	3	4	5	6
Text File	OVERALL	CRUELY-1.TXT	ELLROY-1.TXT	GRAFTO-1.TXT	HIGHSM-1.TXT	LEONAR-1.TXT
Bytes	2.927.976	636.752	715.077	544.023	577.611	454.513
Tokens	487.475	109.350	110.620	92.573	97.606	77.326
Types	29.999	12.575	13.758	11.641	9.597	7.887
Type/Token Ratio	6,15	11,50	12,44	12,57	9,83	10,20
Standardised Type/Token	47,17	47,99	50,01	47,37	45,00	44,48
Av e. Word Length	4,52	4,60	4,73	4,49	4,49	4,30
Sentences	24.702	4.281	5.853	5.298	5.665	3.605
Sent.length	19,62	25,10	18,90	17,40	17,23	21,28
sd. Sent. Length	15,99	19,71	16,44	12,06	12,67	18,22
Paragraphs	283	92	3	22	27	139
Para. length	1.327,29	1.186,13	2,33	4.194,05	3.615,04	551,20
sd. Para. length	1.689,47	915,79	0,58	1.397,01	1.537,02	1.243,09
Headings	0	0	0	0	0	0
Heading length						
sd. Heading length						
1-letter words	28.000	6.102	6.376	5.141	5.898	4.483
2-letter words	120.141	27.860	25.750	24.000	22.822	19.709
3-letter words	69.866	15.280	14.484	13.131	15.097	11.874
4-letter words	46.648	9.762	10.536	8.989	9.554	7.807
5-letter words	60.169	12.965	13.027	11.371	12.493	10.313
6-letter words	49.183	10.464	11.810	8.937	9.701	8.271
7-letter words	43.414	9.529	10.327	7.877	9.206	6.475
8-letter words	28.290	7.102	7.135	5.232	4.863	3.958
9-letter words	18.511	4.201	4.765	3.359	4.019	2.167
10-letter words	11.057	2.755	2.936	2.138	1.975	1.253
11-letter words	6.167	1.648	1.617	1.248	1.035	619
12-letter words	3.057	843	909	604	492	209
13-letter words	1.736	500	546	300	275	115
14(+)-letter words	758	205	256	142	96	59

Imagen 5 - Estadísticas del subcorpus de traducciones TRAD (fase II)

Texto EN	N.º de palabras	Porcentaje del total de palabras del subcorpus
Cornwell	102 582	22,4 %
Grafton	93 587	20,4 %
Ellroy	93 149	20,3 %
Highsmith	92 973	20,3 %
Leonard	76 112	16,6 %
Total EN	458 403	100 %

Tabla 3 - Número de palabras y porcentaje sobre el total de los textos de EN

	N	1	2	3	4	5	6
	N	1	2	3	4	5	6
Text File	OVERALL	LEONAR-1.TXT	HIGHSM-1.TXT	GRAFTO-1.TXT	ELLROY-1.TXT	CORNWE-1.TXT	
By tes	2.597.325	415.913	524.868	516.527	563.362	576.655	
Tokens	467.518	79.441	95.499	94.248	95.232	103.098	
Types	20.827	4.871	7.182	8.723	10.259	8.963	
Type/Token Ratio	4,45	6,13	7,52	9,26	10,77	8,69	
Standardised Type/Token	44,98	38,51	43,19	45,73	51,15	45,29	
Ave. Word Length	4,21	3,86	4,19	4,20	4,46	4,30	
Sentences	32.610	4.004	4.568	7.948	8.376	7.714	
Sent.length	13,85	19,49	20,69	10,54	11,09	13,29	
sd. Sent. Length	12,56	15,91	18,21	7,11	10,13	10,58	
Paragraphs	2.959	129	77	2.321	351	81	
Para. length	157,52	615,82	1.221,73	40,61	271,32	1.272,81	
sd. Para. length	498,53	601,50	1.518,32	47,97	447,60	1.528,26	
Headings	0	0	0	0	0	0	
Heading length							
sd. Heading length							
1-letter words	28.172	5.760	5.554	5.702	5.457	5.699	
2-letter words	71.838	12.892	15.243	14.836	12.588	16.279	
3-letter words	108.108	20.849	23.416	21.533	18.895	23.415	
4-letter words	88.895	16.357	17.534	17.797	18.254	18.953	
5-letter words	55.915	8.930	10.528	11.540	12.714	12.203	
6-letter words	41.394	6.304	7.789	8.351	9.844	9.106	
7-letter words	33.418	4.598	7.242	6.643	7.603	7.332	
8-letter words	18.264	2.086	3.411	3.615	4.624	4.528	
9-letter words	11.260	961	2.875	2.068	2.663	2.693	
10-letter words	5.745	460	1.134	1.209	1.470	1.472	
11-letter words	2.385	126	393	521	590	755	
12-letter words	1.277	83	216	271	298	409	
13-letter words	582	27	135	115	151	154	
14(+)-letter words	184	6	20	34	57	67	

Imagen 6 - Estadísticas del subcorpus de originales EN (fase II)

Los criterios de compilación y la conformación final del corpus se asemejan mucho a los utilizados por Dayrell (2007) para su BBPC (Comparable Corpus of Brazilian Portuguese): este se centra en una única variedad del portugués e incluye libros de ficción narrativa para adultos completos y publicados a partir de los años noventa, y se ha intentado que las traducciones también fueran publicadas en el mismo periodo de tiempo. La lengua de partida de las traducciones es el inglés, y están hechas por traductores profesionales cuya lengua materna es el portugués brasileño. Cuando se escribió el artículo, el corpus contaba con 545 595 palabras en el subcorpus de traducciones y 565 920 palabras en el subcorpus de textos no traducidos, pero solo 5 novelas traducidas y 8 originales (es decir, que, como nuestro corpus, incluía más novelas en el subcorpus de originales, probablemente porque estas eran más cortas):

in line with other corpus-based translation studies which are also based on comparable corpora (see, in particular, Olohan (2003) and Baker (2004)), the translational and the nontranslational subcorpora have been designed to be of similar size in terms of number of words (tokens) rather than with respect to number of texts and/or number of authors/translators. This means that there is an imbalance in terms of the number of texts in each component. (Dayrell 2007: 379-380)

En su opinión, esto puede provocar que un determinado texto tenga más impacto en los resultados del subcorpus con menor cantidad de textos, y, por otra parte, implica que el componente no traducido posee mayor diversidad de temas y autores, dos cuestiones que podrían afectar a los resultados del análisis (2007: 397) y que habrá que tener en cuenta.

3.3.2. Digitalización

Los textos finalmente seleccionados, que están en formato impreso en papel, tienen que convertirse a archivos electrónicos mediante herramientas de digitalización para poder ser leídos por los distintos instrumentos de análisis y procesamiento de texto. Mediante un escáner HP Scanjet 5550c y a través del programa de reconocimiento óptico de caracteres (OCR) ⁷² OmniPage Professional 15.0 (ScanSoft, Inc ⁷³) se digitalizaron todas las páginas de cada libro y se guardaron en el formato de archivo de Omnipage, OPD. A continuación, se realizó el reconocimiento óptico de las imágenes guardadas con el programa, que transformó los ficheros creados a partir de la digitalización en ficheros de texto editable. El programa permite seleccionar y cambiar tipos de zona si no se han reconocido bien (podemos elegir entre texto, tablas y gráficos) y eliminar manchas o sombras de la imagen o incluso partes que no deseamos que se reconozcan, como encabezados o números de página.

⁷² Para una descripción general del funcionamiento de un escáner y los programas de reconocimiento óptico de caracteres, véase Bowker 2002: 25-26.

⁷³ Disponible en <http://www.nuance.es/empresas/producto/omnipage/professional/index.htm> (la empresa se ha unido desde entonces con Nuance Communications).

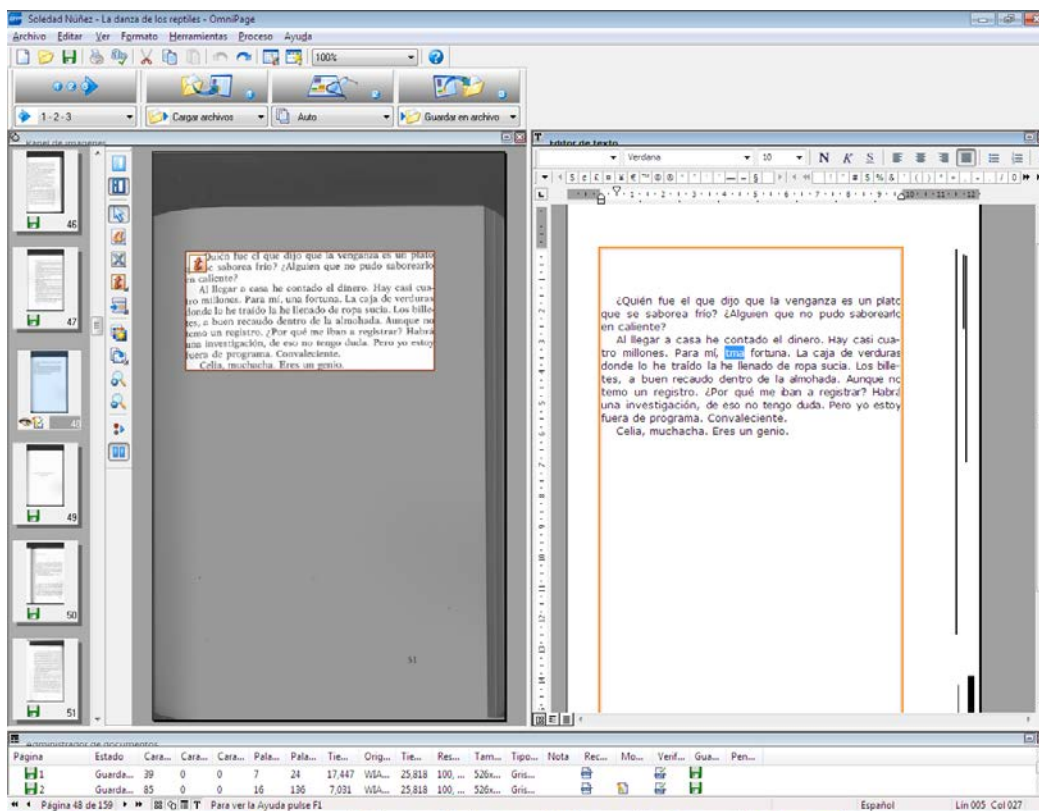


Imagen 7 – Reconocimiento de caracteres con el programa OmniPage

En la imagen 7 podemos ver una zona reconocida como texto en la parte izquierda, y la transcripción de los caracteres realizada por el programa en la parte derecha, donde podemos hacer una primera limpieza del texto y corregir errores que detectemos. Después de esta edición, guardamos los archivos en formato RTF⁷⁴. Estos archivos se revisaron meticulosamente, contrastándolos con los originales, para eliminar elementos como imágenes y cualquier error de digitalización, como omisiones, errores en el reconocimiento de determinadas palabras⁷⁵, palabras separadas en dos por guiones u oraciones cortadas por retornos de carro. Estas características del texto, así como los números de página, no nos interesan, puesto que no pretendemos poner el corpus a disposición del público sino únicamente analizar el texto. Si decidiéramos utilizar el corpus en otros estudios en el futuro, podríamos evaluar si conservar de alguna manera estos elementos o determinadas marcas tipográficas como la cursiva.

Normalmente, los errores pueden corregirse con la función «buscar y reemplazar» de Word y determinadas expresiones regulares. No obstante, mientras que la digitalización de un libro de 250 páginas puede estar lista en una hora, la revisión

⁷⁴ Las características de este formato se encuentran recogidas en Bowker (2002: 40).

⁷⁵ Eder (2013: 607) comenta que «The most frequent confusions will certainly include the following pairs: m/in, l/1, h/li, c/e, and so forth». En nuestros textos también hallamos problemas con el par "m/rn".

toma bastante tiempo más. Un error bastante recurrente que surgió en algunos de los textos fue el reconocimiento de los puntos suspensivos como dos puntos seguidos y un punto final de frase, por ejemplo. Por último, los textos se pasaron por el corrector ortográfico y gramatical de Microsoft Word y se guardaron en formato TXT (sólo texto sin formato)⁷⁶ para posibilitar su reutilización en futuros proyectos y su compatibilidad con la mayoría de las herramientas de procesamiento y análisis lingüístico existentes.

Estos archivos, sin ningún otro tipo de procesamiento posterior, como el que vemos en la imagen 8, son los que se utilizaron con la primera versión del corpus (fase I) en el primer análisis de las perífrasis.

```
RITOS de MUERTE
Alicia Giménez Bartlett
1
    Algún tiempo después de mi segunda separación me empeñé en encontrar una
casita con jardín en la ciudad. Un objetivo difícil, pero lo logré. Era algo
más que un capricho, quise pensar.
```

Imagen 8 – Primera fase del procesamiento de los textos: fragmento de un texto de ES en formato TXT

En la fase II añadimos dos pasos: la alineación de los textos originales en inglés y sus respectivas traducciones al español y el etiquetado morfológico de todos los subcorpus. Explicaremos los pasos que seguimos a continuación.

3.3.3. Alineación

What an ECC-based methodology cannot tell us, however, is why certain patterns occur and how they come about (...). When studying translation as a product entirely in the target language environment, we can only put forward suggestions

⁷⁶ Normalmente, en los libros suele hablarse de este formato como ASCII (American Standard Code for Information Exchange). Según Bowker (2002: 37), este formato puede ser leído y procesado por diversas aplicaciones puesto que fue desarrollado como un estándar reconocido. Sin embargo, Barlow (2003) comenta lo siguiente: «If you are working with English texts, you can use ASCII files, but if you have non-English ASCII files containing accents, etc., then attempts to load them into ParaConc will cause the accented characters to appear in garbled form. You must transform the files into ANSI text format by opening them using a Windows-based word-processor and then saving a copy of the files as Windows (ANSI) text». Esto es justamente lo que hicimos nosotros, dado que en algunos programas, como el ya mencionado ParaConc o TreeTagger, otros formatos nos daban problemas. Zanettin (2012: 74) afirma que ASCII ha quedado ya obsoleto y ha sido reemplazado por Unicode, «a truly multilingual character set which can handle tens of thousands of characters» y que «Different versions of Unicode, or Unicode Transformation Formats (UTFs) have been developed since the 1990s. After reviewing various technical aspects involved in the use of different character encoding formats, McEnery and Xiao (2005) recommend UTF-8 as a suitable format for corpus construction». Este es el formato que se utilizó para la creación del corpus COVALT en el caso de los textos en alemán y en catalán, idiomas que contienen varios signos especiales.

regarding the possible causes that may have led to certain patterns. In order to find an explanation for our results, we would need to construct and analyse in parallel another corpus that would include the source texts of the translational component of ECC. (Laviosa 1998b: 9)

Laviosa se refiere en esta cita al English Comparable Corpus y a la metodología que adoptamos, como explicábamos al principio, en nuestro trabajo de máster. Para pasar de un estudio puramente descriptivo y comparativo a intentar hallar explicaciones a los datos obtenidos o al menos asegurarnos de que las características halladas en los textos traducidos no se deben exclusivamente a la influencia del texto de partida, es necesario contar con dicho texto. Por ese motivo procedimos a la compilación del subcorpus de originales en inglés. No obstante, para poder aprovechar al máximo las ventajas que ofrece el contar con los textos originales, necesitamos poder consultar de manera rápida y simultánea las partes del texto de partida y del texto meta que nos interesan. Aquí es donde entra la fase de alineación del corpus paralelo.

Para que un corpus paralelo funcione, debemos alinear los textos originales y sus traducciones, «that is to produce a link between the two, at the sentence level, word level or some other level» (McEnery y Xiao 2007: 21). De esta manera, podremos localizar todas las apariciones de una determinada palabra o expresión y estudiar las concordancias paralelas para identificar posibles equivalentes en la otra lengua, que, como veremos a lo largo del análisis, puede ser la de partida o la de llegada, dependiendo de cómo configuremos las búsquedas según nuestros objetivos. Como explica Zanettin (2000: 115), «In the first case, a concordance of a source language item or pattern would highlight recurrent translation choices made by the translators; in the second case, a concordance of target language items or patterns would yield as a result the range of source language features from which those items or patterns originated».

De hecho, una vez tenemos identificadas ciertas estructuras equivalentes en ambas lenguas, podemos proceder, de manera cíclica, a analizar otros elementos de la lengua de partida que generen las mismas correspondencias y otros equivalentes de estas correspondencias en los textos de llegada, como sugiere Zanettin (2012: 27), un procedimiento útil en nuestro caso, dado que analizamos elementos que, en principio, no cuentan con un equivalente inmediatamente accesible en la lengua de partida. De hecho, como veremos, en este trabajo hemos utilizado este procedimiento con algunos de los elementos analizados, aunque de manera menos ideal que la que propone este

autor, pues, al no contar con un corpus bidireccional, a veces fue necesario buscar fuera del corpus (en diccionarios, trabajos contrastivos o gramáticas bilingües) las traducciones o equivalentes alternativos en inglés que luego analizaríamos dentro.

Los programas de alineación comparan los textos de partida y de llegada y conectan los fragmentos que consideran equivalentes basándose en distintos métodos. Algunos utilizan estadísticas basadas en el número y la longitud de oraciones, palabras o caracteres y la suposición de que existe una correlación significativa entre estos elementos en originales y traducciones, como por ejemplo Gale y Church (1993) o Mikhailov (2001: 89). Otros analizan la correspondencia de contenido léxico, como Johansson y Hofland (1998: 87), que se basan en un conjunto de *anchor words*, es decir, palabras frecuentes con una correspondencia clara en la otra lengua, o el programa Translation Corpus Aligner, descrito por Izquierdo, Hofland y Reigem (2008: 37-38) para la alineación de P-ACTRES. No obstante, los mejores resultados se obtienen mediante la combinación de ambas técnicas y, de hecho, en estos dos últimos artículos, los autores comentan que, además de puntos de anclaje basados en lexicones bilingües, palabras similares, números, mayúsculas y signos de puntuación, los programas utilizados para la alineación han tenido en cuenta también, de manera secundaria, el número de caracteres de las distintas frases o unidades.

Por supuesto, todos los métodos indicados tienen un porcentaje de fallos y es necesario revisar manualmente la alineación para realizar la corrección de los errores. Esta tarea es bastante pesada, como hemos comprobado de primera mano tanto para la compilación de este corpus como durante la alineación de un conjunto bastante más extenso de textos para el corpus COVALT (Marco 2012: 232). Podemos tardar entre un par de horas y quince, dependiendo de la extensión y el estilo de escritura y de traducción de cada novela. Sin embargo, en nuestra opinión, no es un paso opcional, como puede serlo la revisión del etiquetado (del que hablaremos en un apartado posterior) porque una alineación incorrecta es complicada de subsanar en un corpus indexado y no permite realizar el trabajo cualitativo de comparación de concordancias paralelas.

The high quality needed for descriptive translation research can only be obtained through manual alignment editing, as opposed to corpus-based machine (assisted) translation, which relies on automation and data quantity and for whose purposes automatic alignment techniques provide viable results. This is especially true for

corpus based studies of genres such as fiction and news writing, whose language is often rather “noisy”, that is, resistant to automatic alignment. (Zanettin 2013:30)

El programa utilizado para la alineación es Déjà Vu X2 Professional v. 8.0.585, un sistema de traducción asistida de Atril Solutions⁷⁷ que permite la creación de proyectos de alineación para alimentar memorias de traducción. En Utko (2004: 200-201) se explica que algunos programas de alineación —en ese caso, el Vanilla Aligner— requieren cierto grado de preprocesamiento, como la inclusión de marcas de párrafo y oración. En nuestro caso, los textos se introdujeron en el programa sin marcas de este tipo, lo que ahorra el tiempo de preprocesamiento pero probablemente genera más errores de alineación. Los delimitadores de oraciones se seleccionan antes de la alineación, y para cada lengua podemos seleccionar los que más convengan.

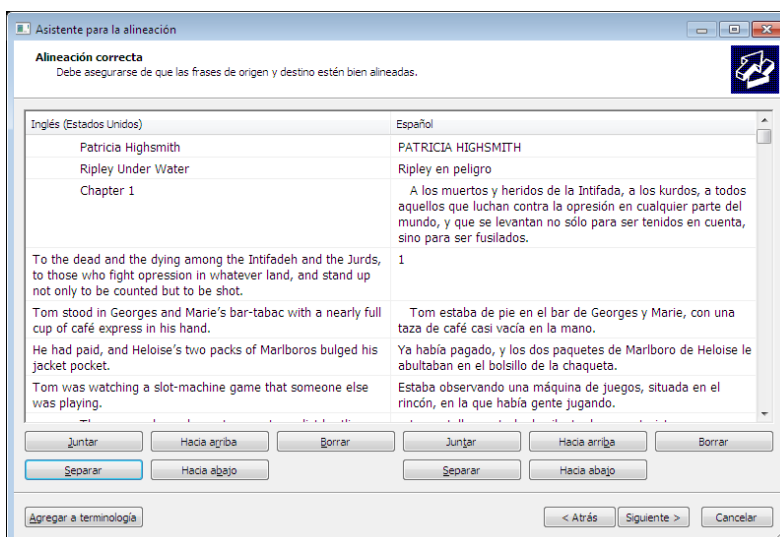


Imagen 9 – Asistente para la alineación de Déjà Vu: fragmentos desordenados

Una vez hemos seleccionado las lenguas de partida y de llegada y los textos que queremos alinear, se crea el proyecto de alineación y se abre el asistente de alineación, del que presentamos una captura en la imagen 9.

⁷⁷ Disponible en <http://www.atril.com/>

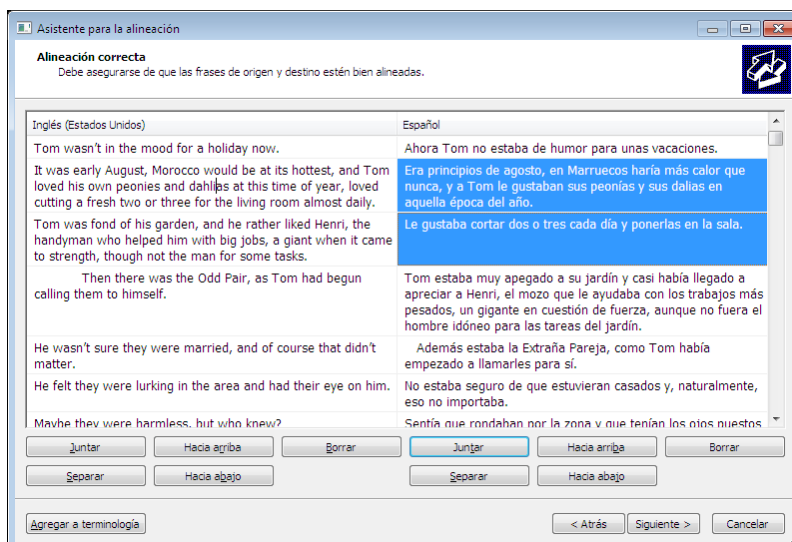


Imagen 10 - Asistente para la alineación de Déjà Vu: fragmento dividido

El error de alineación que observamos en la imagen 10 puede subsanarse mediante los botones «hacia arriba» y «hacia abajo», que permiten cambiar el orden de los fragmentos seleccionados.

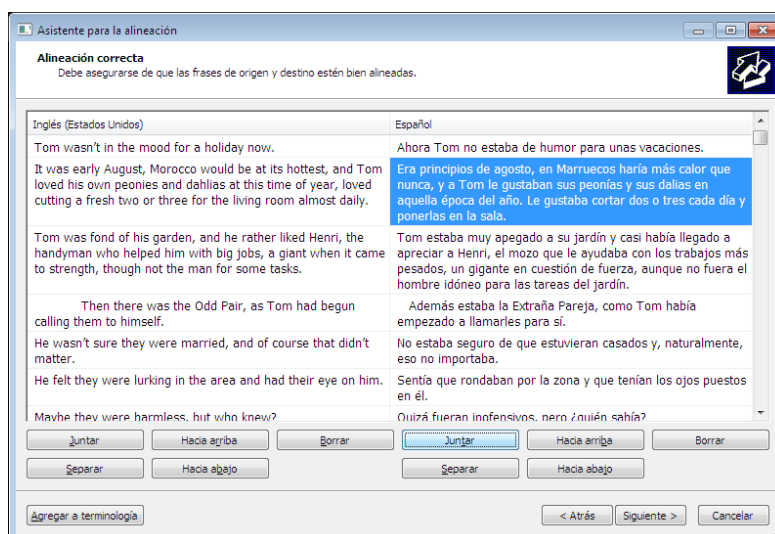


Imagen 11 - Asistente para la alineación de Déjà Vu: fragmento unido

Haríamos lo mismo si fuera a la inversa, es decir, si el traductor hubiese unido en una misma frase dos del texto original, pues tomaríamos como punto de referencia el texto en el que hay una sola unidad y alinearíamos esa unidad con todo lo que le corresponda en el otro texto, como en los ejemplos siguientes:

- 1) “Reeves here,” said Reeves Minot on a clear connection.
Aquí Reeves —dijo Reeves Minot. Se le oía con bastante claridad.

- 2) This morning, believe it or not, an invitation came from the Pritchards—you know, the American pair. For us to come for a drink around six tonight.

—Esta mañana, aunque te parezca increíble, ha llegado una invitación de los Pritchard para que vayamos a tomar una copa allí hacia las seis.

Como comenta Utka (2004: 201), la segmentación puede hacerse siguiendo un criterio más o menos estricto: «In many cases, it is reasonable to treat a semi-colon or a colon as the end of a sentence». Por otra parte, hay que tener cuidado con los puntos que cumplen otros objetivos que no son el de marcar el final de una oración (Gale y Church 1993: 79), como en el caso de las cifras y las abreviaturas (p. ej. «Mme. Annette»), pues el programa puede interpretar que debe dividir el fragmento en ese punto de la oración.

Hay ocasiones en las que a un fragmento en una de las columnas no corresponde ninguno en la otra columna. Esto sucede cuando el traductor se ha dejado algo por traducir involuntariamente o ha realizado una omisión intencionada o, por el contrario, ha añadido material textual por alguna razón. El primer impulso es dejar un espacio en blanco en el fragmento correspondiente de la columna opuesta. Sin embargo, el programa no interpreta lo que nosotros queremos, y elimina el par de fragmentos en cuestión al terminar la alineación. Por tanto, lo que hay que hacer es añadir algún signo de puntuación o símbolo para que Déjà Vu no elimine dichos fragmentos. Lo primero que se nos ocurrió y que de hecho, implementamos para este corpus en su momento, fue marcar los fragmentos en blanco con puntos suspensivos. Sin embargo, si tuviésemos que modificar o ampliar el corpus hoy en día, no sería esta la decisión que tomaríamos, pues los puntos suspensivos son un signo de puntuación que puede interesarnos estudiar en determinado momento, y esta manera de marcar las omisiones o adiciones generaría ruido innecesario⁷⁸. La mejor opción sería un asterisco (*), que se etiquetaría como símbolo y no interferiría en las búsquedas que pudiésemos hacer sobre puntuación en el corpus.

Otra cuestión que gestionaríamos de manera diferente si tuviéramos que compilar nuevamente el corpus sería la de las notas a pie. Dada nuestra falta de experiencia investigadora en el momento de iniciar la creación del corpus, ciertos factores —el hecho de que ni los originales en español ni en inglés las contenían, en las

⁷⁸ Afortunadamente, no interfiere en los cómputos de palabras, pues WordSmith no cuenta como tokens los signos de puntuación: «A sequence of only 'letter' characters delimited by space or punctuation will be treated as a 'word' and counted separately», en palabras de Mike Scott citadas en un correo de Laurence Anthony de fecha 11/28/11 al foro público de debate de Google Groups AntConc-discussion. Recuperado el 11 de septiembre de 2015 de https://groups.google.com/forum/#!msg/antconc/IZ_BT7zdP0w/HX-AVD9xXq0J

traducciones apenas había alguna, y, a diferencia de las omisiones o adiciones, las notas no nos interesaban tanto para el análisis de elementos únicos o típicos, ya que, en principio, estos no tendrían por qué aparecer en las traducciones ni, por tanto, estar explicados con una nota del traductor—, decidimos no preocuparnos por el tema y omitirlas de los textos editados, como hicimos, por ejemplo, con los números de página. Ahora, por supuesto, las conservaríamos, ya que pueden resultar una información muy valiosa a la hora de analizar el proceso traductor. El sistema que adoptaríamos sería copiarlas dentro del cuerpo del texto, entre paréntesis, detrás de la llamada, que es el que realizamos para la alineación de los textos del corpus COVALT y el que mencionan Frankenberg-García y Santos (2003: 4): «The notes themselves are inserted at the point where their identifiers appear».

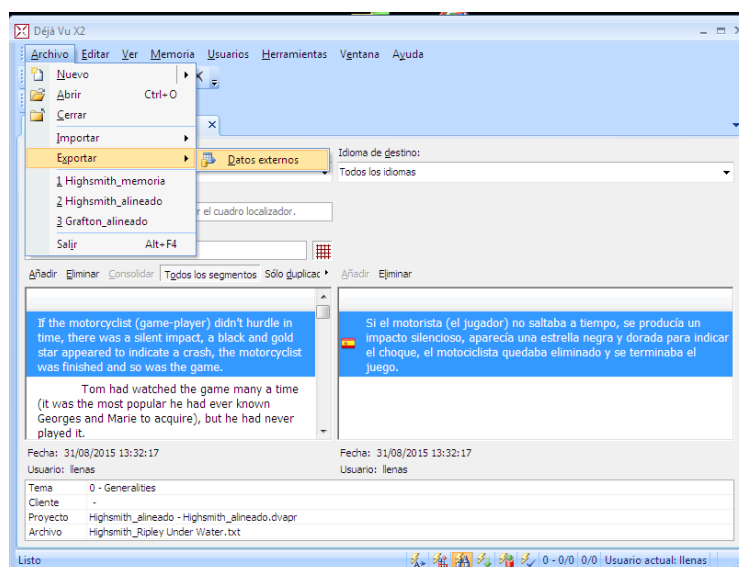


Imagen 12 – Exportación de una memoria de traducción de Déjà Vu

Cuando terminemos la revisión de la alineación, el programa nos pedirá guardar los segmentos alineados en una memoria de traducción o, si esta no existe, crearla. Los archivos con los que nos quedaremos al finalizar este proceso serán el del proyecto alineado (DVAPR) y los cuatro de la memoria de traducción: dos DVMDI, un DVMDX y un DVMDB. Este último es el que deberemos abrir para exportar los segmentos alineados guardados en la memoria a un archivo de texto, como se observa en la imagen 12. Para que dicha exportación tenga éxito, deberemos seleccionar el formato al que queremos exportar (TXT), los campos («segmento principal» y «traducción»), la combinación de lenguas y el delimitador, es decir, el signo que separará el fragmento

original del traducido (nosotros elegimos el punto y coma). El resultado final será el siguiente:

```
"Source";"Target"  
"\"L\" is for LAWLESS";"L DE LEY"  
"Sue Grafton";"SUE GRAFTON"  
"Chapter 1";"Capítulo 1"  
"I don't mean to bitch, but in the future I intend to hesitate before I do a  
favor for the friend of a friend. "; " No es por quejarme, pero en lo sucesivo  
me lo pensaré dos veces antes de hacer un favor a los amigos de los amigos. "  
"Never have I taken on such a load of grief. ";"Jamás lo he lamentado tanto. "  
"At the outset, it all seemed so innocent. "; "Al principio, todo parecía de lo  
más inocente. "
```

Imagen 13 – Segunda fase del procesamiento de los textos: fragmentos de EN y TRAD alineados en un único archivo

3.3.4. Segmentación

Para poder ser etiquetados y utilizados en programas de concordancias como ParaConc⁷⁹ o de gestión de corpus como Corpus Workbench⁸⁰, los archivos alineados necesitan ser segmentados con anterioridad. Segmentar quiere decir dividir por lenguas los fragmentos alineados (que, recordemos, se encuentran ahora en la misma línea separados mediante un punto y coma) sin perder la información de la alineación, que indica la correspondencia de un fragmento del texto original con otro del texto meta. No olvidemos que solo podemos etiquetar en una única lengua por archivo. Por otro lado, los programas de concordancias están preparados para acceder a dos archivos diferentes a la hora de trabajar con textos paralelos.

Este proceso de segmentación se ha realizado de manera automatizada mediante un script⁸¹, o secuencia de comandos, que hemos aplicado a cada texto alineado por Déjà Vu (cambiándole la extensión en primer lugar de TXT a ALG). Este script, que

⁷⁹ Un programa de concordancias para textos paralelos desarrollado por Michael Barlow y distribuido por Athelstan (Barlow 2003). En este trabajo utilizamos la versión 1, compilación 269 (2004).

⁸⁰ *The IMS Open Corpus Workbench* es un software desarrollado por Oliver Christ y Bruno Schulze en el marco del proyecto TC del Instituto de Lingüística Computacional de la Universidad de Stuttgart Disponible en <http://cwb.sourceforge.net/>

⁸¹ Agradecemos a Carme Colominas, del grupo GLICOM de la Universitat Pompeu Fabra, su ayuda durante esta etapa de la compilación del corpus. En marzo de 2012 impartió un taller en la Universitat Jaume I sobre preprocesamiento y explotación de corpus y la utilización del Corpus Query Processor (CQP), invitada por el grupo COVALT, y nos facilitó la primera versión del script ([separab3.pl](#)) con el que pudimos empezar a segmentar nuestros textos. Posteriormente, sin embargo, hubo que cambiar un poco el script para que no añadiera el nombre del archivo y el número de línea o segmento a cada etiqueta de apertura de segmento (p. ej. <s__Grafton_alineado-1>), pues, de lo contrario, había que eliminar esta información antes del etiquetado.

aplicamos mediante el programa Strawberry Perl⁸², ejecuta una serie de instrucciones cuyo resultado final es la generación de dos archivos individuales, uno con el texto original en inglés (extensión TP) y otro con el texto traducido al español (extensión TA). A continuación reproducimos los comandos necesarios para modificar la extensión de uno de nuestros archivos y que Strawberry Perl pueda leerlo (a) y para que, a partir del mismo, se generen los archivos TP y TA (b).

```
(a) C:\Strawberry\perl\bin>rename Grafton_alineado.txt Grafton_alineado.alg
(b) C:\Strawberry\perl\bin>perl separab3.pl Grafton_alineado
```

La diferencia con los textos de los que disponíamos antes de comenzar el proceso de alineación es que estos archivos contienen ahora información sobre cómo debe agrupar el programa de concordancias un determinado fragmento de TP con el fragmento correspondiente de TA. Nos referimos a una información muy básica: cada segmento en el que se fragmentó el texto original durante el proceso de alineación se encuentra en una línea distinta del archivo TP, y el segmento correspondiente del texto meta se encuentra en la misma línea del archivo TA, en ambos casos, precedidos por una etiqueta de apertura de frase (<s>) y seguidos por una etiqueta de cierre (</s>).

```
#TEXT TT3 Sue-Grafton L-de-ley

<s>
Capítulo 1
</s>
<s>
  No es por quejarme, pero en lo sucesivo me lo pensaré dos veces antes de
  hacer un favor a los amigos de los amigos.
</s>
<s>
  Jamás lo he lamentado tanto.
</s>
<s>
  Al principio, todo parecía de lo más inocente.
</s>
-----
#TEXT ST3 Sue-Grafton "L"-is-for-LAWLESS

<s>
Chapter 1
</s>
<s>
I don't mean to bitch, but in the future I intend to hesitate before I do a
favor for the friend of a friend.
```

⁸² Una distribución del lenguaje de programación Perl (creado originalmente para el sistema operativo Unix) que permite trabajar con este lenguaje en el entorno de Windows. Disponible en <http://strawberryperl.com/>

```
</s>  
<s>  
Never have I taken on such a load of grief.  
</s>  
<s>  
At the outset, it all seemed so innocent.  
</s>
```

Imagen 14 – Tercera fase del procesamiento de los textos: fragmentos de EN y TRAD segmentados en dos archivos

3.3.5. Etiquetado

Teniendo en cuenta las necesidades específicas del estudio, estos textos pueden etiquetarse y lematizarse para, durante la etapa de análisis de los datos, poder recuperar información morfosintáctica de cada elemento lingüístico y buscar, además de elementos léxicos, categorías y estructuras sintácticas o secuencias de determinadas categorías (por ejemplo, todos los sustantivos precedidos por adjetivos de un corpus). Pero este no es el único tipo de etiquetado que existe. Perea (2005: 1) define el proceso de etiquetado de la siguiente manera:

Etiquetar un texto consiste en incorporarle información, normalmente acerca de los elementos que lo componen. Esta información puede referirse tanto al texto como conjunto –autor del texto, fecha de publicación, tipología textual– como a los distintos elementos del mismo (por ejemplo, diferenciación de los elementos de un texto: título, resumen, cuerpo, etc.).

Sin embargo, los tipos de etiquetado más habituales son los que se realizan sobre las palabras que conforman el texto (Macken, De Clercq y Paulussen 2011: 383): el etiquetado morfológico (o *part-of-speech tagging*)⁸³, que asigna a cada palabra — también a los signos de puntuación— una etiqueta con la categoría lingüística a la que pertenecen (p. ej. sustantivo o adjetivo) y el lematizado, que le asigna su lema o forma básica correspondiente (lo que permite buscar todas las formas conjugadas de un verbo sin necesidad de introducirlas una a una)⁸⁴. Esta información se añade al texto, unida a

⁸³ En inglés, McEnery, Xiao y Tono (2006: 22, 29) distinguen entre *corpus mark-up* (la adición de información sobre el texto, como la que describe Perea) y *corpus annotation* (la adición de información lingüística).

⁸⁴ A veces, como veremos también en el caso del etiquetado morfosintáctico, la lematización no es deseable, pues puede ocultar las frecuencias de las distintas formas lematizadas, que a veces muestran comportamientos o significados muy distintos entre sí (Utko 2004: 206).

la palabra mediante un guión bajo u otro símbolo, o separada de esta mediante una marca de tabulación.

Esta información puede añadirse hoy en día de manera automática mediante programas informáticos conocidos como etiquetadores. La ventaja es que el proceso se realiza de manera casi instantánea; la desventaja es que las etiquetas que se aplican vienen predeterminadas. De hecho, esa es una de las críticas que se realizan a la idea de anotar los textos con información lingüística: la clasificación siempre vendrá dada por el lingüista que la realice, y a los usuarios del corpus se les impondrán estas categorías.

Por este motivo, hay autores que abogan por dejar el texto «as it is» (Sinclair 1991: 21, 29; Tognini-Bonelli 2001: 74). De esta manera, se evita el riesgo de asumir modelos lingüísticos heredados que no se correspondan con la evidencia obtenida de los corpus. El etiquetado siempre podrá añadirse para estudios específicos que lo requieran y cada investigador podrá utilizar los códigos lingüísticos que considere adecuados.

No obstante, también hay defensores del proceso de etiquetado (McEnery, Xiao y Tono 2006: 30), dado que aporta valor añadido a la información proporcionada por los textos, facilita la explotación del corpus en determinadas cuestiones (como la extracción de patrones gramaticales concretos), el resultado del proceso se aprovecha, pues puede ser reutilizado en proyectos posteriores para fines completamente distintos y, en respuesta a la opinión de que el lingüista impone su punto de vista subjetivo mediante las etiquetas, algunos creen que «la anotación es un excelente medio de dejar constancia de una visión sobre el lenguaje, lo que es un claro beneficio frente a la subjetividad» (Villayandre Llamazares 2010: 53).

Zanettin (2013: 30) opina que el debate no tiene demasiado sentido, pues «Linguistic annotation may not be necessary or even useful for many types of investigations, but it can be extremely practical and sometimes necessary when looking for regularities beyond the lexical level»; y, por otra parte, como veremos más adelante, normalmente es posible consultar un corpus anotado visualizando las etiquetas o no, dependiendo de nuestras necesidades.

Por lo tanto, como explican Zanettin (2012: 79) y Villayandre Llamazares (2010: 53), existen múltiples niveles de anotación:

- ✓ Textos *raw* o sin ningún tipo de etiquetado, que pueden explotarse perfectamente mediante distintas herramientas de análisis textual como WordSmith Tools.

- ✓ Adición de un encabezado con información extralingüística y metatextual, como pueden ser «título de la obra», «traductor», «lengua» o «género». Incorporar estos datos puede resultar útil para realizar búsquedas solo en determinadas secciones de un corpus extenso (Zubillaga, Sanz y Uribarri 2015: 82).
- ✓ Marcado de las distintas partes del texto (p. ej. párrafos o capítulos).
- ✓ Marcado de convenciones tipográficas como cursiva o itálica.
- ✓ Información lingüística sobre lema, función sintáctica y morfología (*POS tagging* o anotación morfosintáctica).
- ✓ Anotación sintáctica (también llamada *parsing*) sobre categorías que reflejan la estructura de la oración (p. ej. «frase preposicional»).
- ✓ Anotación semántica o adición de información sobre el significado de las palabras o su pertenencia a un determinado campo semántico.
- ✓ Otros tipos de anotación para fines específicos, como el marcado de recursos estilísticos o el diálogo en las novelas.

En cuanto al etiquetado sobre características del texto en conjunto —lo que McEnery, Xiao y Tono denominan *corpus mark-up*—, en la actualidad existen varios estándares de codificación, como TEI (Text Encoding Initiative) o CES (Corpus Encoding Standard), un formato simplificado de la TEI diseñado específicamente para corpus. Villayandre Llamazares (2010: 47) explica que en la TEI, los elementos codificados se marcan con etiquetas de inicio y cierre, como vimos que hacía el script de segmentación en nuestros archivos:

Así, los textos se conciben como documentos que se estructuran en dos partes: cabecera (head), donde se incorpora de forma obligatoria la información bibliográfica completa del archivo electrónico y, opcionalmente, otro tipo de informaciones (como fecha de los cambios, lenguas del texto, etc.); y cuerpo (body) o texto propiamente dicho, en el que se codifican básicamente elementos estructurales (capítulos, párrafos, oraciones, palabras).

Este estándar se expresa normalmente en lenguaje XML (Extensible Markup Language)⁸⁵, como acabamos de ver, pero puede convertirse fácilmente a otros

⁸⁵ Según Pes (2014-2015), «XML (*eXtensible Markup Language*, Lenguaje de Marcado eXtensible) es un lenguaje desarrollado por W3C (*World Wide Web Consortium*) que está basado en SGML (*Standard Generalized Markup Language*, Lenguaje de Marcado Generalizado Estándar)». Se trata de un metalenguaje que sirve para diseñar

formatos de anotación como el requerido por otras herramientas como Corpus Workbench, es decir, una palabra por línea y etiquetas separadas por tabulaciones (Zanettin 2012: 96).

Nuestro corpus, sin embargo, solo contiene unos cuantos atributos estructurales en forma de encabezados con algunos datos contextuales como el código identificador del texto, el nombre del autor y el título de la novela, y, por supuesto, las marcas de inicio y final de frase, texto y corpus. No hemos marcado capítulos ni otras subdivisiones internas pues en su momento no nos parecieron relevantes para el tipo de análisis que queríamos realizar (y porque, dado el reducido tamaño del corpus, pensamos que siempre tendríamos los originales a mano para consultar este tipo de información); pero, si tuviéramos que hacer público el corpus o simplemente ampliarlo, sería interesante aumentar la cantidad de datos incluidos, pues este tipo de información adicional puede resultar útil, como decíamos anteriormente, para realizar búsquedas solo en determinadas partes del corpus, o porque, como afirma Laviosa (1998b: 9), determinados atributos extratextuales como el género del traductor o la dirección de traducción son variables que pueden estar relacionadas con las características de los textos traducidos, y, por tanto, tienen que estar disponibles para los posibles usuarios futuros del corpus.

Otro motivo por el cual probablemente no nos atrevimos en su momento a adentrarnos en este tipo de anotación es la complejidad del estándar en sí. Hardie (2014a) defiende un estándar TEI más «modesto» para corpus pequeños como el nuestro cuyos fines se limitan a investigaciones individuales. Hardie lo explica de la siguiente manera (2014a: 76):

The TEI and CES are both products of their time. In the period from 1987 to 1996, corpus creation was not a task to be undertaken lightly. Collecting any sizeable amount of text in electronic form was still fraught with difficulties. This being the case, it was by and large safe to assume that the ‘average’ corpus would be widely distributed and re-used. Therefore, dense and careful markup of as many potentially useful features as possible was highly desirable – since the requirements of the corpus user could not be precisely predicted. Moreover, given the scale of corpus building projects, it was also safe to assume that any corpus construction

lenguajes de marcado para fines específicos mediante etiquetas definidas por el usuario, que se escriben utilizando los signos «menor que» (<), «mayor que» (>) y la barra inclinada (/).

team would have access to personnel with a substantial degree of technical knowhow: computer scientists, or at least programmers.

Hoy en día, sin embargo,

A much larger number of corpora are created by individual researchers for their own research interests, with no expectation that the resource will ever be used by anyone else. First, there is no especial need for any markup directed at the speculated requirements of end-users, since there is no end-user other than the corpus creator. Second, when a corpus is created by a single linguist, the assumption of a high degree of technical expertise related to computers is by no means always met.

La propuesta «Modest XML for Corpora» se describe en este artículo, y es algo que no conocíamos en el momento de crear nuestro corpus pero que recomendamos para investigadores noveles en este campo, pues es un estándar mucho más sencillo y accesible que puede servir de base para, posteriormente, empezar a realizar un marcado más sofisticado. Según estas sugerencias, aparte de lo que ya tenemos hecho, solo nos faltaría marcar los párrafos, los saltos de página y los elementos omitidos del texto, como imágenes, y, a partir de estos elementos básicos, todo lo que nos pueda parecer relevante.

Como hemos mencionado anteriormente, un corpus cuyas palabras han sido lematizadas o etiquetadas con información morfosintáctica ofrece la posibilidad de realizar búsquedas más precisas, no sólo de unidades léxicas sino también de patrones gramaticales. Teniendo en mente la clase de elementos con los que íbamos a trabajar, la opción de etiquetar y lematizar el corpus parecía bastante razonable, ya que para encontrar todas las apariciones de una perífrasis, por ejemplo, hay que asegurarse de incluir todas las flexiones del verbo auxiliar en la ventana de búsqueda y esto, además de tedioso, a veces no es posible en algunos programas como WordSmith, que no nos permitía incluir más de quince alternativas en las expresiones de búsqueda. Además, es difícil buscar categorías específicas como «infinitivo» o «gerundio».

El etiquetado morfológico de nuestro corpus ha sido realizado mediante el etiquetador TreeTagger⁸⁶ (Schmid 1994, 1995)⁸⁷, una aplicación gratuita desarrollada por el Instituto de Lingüística Computacional de la Universidad de Stuttgart que anota con etiquetas morfosintácticas y lematiza. Es independiente de la lengua a la que se aplique pero viene con ficheros de parámetros para distintas lenguas y un módulo de entrenamiento. Nosotros hemos utilizado la versión para Windows, complementada con una interfaz gráfica creada por Ciarán Ó Duibhín⁸⁸.

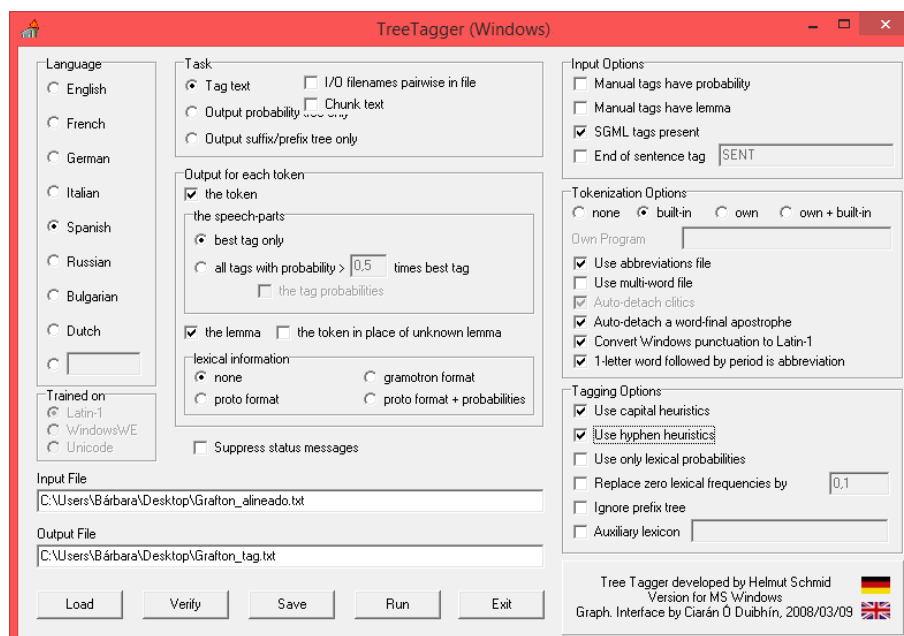


Imagen 15 – Interfaz para Windows del programa de etiquetado morfológico y lematización TreeTagger

Como vemos en la captura de la imagen 15, la interfaz es muy sencilla de comprender y utilizar. Hay que recordar seleccionar el idioma correcto en la columna de la izquierda y cambiar la extensión del archivo que queremos etiquetar de TP o TA a TXT, para que TreeTagger pueda procesarlo. Marcando la opción «SGML tags present», TreeTagger no etiqueta las etiquetas <s> y </s>, lo que ahorra trabajo de revisión posterior. Si los textos contuvieran encabezados (en nuestro caso, los tienen: «#TEXT TT3 Sue-Grafton L-de-ley»), se etiquetarían también, por lo que habría que

⁸⁶ Software (etiquetador morfosintáctico y lematizador) desarrollado por Helmut Schmid en el marco del proyecto TC del Instituto de Lingüística Computacional de la Universidad de Stuttgart. Disponible en <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>

⁸⁷ Antes de decidimos por TreeTagger, realizamos pruebas con otros programas de etiquetado, como PetraTag (Perea Sardón 2005) y Connexor Machine Syntax (disponible en <https://www.connexor.com/nlplib/?q=demo/syntax>). En nuestro trabajo de máster explicamos los motivos por los cuales finalmente decidimos no utilizarlos.

⁸⁸ Los archivos y las instrucciones de cómo instalar TreeTagger desde el principio para que funcione con esta interfaz se encuentran en <http://www.smo.uhi.ac.uk/~oduibhin/oideasra/interfaces/wintntinterface.htm>

borrarlos antes y volver a añadirlos manualmente después. El formato de salida es el texto verticalizado y la información morfológica (segunda columna) separada de la palabra (primera columna) y del lema (tercera columna) por tabulaciones, como se ve a continuación:

```
<text id="TT3" author="SueGrafton" title="Ldeley">
<s>
Capítulo NC      capítulo
1 CARD  1
</s>
<s>
No NEG      no
es VSfin   ser
por        PREP   por
quejarme  VLinf   quejarme
,         CM
pero      CCAD   pero
en PREP   en
lo ART    el
sucesivo ADJ    sucesivo
me PPX    yo
lo PPC    él
pensaré  VLfin   pensar
dos      CARD   dos
veces    NC      vez
antes    ADV    antes
de CSUBI de
hacer    VLinf   hacer
un ART   un
favor   NC      favor
a PREP  a
los     ART    el
amigos  NC      amigo
de PREP de
los     ART    el
amigos  NC      amigo
.       FS
</s>
```

Imagen 16 – Cuarta fase del procesamiento de los textos: fragmento de TRAD etiquetado con TreeTagger

TreeTagger cuenta con distintos conjuntos de etiquetas (*tagsets*) para cada idioma, por lo que las etiquetas no siempre coincidirán en inglés y en español, ni tampoco la información etiquetada. Por ejemplo, «PDEL» es una etiqueta exclusiva del español, puesto que sirve para marcar la combinación de la preposición *de* y el artículo *el* (*del*).

El problema del etiquetado automático es que nunca puede ser preciso al cien por cien. Izquierdo, Hofland y Reigem (2008: 39), que también utilizaron TreeTagger para etiquetar el corpus P-ACTRES, comentan que el programa alcanza, en su caso, una precisión del 95 % (en su caso no hubo revisión posterior por falta de tiempo y recursos). Perea (2005: 3) señala que, aunque muchos etiquetadores alcanzan un 97 %

de precisión, ello «deja un margen de error del 3% que, aplicado a una oración media de unas 30 palabras, implica estadísticamente que cada oración incluye una palabra mal etiquetada». Bowker (2002: 69) comenta, en la misma línea, que un 95 % de precisión es razonable para muchos objetivos pero «for sentences containing 14 or more words, the level of accuracy is typically less than 50 percent».

De manera general, los errores pueden deberse a casos de homografía (por ejemplo, *meta* puede ser tanto sustantivo como verbo conjugado), pero también puede ser simplemente porque el programa no conozca la palabra. En estos casos, en la columna del lema aparece la indicación «unknown» o la misma palabra, dependiendo de los parámetros que hayamos especificado al etiquetar, lo cual, como veremos en los capítulos de análisis cuando expliquemos los problemas de etiquetado, puede ser muy problemático, dado que al buscar una palabra por su lema, estos casos no aparecerán y no nos enteraremos de ello. Por poner un par de ejemplos hallados en nuestro corpus, el programa etiqueta la palabra *ocupé* correctamente como «VLfin» (verbo léxico conjugado), pero no es capaz de asignarle el lema *ocupar*, con lo que las búsquedas que realicemos sobre ese lema no incluirán esta forma. Un error del primer tipo es el etiquetado del sustantivo *matrimonio* como «VLfin» y su lematizado como *matrimoniar*.

Sin embargo, McEnery, Xiao y Tono (2006: 75-76) no parecen darle mucha importancia a los errores, siempre y cuando el tamaño del corpus sea suficiente para neutralizar su efecto: «With the same precision rate of annotation, a corpus of one hundred thousand words is clearly more reliable than a corpus containing merely a few thousand words, assuming that the errors are relatively random». Corpas Pastor (2008: 220), por ejemplo, tampoco corrige el etiquetado de sus textos, pero lo justifica asimismo por la gran cantidad de datos con los que trabaja. Izquierdo, Hofland y Reigem (2008: 39) resumen así su postura: «There is, today, no program which is 100 percent accurate, so it seems more practical to manage with this (95 percent accurate) program, than do nothing but wait for a perfect tool».

Nosotros consideramos que, dado el reducido tamaño del corpus y nuestra pretensión de exhaustividad, la cantidad de errores que nos fuimos encontrando no era aceptable. Por supuesto, sería posible buscar y borrar las etiquetas equivocadas o corregir las palabras desconocidas con funciones como «buscar y reemplazar» de algunos editores de texto o programas específicos como InfoRapid Search and Replace. No obstante, como veremos en el análisis de los elementos, hemos ido detectando los

errores a medida que lo llevábamos a cabo. El tiempo que hubiésemos perdido deteniendo el análisis y corrigiendo todos los errores del corpus no estaba justificado, pues nuestro corpus solo estaba pensado para realizar búsquedas de estos tres tipos de elementos en concreto. Para trabajos futuros nos planteamos, con más tiempo, proceder a una revisión y corrección exhaustiva del etiquetado para poder aprovecharlo en la búsqueda de patrones gramaticales. De esta manera, no hemos podido explotarlo al máximo en el análisis, pero nos ha servido para conocer el programa, sus fortalezas y puntos flacos.

3.3.6. Indexación

En este apartado explicaremos los pasos que seguimos para poder introducir nuestros textos en el servidor y consultarlos mediante el paquete de herramientas de procesamiento, gestión y análisis textual IMS Open Corpus Workbench (CWB). Corpus Workbench (Evert y Hardie 2011) puede gestionar corpus de hasta dos mil millones de palabras con anotaciones lingüísticas. Más adelante hablaremos de su componente de búsquedas, Corpus Query Processor (CQP).

Este trabajo comenzó con programas de concordancias independientes, primero WordSmith para las búsquedas en los textos del corpus comparable, y, con la adición del componente paralelo, también ParaConc. No obstante, gracias a nuestra colaboración con el grupo COVALT de la Universitat Jaume I, que trabaja en la creación y explotación de un corpus de literatura traducida al valenciano, tuvimos acceso a CWB y a formación sobre su utilización⁸⁹. Nos pareció una herramienta más flexible y poderosa, aunque requiere un período de aprendizaje más amplio que otras y un administrador con conocimientos de informática para gestionar el servidor, por lo que para un usuario individual puede resultar menos accesible.

Si trabajamos con un programa de concordancias autónomo, como WordSmith o ParaConc, el paso de la indexación no es necesario. Para aclarar lo que significa indexar un corpus, nos remitimos a la explicación de Bowker (2002: 53):

⁸⁹ Agradecemos a Adriano Ferraresi, de la Escuela Superior de Lenguas Modernas, Traductores e Intérpretes de la Universidad de Bolonia (sede de Forlì), la ayuda que nos prestó durante el taller impartido para el grupo COVALT en diciembre de 2012 y su entusiasta colaboración en la resolución de todas las dudas y problemas que nos surgieron acerca del proceso de etiquetado e indexación de nuestros respectivos corpus.

Some concordancers operate by searching through the entire corpus from beginning to end every time a search pattern is entered. Others work by first creating an index of all the words in the corpus along with a record of the location of each occurrence (e.g., line number). The index must be created prior to conducting any searches in a given corpus, but once created, it can be consulted during all subsequent searches for particular items in that corpus.

Si el corpus no estuviera indexado y fuese muy extenso, las búsquedas llevarían mucho tiempo, sobre todo en sistemas que, como CWB, funcionan a través de un servidor. La desventaja es que la indexación es un proceso que se aplica a todo el corpus en conjunto, por lo que habrá que volverlo a llevar a cabo cada vez que decidamos añadir o quitar algún texto o, simplemente, modificarlo.

Como acabamos de mencionar, lo primero que necesitamos para utilizar CWB, que está preparado para trabajar con el sistema operativo UNIX, es un servidor donde instalarlo y un administrador que sepa gestionarlo y otorgarnos permisos de acceso. En segundo lugar, utilizaremos el programa WinSCP⁹⁰, un cliente FTP, SFTP y SCP para Windows, que nos permitirá acceder al servidor (con las credenciales que nos haya proporcionado el administrador) y copiar y trasladar de manera muy intuitiva archivos de nuestro ordenador al servidor y viceversa, no solo durante el proceso de indexación, sino más adelante, cuando realicemos las búsquedas en el corpus y queramos guardar y editar los resultados.

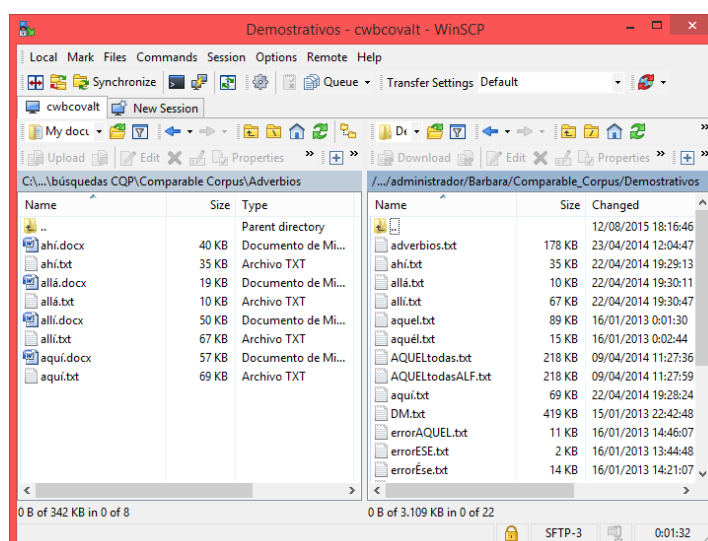
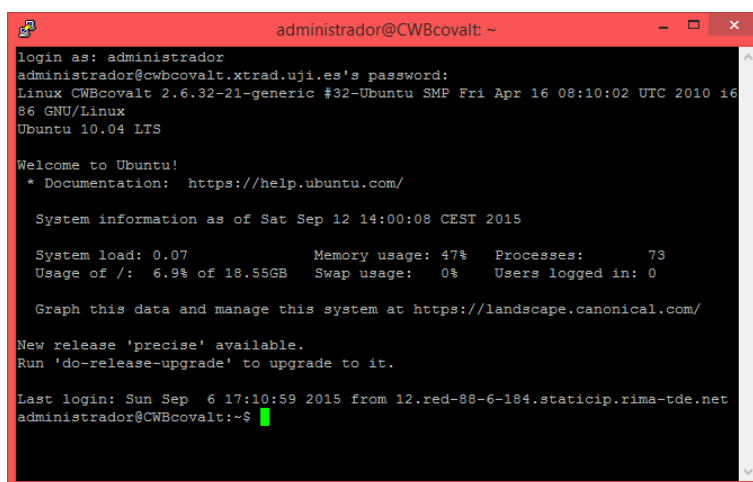


Imagen 17 – Interfaz del programa WinSCP

⁹⁰ Creado por Martin Prikryl. Puede descargarse de la siguiente página: <https://winscp.net/eng/download.php>

En la ventana de la izquierda de la imagen 17 vemos nuestro sistema de carpetas local (en nuestro ordenador) y en la de la derecha, el sistema de carpetas del servidor remoto. Lo único que hay que hacer para transferir archivos de una a otra es arrastrarlas como si se tratara del explorador de Windows.

Para indexar el corpus y, posteriormente, extraer información mediante la herramienta de búsquedas, deberemos también instalar el programa Putty⁹¹. Este programa funciona mediante la introducción de líneas de comando en una ventana similar a la consola cmd o «símbolo del sistema» de Windows, que tiene el aspecto que puede observarse en la imagen 18.



```
administrador@CWBcvalt: ~
login as: administrador
administrador@cwbcvalt.xtrad.uji.es's password:
Linux CWBcvalt 2.6.32-21-generic #32-Ubuntu SMP Fri Apr 16 08:10:02 UTC 2010 i686 GNU/Linux
Ubuntu 10.04 LTS

Welcome to Ubuntu!
* Documentation: https://help.ubuntu.com/

System information as of Sat Sep 12 14:00:08 CEST 2015

System load: 0.07          Memory usage: 47%        Processes:      73
Usage of /:  6.9% of 18.55GB Swap usage:   0%         Users logged in: 0

Graph this data and manage this system at https://landscape.canonical.com/

New release 'precise' available.
Run 'do-release-upgrade' to upgrade to it.

Last login: Sun Sep  6 17:10:59 2015 from 12.red-88-6-184.staticip.rima-tde.net
administrador@CWBcvalt:~$
```

Imagen 18 – Pantalla de acceso a CWB a través del cliente Putty

Los archivos que introduciremos en el servidor, en el formato que es capaz de entender CWB, son los etiquetados con TreeTagger, es decir, con los datos alineados en tres columnas. Sin embargo, antes tendremos que añadir etiquetas de apertura y cierre de texto al principio y al final de cada uno (<text> y </text>) e introducir en la primera etiqueta el encabezado del texto nuevamente en el formato correcto (como explicamos en el apartado dedicado al etiquetado, nosotros solo hemos incluido un nombre de archivo identificativo, el nombre del autor y el título de la novela).

En segundo lugar, tendremos que unirlos en un único archivo TXT por lengua; uno para los textos originales en español y, en el caso del corpus paralelo, deberemos crear uno para el inglés y otro para el español. Para ello utilizaremos el comando «cat *.txt > corpus.txt», en el que el asterisco significa que el programa abrirá todos los documentos que terminan con esa extensión, y el signo > significa que el resultado

⁹¹ Un cliente Telnet y SSH gratuito para Windows y Unix creado por Simon Tatham. Disponible en <http://www.chiark.greenend.org.uk/~sgtatham/putty/>

deberá guardarse en un archivo de nombre «corpus.txt», o el nombre que le queramos asignar a cada conjunto de textos de uno de los componentes monolingües del corpus.

Una vez tengamos estos archivos conjuntos, añadiremos las etiquetas <corpus> y </corpus> al principio y al final de cada documento y ya los tendremos listos para la indexación.

```

<corpus>

<text id="ES1" author="Alicia-Giménez-Bartlett" title="Ritos-de-muerte">

1 CARD 1
Algún QU alguno
tiempo NC tiempo
después ADV después
de PREP de
mi PPO mío
segunda ORD segundo
separación NC separación
me PPX yo
empeñé VLfin empeñar
en PREP en
encontrar VLinf encontrar
una ART un
casita NC casa
conPREP con
jardín NC jardín
en PREP en
la ART el
ciudad NC ciudad
. FS .

(...)

</text>

</corpus>

```

Imagen 19 – Quinta fase del procesamiento de los textos: fragmento de ES preparado para la indexación

Para indexarlos, utilizamos un editor de textos llamado Pico (siempre a través de la línea de comandos en Putty) para adaptar el script de indexación «cwbify_standard_corpus.pl» a los datos de nuestro corpus, con el objeto de que, cuando empecemos con las búsquedas, CWB sepa que puede buscar por los datos que hemos introducido en los encabezados (en nuestro caso, «text_id», «text_author» y «text_title») y lo hacemos ejecutable mediante el comando «chmod +x cwbify_standard_corpus.pl». A continuación, indexamos el primer subcorpus (por ejemplo, el de originales en inglés) mediante un comando que incluya los siguientes datos (aquí reproducimos los de nuestro corpus):

<code>./cwbify standard_corpus.pl -l en</code>	(la lengua del subcorpus)
<code>-d /home/cwbdata_en</code>	(el directorio donde lo vamos a guardar)
<code>-c EN_BAR</code>	(la carpeta que crearemos)
<code>-n "Source Original English Texts"</code>	(una descripción cualquiera del subcorpus)
<code>par_en_bar.tgd</code>	(el nombre del archivo que queremos indexar)

Hacemos lo mismo para el subcorpus de traducciones al español (`./cwbify standard_corpus.pl -l es -d /home/cwbdata_es -c ES_BAR -n "Target texts translated to Spanish" par_trad_bar.tgd`) y alineamos ambos mediante el comando `«cwb-align par_en_bar par_trad_bar s»`. Por último, editamos con Pico el archivo de registro de cada uno de los corpus, (es decir, tendremos que hacerlo dos veces) —al que accedemos tecleando `«pico /usr/local/share/cwb/registry/par_en_bar»`—, escribiendo al final `ALIGNED` y el nombre del archivo con el que queremos que se alinee. Tras realizar este paso en las dos direcciones, el último comando (que también utilizaremos las dos veces que realicemos el paso anterior) es `«cwb-align-encode -D out.align»`. Tras estos pasos ya podremos consultar el corpus en el servidor.

3.4. Análisis del corpus

Además de saber compilar, alinear, etiquetar e indexar (si fuera el caso) un corpus, es muy importante utilizar las herramientas adecuadas para explotarlo. Como ya hemos adelantado, los primeros análisis los realizamos con programas de concordancias independientes (denominados en inglés *concordancers*), más específicamente, con WordSmith Tools y ParaConc, que ya hemos ido presentando a lo largo de las páginas anteriores (v. notas 24 y 34). Dice Zanettin (2012: 116) al respecto de estos programas:

They offer an all-in-one solution to text retrieval and display. Users select their corpus by choosing files and directories, and queries are performed on the fly. Concordancers are readily available (...) and do not require much technical know-how. (...) However, when not working with indexed data, concordancers tend to be slow, hence they perform better with small corpora. Furthermore, while some

concordancers can make some use of in-line annotation, they cannot exploit it to its full potential like corpus management systems...

En la imagen 20 podemos ver a lo que se refiere Zanettin cuando habla de la selección del corpus mediante un sistema de directorios y archivos, como en cualquier programa de Windows.

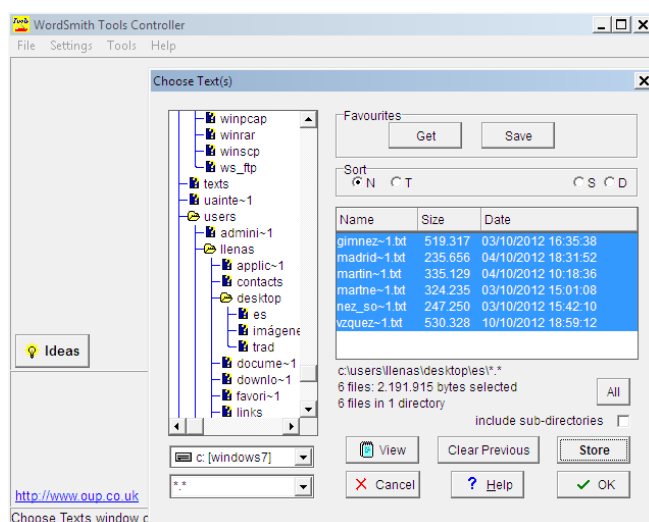


Imagen 20 – Ventana de selección de textos de WordSmith Tools

Por supuesto, la interfaz mucho más intuitiva de estos programas es una ventaja, ya que, como hemos visto, una aplicación como CWB no tiene botones que pulsar y ni siquiera permite utilizar un ratón; todo, incluso el acceso a los diferentes recursos textuales y aplicaciones, se hace tecleando comandos en una ventana. Además, al estar alojados los textos en un servidor remoto, cualquier fallo en la conexión, cambio en un puerto o problema con las credenciales de acceso puede hacer que no podamos utilizarlos (como, de hecho, nos ha sucedido alguna vez).

Con respecto a la facilidad de uso de esta aplicación, tenemos que mencionar que, al parecer, CQP (el componente de CWB utilizado para realizar las búsquedas en corpus) cuenta con un nuevo sistema de análisis a través de una página web interactiva, con una interfaz de usuario muy parecida a la del BNC, más flexible e intuitiva que la que utilizamos actualmente. Según Hardie (2012: 380) «in making a sophisticated query system accessible to untrained users, CQPweb combines ease of use, power and flexibility to a very high degree». Aun así, la figura del administrador, alguien con conocimientos de informática encargado de gestionar los corpus que se suben a la web, no desaparece.

3.4.1. WordSmith y ParaConc

Estos dos programas pertenecen a lo que Hardie (2012: 382) denomina «second generation, and later third-generation tools», es decir, herramientas un poco menos potentes y sofisticadas, pero mucho más accesibles al público en general, creadas para satisfacer la demanda de la mayor parte de lingüistas que no cuentan con amplios conocimientos de informática y programación. Son programas específicos para el análisis de corpus, y una de sus herramientas fundamentales son los buscadores de concordancias. Estos facilitan el acceso a la información contenida en los textos identificando todas las ocurrencias de una palabra o expresión solicitada por el usuario y proporcionando esta información en un listado de muy fácil lectura e interpretación. Por ejemplo, en la imagen 21, vemos la expresión regular introducida en la ventana de búsqueda y en la 22 el resultado de dicha búsqueda en forma de concordancias.

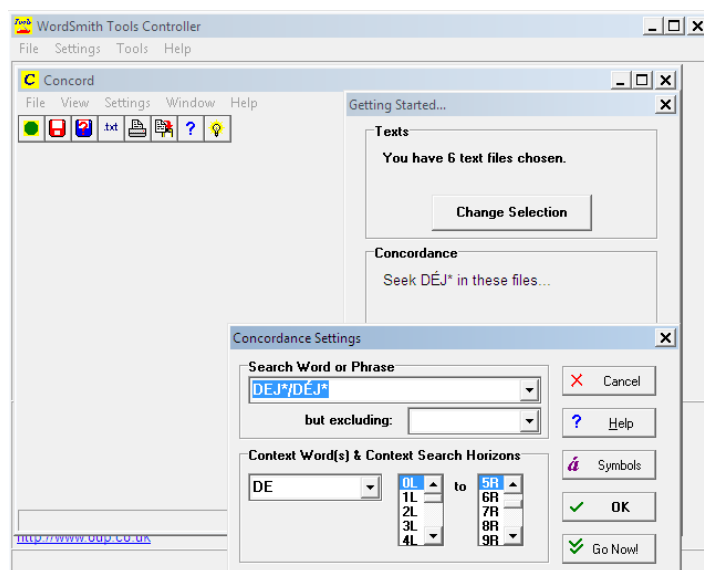


Imagen 21 – Ventana de búsqueda de concordancias de WordSmith Tools

Todas las apariciones de la cadena de caracteres buscada aparecen alineadas una debajo de la otra, en el centro de la ventana, resaltadas en un color distinto y acompañadas a ambos lados por su respectivo contexto, es decir, de las palabras que las preceden y las siguen (la cantidad de palabras o caracteres dependerá de lo que hayamos especificado). Este tipo de visualización se denomina KWIC (o *keyword in context*). El contexto puede ampliarse a varias líneas si es necesario.

The screenshot shows a concordance window with the following columns: N, Concordance, Set, Tag, Word No., File, and %. The text in the concordance column is truncated and includes various uses of the verb 'deja' in different contexts.

N	Concordance	Set	Tag	Word No.	File	%	
1	...de la historia de ayer? La historia de la puta arrastrada por los pelos...	---	Dejato	de la historia de ayer, cofito! —La agarró de la muñeca con brusco	48.407	s\martin-1.txt	86
2	...de pasame la noche llorqueando, lo increpé en mi estilo más coloquial.	---	Dejese	de hostias, Fermín! La vida que de verdad cuenta es la que aún n	63.619	l\gimnez-1.txt	75
3	...le come el coco a Ventura, esto es el acabose ¡Madre mía, qué castigo!	---	Deja	de decir tonterías de una vez, Richi —dijo Gallardo. —¿Tonterías?	5.426	l\madrid-1.txt	23
4	...os dónde vive el señor Ventura. Eso no lo sabe nadie más que el diablo.	---	Deja	de decir tonterías, Jesucristo. Tú te has tirado de confite de la Bnga	10.775	l\madrid-1.txt	26
5	...ado, se puso junto a él. —Pues muy bien, ¿y a mí que me cuentan?	---	Deja	de hacerte el gracioso y vámonos. Esta vez no vas a salir con tanta	71.447	l\gimnez-1.txt	84
6	...te hubiera conocido Ávalos, presidirías el Valle de los Caídos, mi niña.	---	Dejate	de culturas, chotaza. Macarena se estremeció al escuchar	25.893	l\martne-1.txt	48
7	...que y los demás trataban de resituarse en el mundo y en la habitación.	---	Dejemos	de lado las aficiones del detective Carvalho y cuéntenos los tér	58.647	l\wzquez-1.txt	67
8	...tipos. —La verdad, comisario, es que no sé cómo te puedo ayudar.	---	Dejalo	de mi cuenta. Este asunto necesita un toque de sutileza de inge	35.936	l\nez_so-1.txt	85
9	...s. Yo me quedé de pie. —Bueno, Tajón, ¿que pasa con mis derechos?	---	Dejate	de derechos y empieza por decime lo que hiciste ayer. ¿Te acue	22.944	l\madrid-1.txt	55
10	...turo no existe y el pasado lo borro cada día cuando me voy a dormir.	---	Dejate	de hostias. Cuenta. Y no te ahorres detalles. Hay una biografi	5.838	l\martin-1.txt	10
11	...ste tío es un gilipollas —dijo el subinspector Garzón a la salida del bar.	---	Dejese	de hostilidades y vayamos a lo nuestro. ¿Qué hora es? —Las	78.280	l\gimnez-1.txt	92
12	...mientos los hechos desde el principio. ¿Quiere que hagamos una pizza?	---	Dejese	de pizzas, Petra, no es el momento. Nos colocamos el uno fr	77.301	l\gimnez-1.txt	91
13	...siento! —Desde que ese tipo empezó a hacer de las suyas, la Policía no deja de acosarnos, persiguiendo macarras y camellos. Están seguros de	---	Deja	de acacosarnos, persiguiendo macarras y camellos. Están seguros de	54.014	l\martin-1.txt	96
14	...horas por la calle. Me puse de pie y subí ligeramente la voz. —Si no deja de interrumpir tendré que echarla, señora. Apretó la boca y dijo alg	---	Deja	de interrumpir tendré que echarla, señora. Apretó la boca y dijo alg	3.383	l\gimnez-1.txt	4
15	...Por fin, oí un ruido, como un gemido. —Carmen, ábreme de una vez. Deja de jugar.	---	Deja	de jugar. Sonó el pestillo, el pasador de seguridad y, por fin, la p	31.131	l\martin-1.txt	57
16	...que a mi edad... El hombre lo miró con furia desbordada. —¡Estúpido! Deja de llamarme papi. ¡Dios, qué castigo! Tener un hijo como un armario y	---	Deja	de llamarme papi. ¡Dios, qué castigo! Tener un hijo como un armario y	19.269	l\nez_so-1.txt	46
17	...la intentar determinar un retrato robot. Pero lo cierto es que la psicología no deja de parecer una ciencia mezcla de intuición y fantasía, es decir, lo men	---	Deja	de parecer una ciencia mezcla de intuición y fantasía, es decir, lo men	16.088	l\gimnez-1.txt	19
18	...seductora, la gran tentadora. «Lo mirarás, pero no lo tocarás.» En cuanto deja de ser niña, su única preocupación consiste en maquillarse, peinarse	---	Deja	de ser niña, su única preocupación consiste en maquillarse, peinarse	2.812	l\martin-1.txt	5
19	...e una porción y adelanta. —Cierto. Podría haber sucedido así. Pero no deja de ser un error técnico desconocer el nombre del veneno que utilizas.	---	Deja	de ser un error técnico desconocer el nombre del veneno que utilizas.	81.210	l\wzquez-1.txt	92
20	...chóferes en la empresa. —¿Sabe como llaman a esto? Financiación. No deja de tener su gracia. ¿Verdad? Tenía la boca grande, bolsas bajo los oj	---	Deja	de tener su gracia. ¿Verdad? Tenía la boca grande, bolsas bajo los oj	6.890	l\madrid-1.txt	17
21	...a saber que una línea de ordenador es frágil como el cristal y su borrado no deja huella. «Diano de una investigación. Viaje al submundo del topless y e	---	Deja	de una investigación. Viaje al submundo del topless y e	27.614	l\martin-1.txt	51
22	...ón. —¡Ah no?, ¿por qué? —Porque pone demasiada emotividad, se deja llevar por los sentimientos de las víctimas, y debe encontrar un culpabl	---	Deja	de llevar por los sentimientos de las víctimas, y debe encontrar un culpabl	60.877	l\gimnez-1.txt	72
23	...s y mucho dinero. Tu jefe era un hampon. El local era la tapadera. —Me deja usted de piedra. —No manté. Acababa de descubrirlo hacia muy poca	---	Deja	de piedra. —No manté. Acababa de descubrirlo hacia muy poca	12.909	l\nez_so-1.txt	30
24	...Si quieres, te doy su teléfono. O si quieres, le llamo yo mismo. —No, deja. Castellet está fuera de España. Tampoco tiene mucha importancia, n	---	Deja	de España. Tampoco tiene mucha importancia, n	23.330	l\martin-1.txt	43
25	...carpeta. Se trataba de un dossier del que iba seleccionando unas fotos que dejaba a un lado de la mesa. Por último, cerró la carpeta y me miró. Al fi	---	Deja	de la mesa. Por último, cerró la carpeta y me miró. Al fi	20.412	l\nez_so-1.txt	48
26	...a escasamente secretos aunque no diferenciadamente públicos o privados, dejaba actuar a las cámaras de televisión que con sus reflectores convertía	---	Deja	de televisión que con sus reflectores convertía	25.257	l\wzquez-1.txt	29
27	...e encontraba ya en posición. Una corona de flores en embutía su cabeza y dejaba caer algunas mechadas de pelo sobre su frente, y ella estiraba la ban	---	Deja	de pelo sobre su frente, y ella estiraba la ban	25.282	l\martin-1.txt	46
28	...os no estaban con ellos, cuando estaban en el colegio o de excursión, y dejaba caer uno de sus disparates. —Hoy me he cagado encima de do	---	Deja	de sus disparates. —Hoy me he cagado encima de do	42.351	l\martin-1.txt	75
29	...n Chicote y luego nos vamos a comer algo. A mí, ni me parecía ni me dejaba de parecer. Estaba en su coche, dejándome llevar, extrañado por la	---	Deja	de parecer. Estaba en su coche, dejándome llevar, extrañado por la	6.845	l\martin-1.txt	16
30	...en el Ensanche, con Avelino, como su época de plenitud. Al fin solos, ella dejaba de ser criada para ser esposa, madre y enfermera, al fin hablando f	---	Deja	de ser criada para ser esposa, madre y enfermera, al fin hablando f	41.810	l\martin-1.txt	74
31	...o demoes que les hicieran olvidar que estaban pagando por entrar en ellos no dejaba de ser una tontería, una variante más. ¡Pero yo me había quedado	---	Deja	de ser una tontería, una variante más. ¡Pero yo me había quedado	5.645	l\nez_so-1.txt	13
32	...r sustituciones de imagen. Pero todo estaba contado de modo que se dejaba entrever más de lo que se contaba. Un lector inocente, si es que se	---	Deja	de lo que se contaba. Un lector inocente, si es que se	45.816	l\wzquez-1.txt	84
33	...presión que había creado poniendo un sobre en la mano de Carvalho que le dejaba libre el vaso de whisky. —Para sus gastos por Madrid. Dinero d	---	Deja	de whisky. —Para sus gastos por Madrid. Dinero d	15.978	l\wzquez-1.txt	58
34	...d por saber cómo se comportaba cuando su padre y el entorno de su padre dejaban de ser el referente de su vida. Lázaro Conesal apenas probó el jere	---	Deja	de ser el referente de su vida. Lázaro Conesal apenas probó el jere	45.552	l\wzquez-1.txt	12

Imagen 22 – Ventana de resultados de la búsqueda de concordancias con WordSmith Tools

ParaConc es muy similar a WordSmith. Esta herramienta también ofrece la posibilidad de obtener la frecuencia de aparición de unidades específicas, y de ver palabras clave en contexto, función fundamental que permite localizar con facilidad los elementos estudiados y analizar de manera sistemática sus patrones de utilización. La diferencia es que está diseñada para trabajar con textos paralelos.

The screenshot shows the ParaConc interface with two text files loaded: 'L\isfor-LAWLESS SUE GRAFTON Capítulo 1' and 'L\isfor-LAWLESS Sue Grafton Chapter 1'. A window titled 'Alignment Spanish (Modern Sort) - English (United States) (Grafton_ES.txt - Grafton_EN.txt): Segments' displays the aligned text. The alignment is shown with vertical lines indicating corresponding lines in both languages.

Spanish (Modern Sort)	English (United States)
L\isfor-LAWLESS	L\isfor-LAWLESS
SUE GRAFTON	Sue Grafton
Capítulo 1	Chapter 1
No es por quejarme, pero en lo sucesivo me lo pensaré dos veces antes de hacer un favor a los amigos de los amigos.	I don't mean to bitch, but in the future I intend to hesitate before I do a favor for the friend of a friend.
Jamás lo he lamentado tanto.	Never have I taken on such a load of grief.
Al principio, todo parecía de lo más inocente.	At the outset, it all seemed so innocent.
Juro que no había manera de adivinar lo que iba a ocurrir.	I swear there's no way I could have guessed what was coming down.
Estuve a un paso de la muerte, lo que es quizá peor (para los que como yo padecen dentifonía), a un pelo de perder los dos incisivos superiores.	I came this close to death and, perhaps worse (for my fellow dental phobics), within a hairbreadth of having my two front teeth knocked out.
Y todavía tengo en la cabeza un chichón del tamaño de un puño.	Currently I'm sporting a knot on my head that's the size of my fist.
Y todo por un trabajo por el que ni siquiera me pagaron.	And all this for a job for which I didn't even get paid!
El caso me llamó la atención por culpa de mi casero, Henry Pitts, del que todos saben que estoy medio enamorada desde hace años.	The matter came to my attention through my landlord, Henry Pitts, whom everybody knows I've been half in love with for years.
Que tenga ochenta y cinco años (sólo cincuenta más que yo) no parece haber modificado nunca el impacto básico de su atractivo.	The fact that he's eighty-five (a mere fifty years my senior) has never seemed to alter the basic impact of his appeal.
Es un encanto y casi nunca me pide nada, de manera que era imposible negarse.	He's a sweetheart and he seldom asks me for anything, so how could I refuse?
Sobre todo porque su petición parecía muy inofensiva en principio, sin nada que permitiera entrever los problemas que desencadenaría.	Especially when his request seemed so harmless on the face of it, without the faintest suggestion of the troubles to come.
Era el jueves veintiuno de noviembre, una semana antes del Día de Acción de Gracias, y se estaban ultimando los preparativos de la boda.	It was Thursday, November twenty-first, the week before Thanksgiving, and wedding festivities were just getting under way.
William, el hermano mayor de Henry, iba a casarse con mi amiga Rosie, que dirige una anticuada casa de comidas que hay en mi barrio.	Henry's older brother William was to marry my friend Rosie, who runs the tacky little tavern in my neighborhood.
El local de Rosie cierra tradicionalmente el Día de Acción de Gracias y la propietaria estaba radiante por poder casarse con William sin necesidad de perder dinero.	Rosie's restaurant was traditionally closed on Thanksgiving Day, and she was feeling smug that she and William could get hitched without her losing any business.
Se las había ingeniado para prescindir de la iglesia proyectando la ceremonia y el	With the ceremony and reception being held at the restaurant, she'd managed to

Imagen 23 – Textos alineados en ParaConc

En este programa, el primer paso es cargar dos archivos, el original y la traducción, a la vez, seleccionando la lengua de cada uno y, si está alineado o no y de qué manera (mediante línea nueva o símbolos delimitadores, por ejemplo). El programa

incorpora su propia herramienta de alineación, pero a nosotros nos ha resultado más sencillo alinear los textos previamente con Déjà Vu (y segmentarlos) que hacerlo a través de este alineador. En la imagen 23 podemos ver los textos alineados tal como se muestran en ParaConc. Mediante el diálogo de búsqueda ya podemos empezar a realizar consultas. Por ejemplo, en la captura de la imagen 24 hemos elegido buscar todos los casos de la perífrasis «dejar de + infinitivo».

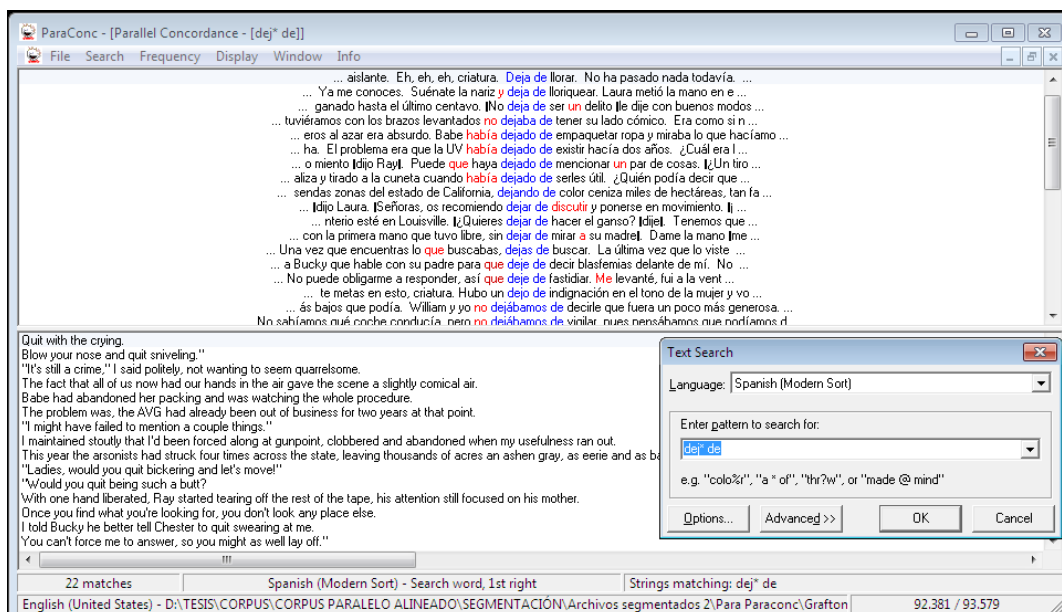


Imagen 24 – Concordancias paralelas en ParaConc

Como puede observarse, la ventana de arriba muestra los ejemplos de uso en contexto en el subcorpus de traducciones, de la misma manera en la que lo hacía WordSmith, con la expresión de búsqueda en el centro resaltada con otro color y rodeada de su contexto (que también podemos ampliar a voluntad). La diferencia es que la primera línea de arriba, que incluye el fragmento «Deja de llorar», se corresponde con la primera línea de la ventana de abajo «Quit with the crying», y así sucesivamente. De esta manera, podemos analizar las correspondencias de esta perífrasis en las concordancias paralelas del corpus de textos originales en inglés. Las palabras marcadas en otro color cerca de la expresión resaltada muestran posibles colocaciones⁹².

Otra función particularmente interesante de este programa es la denominada *hot words*, que le permite sugerir posibles traducciones para nuestra expresión de búsqueda

⁹² Según Bowker (2002: 64), las colocaciones son «characteristic co-occurrence patterns of words (...) words that “go together” or “are found in each other’s company” (...) words that appear together with a greater than random probability. Because language is not random, certain words tend to cluster together, and some of these clusters form collocations».

basándose en el análisis de las concordancias paralelas. ParaConc nos ofrece una lista de candidatos y nosotros podemos seleccionar cualquiera de ellos para que aparezca resaltado en la ventana inferior y se nos facilite así la búsqueda de regularidades en la traducción y, en consecuencia, posibles equivalentes.

3.4.2. Corpus Query Processor

Los dos programas descritos en el apartado anterior son perfectamente válidos para analizar nuestro corpus. No obstante, como ya hemos comentado, nos pareció que IMS Open Corpus Workbench ofrecía nuevas posibilidades y mayor flexibilidad en el lenguaje de búsqueda (Zanettin 2012: 184), a pesar de las desventajas ya señaladas con respecto a la dificultad que entraña comenzar a utilizarlo y su falta de interfaz gráfica.

CWB es un conjunto de herramientas de gestión y explotación de corpus bastante más poderoso, puesto que, a diferencia de WordSmith, es más que un extractor de concordancias: entre otras funciones, permite la codificación, la indexación y la manipulación de grandes corpus de texto etiquetado. Una de sus herramientas es el IMS Corpus Query Processor o CQP (Evert et ál. 2010), que se utiliza para realizar búsquedas en los textos mediante comandos y expresiones regulares.

```

administrador@CWBcovalt: ~/Barbara/Parallel_Corpus/Demostrativos
Usage of /: 6.1% of 18.55GB  Swap usage: 0%  Users logged in: 0

Graph this data and manage this system at https://landscape.canonical.com/

172 packages can be updated.
121 updates are security updates.

New release 'precise' available.
Run 'do-release-upgrade' to upgrade to it.

Last login: Thu Apr 10 11:44:44 2014 from 58.253.14.62.static.jazztel.es
administrador@CWBcovalt:~$ cd Barbara/Parallel_Corpus/Demostrativos/
administrador@CWBcovalt:~/Barbara/Parallel_Corpus/Demostrativos$ cqp -e -C
Welcome to CQP -- the Colourful Query Processor.
[30: B  31: NBU$  32: NBU$  33: NBU$  34: NBU$  35: NBU$  36: NBU$  37: NBU$
]
[40: NBU$  41: NBU$  42: NBU$  43: NBU$  44: NBU$  45: NBU$  46: NBU$  47: B
]
[no corpus]> PAR_TRAD_BARBARA
PAR_TRAD_BARBARA> set PrintStructures text_author
PAR_TRAD_BARBARA> set PrintOptions num
PAR_TRAD_BARBARA> set PrintOptions hdr
PAR_TRAD_BARBARA> set Context 2s
PAR_TRAD_BARBARA>

```

Imagen 25 – Acceso a CQP mediante Putty

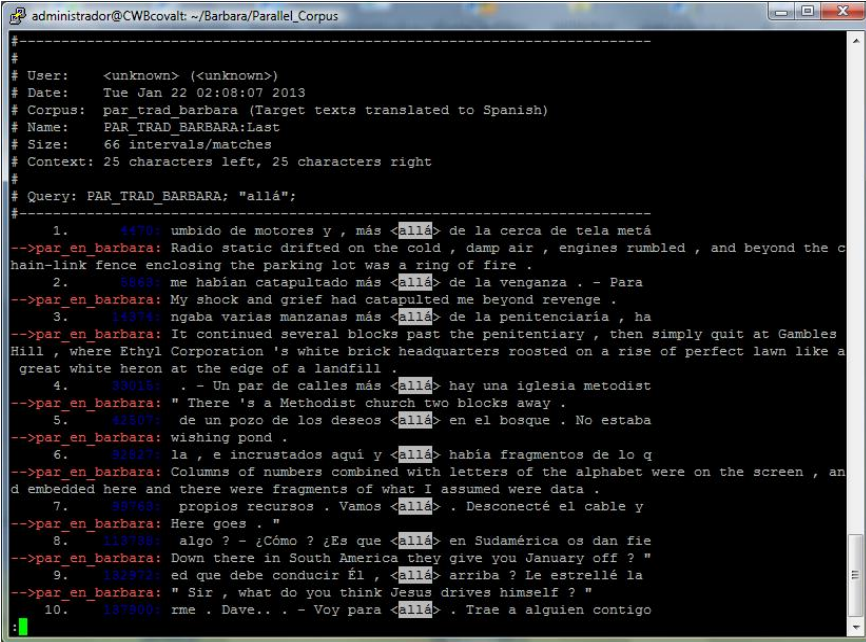
Accedemos al servidor de CWB mediante la utilización, una vez más, del programa Putty y los datos proporcionados por el administrador del servidor (nombre,

contraseña y puerto). Como vemos en la imagen 25, una vez dentro, entramos en el directorio en el que tenemos nuestras carpetas con los textos del corpus y activamos CQP tecleando «cqp -e -C» (esta última «C» es opcional, pero nos permite acceder a la versión coloreada del procesador de búsquedas). A continuación, seleccionamos el corpus en el que queremos realizar la búsqueda (en este caso, «PAR_TRAD», el de textos traducidos), y activamos todas las opciones que deseemos para visualizar los resultados. Por ejemplo, «set PrintStructures text_author» nos permitirá saber de qué texto proviene cada una de las concordancias que arrojen las búsquedas, pues al principio aparecerá el nombre del autor o la autora. Con «set PrintOptions hdr» lograremos que las concordancias vayan precedidas de un encabezado en el que se resumirán datos como la fecha de la búsqueda, el nombre del corpus o el número de resultados que coinciden con nuestra búsqueda, lo que puede ser útil si queremos guardar la búsqueda y consultarla más adelante. Mediante «set PrintOptions num» se numerarán las concordancias y «set Context», seguido del número que consideremos conveniente, establecerá el tamaño del contexto que rodea el término de búsqueda. Por ejemplo, si lo establecemos en «100», serán cien caracteres en total, si tecleamos «5 s», serán cinco oraciones. También podemos seleccionar contexto a la derecha o a la izquierda indicándolo mediante un comando como este: «Set LeftContext 50». Es importante contar con el contexto suficiente para poder analizar las soluciones de traducción sin constreñirnos al límite de la palabra o la frase (Granger y Tribble 1998: 209)⁹³. Al analizar las concordancias en los archivos exportados, nos dimos cuenta de que es mucho más útil establecer un contexto por caracteres y no por frases, ya que de la primera forma se preserva el formato KWIC (las palabras claves alineadas en medio del contexto, fácilmente identificable), mientras que de la segunda manera, las palabras podían aparecer a cualquier altura de la página, lo que dificultaba el análisis visual de las concordancias.

Ahora ya podremos empezar la búsqueda. La estrategia más sencilla es simplemente introducir una palabra entre comillas (los comandos deben terminar siempre con punto y coma pero CQP los añade automáticamente, siempre que hayamos

⁹³ Aunque en este punto hay opiniones divergentes: Ebeling comenta lo siguiente en un trabajo de 1998 en el que se vale del ENPC para estudiar construcciones presentativas en inglés y noruego: «I should stress that I have made no attempt to look at a wider context to see if the information added or lost can be found outside the sentence translated, or can be recovered by social or cultural knowledge or logical inference. The latter, to my mind, lies outside the scope of contrastive linguistics, at least at the level of syntax» (1998: 6).

introducido la *-e* inicial al abrirlo). Por ejemplo, teclear “allá” equivale a introducir la siguiente expresión: [word=“allá”].



```

administrador@CWBcovalt: ~/Barbara/Parallel_Corpus
#-----#
#
# User:      <unknown> (<unknown>)
# Date:     Tue Jan 22 02:08:07 2013
# Corpus:   par_trad_barbara (Target texts translated to Spanish)
# Name:     PAR_TRAD_BARBARA:Last
# Size:     66 intervals/matches
# Context:  25 characters left, 25 characters right
#
# Query:    PAR_TRAD_BARBARA: "allá";
#-----#
1.      4470:  umbido de motores y , más <allá> de la cerca de tela metá
-->par_en_barbara: Radio static drifted on the cold , damp air , engines rumbled , and beyond the c
hain-link fence enclosing the parking lot was a ring of fire .
2.      4869:  me habían catapultado más <allá> de la venganza . - Para
-->par_en_barbara: My shock and grief had catapulted me beyond revenge .
3.      14374: ngaba varias manzanas más <allá> de la penitenciaría , ha
-->par_en_barbara: It continued several blocks past the penitentiary , then simply quit at Gambles
Hill , where Ethyl Corporation 's white brick headquarters roosted on a rise of perfect lawn like a
great white heron at the edge of a landfill .
4.      30913:  . - Un par de calles más <allá> hay una iglesia metodist
-->par_en_barbara: " There 's a Methodist church two blocks away .
5.      32307:  de un pozo de los deseos <allá> en el bosque . No estaba
-->par_en_barbara: wishing pond .
6.      31817:  la , e incrustados aquí y <allá> había fragmentos de lo q
-->par_en_barbara: Columns of numbers combined with letters of the alphabet were on the screen , an
d embedded here and there were fragments of what I assumed were data .
7.      38763:  propios recursos . Vamos <allá> . Desconecté el cable y
-->par_en_barbara: Here goes . "
8.      119798: algo ? - ¿Cómo ? ¿Es que <allá> en Sudamérica os dan fie
-->par_en_barbara: Down there in South America they give you January off ? "
9.      139973: ed que debe conducir él , <allá> arriba ? Le estrellé la
-->par_en_barbara: " Sir , what do you think Jesus drives himself ? "
10.     137309: rme . Dave . . - Voy para <allá> . Trae a alguien contigo

```

Imagen 26 - Listado de concordancias de la palabra allá en CQP

A continuación incluimos algunos ejemplos del tipo de búsquedas que hemos realizado durante el análisis. Para más detalles, puede consultarse Evert et ál. (2010):

- [word=“aterriz.*”]: buscará todas las palabras que empiecen por esas letras (*aterrizaje*, *aterrizaron*, etc.). Si en lugar de un asterisco utilizáramos un signo de interrogación tras el punto, solo nos mostraría las palabras que contuvieran una única letra más después de la zeta, como *aterrizó*, y si fuera una signo más (+), una o más letras.

- [word=“aterriz.*” %c]: si no introducimos el signo de porcentaje y la letra *c*, solo obtendremos aquellas concordancias que empiecen por minúscula. De esta manera, también nos saldrán nombres propios y palabras que se encuentren al principio de una oración. Añadiendo el signo de porcentaje seguido de la letra de (%d), CQP no distinguirá entre palabras acentuadas y no acentuadas.

- “dejar” []{0,2} “de”: Los corchetes indican que, intercalado entre la primera y la última palabra de la expresión regular, puede ir un elemento. Podemos indicar entre llaves cuántas palabras como mínimo y como máximo queremos incluir.

- [word=“mand(ó|o)n(a|e)?(s)?”]: Mediante barras y paréntesis podemos incluir varias opciones de búsqueda en una misma expresión: aquí, por ejemplo, le

estamos pidiendo al programa que busque casos tanto de *mandón*, como de *mandones*, *mandona* o *mandonas*. Las barras dividen las distintas alternativas y el signo de interrogación indica que lo que está entre paréntesis es opcional.

- **[lem="comer"]**: También podemos realizar búsquedas por lemas o etiquetas. Esta expresión hallará todas las formas conjugadas del verbo, y **«[pos="V.*"]»**, todos los verbos etiquetados como tales (el asterisco se incluye porque existen varias etiquetas diferentes para clasificar los verbos pero en este caso no nos interesa buscar ninguna en concreto).

- **[word="suelo" & pos!="N.*"]**: El signo de exclamación nos permitirá buscar casos en los que *no* se cumplen determinadas condiciones. Por ejemplo, la expresión anterior hallará todos los casos de *suelo* que no sean sustantivos.

Las concordancias pueden reordenarse por orden alfabético mediante el comando **«sort by word»**, sin olvidarnos de añadir **«%c»** al final para que el programa tenga en cuenta las palabras con mayúsculas y minúsculas. El comando **«sort by word %c matchend[1]»** ordenará las concordancias por la primera palabra a la derecha y **«sort by word %c match[-1]»**, por la primera a la izquierda; **«sort Last match by "text_author"»** lo hará por autor, mientras que si escribimos **«group Last match text_author»**, el programa nos proporcionará las estadísticas de distribución por autor (o texto, si utilizamos "text_id").

```
administrador@CWbcovalt: ~/Barbara/Paralel_Corpus
#-----
#
# User: <unknown> (<unknown>)
# Date: Tue Jan 22 02:09:04 2013
# Corpus: par_trad_barbara (Target texts translated to Spanish)
# Name: PAR_TRAD_BARBARA:Last
# Size: 66 Intervals/matches
# Context: 25 characters left, 25 characters right
#
# Query: PAR_TRAD_BARBARA: "allá";
#-----
1. 4470: ores/NC y/CC ,/CM más/ADV <allá/ADV> de/PREP la/ART cerca/ADV
-->par_en_barbara: Radio/NP static/NN drifted/VVD on/IN the/DT cold/JJ ,/, damp/JJ air/NN ,/, engin
es/NNS rumbled/VVD ,/, and/CC beyond/IN the/DT chain-link/NN fence/NN enclosing/VVG the/DT parking/
NN lot/NN was/VBD a/DT ring/NN of/IN fire/NN ./SENT
2. 5883: catapultado/VLadj más/ADV <allá/ADV> de/PREP la/ART venganza/
-->par_en_barbara: My/PPS shock/NN and/CC grief/NN had/VHD catapulted/VVN me/PP beyond/IN revenge/N
N ./SENT
3. 14374: as/QU manzanas/NC más/ADV <allá/ADV> de/PREP la/ART penitenci
-->par_en_barbara: It/PP continued/VVD several/JJ blocks/NNS past/IN the/DT penitentiary/NN ,/, the
n/RB simply/RB quit/VVN at/IN Gambles/NP Hill/NP ,/, where/WRB Ethyl/NP Corporation/NP 's/POS white
/JJ brick/NN headquarters/NNS roosted/VVN on/IN a/DT rise/NN of/IN perfect/JJ lawn/NN like/IN a/DT
great/JJ white/JJ heron/NN at/IN the/DT edge/NN of/IN a/DT landfill/NN ./SENT
4. 38015: e/PE calles/VLfin más/ADV <allá/ADV> hay/VHfin una/ART iglesi
-->par_en_barbara: "/` There/EX 's/VBZ a/DT Methodist/JJ church/NN two/CD blocks/NNS away/RB ./SEN
T
5. 41307: de/PREP los/ART deseos/NC <allá/ADV> en/PREP el/ART bosque/NC
-->par_en_barbara: wishing/VVG pond/NN ./SENT
6. 42427: tados/VLadj aquí/ADV y/CC <allá/ADV> había/VHfin fragmentos/N
-->par_en_barbara: Columns/NNS of/IN numbers/NNS combined/VVN with/IN letters/NNS of/IN the/DT alph
abet/NN were/VBD on/IN the/DT screen/NN ,/, and/CC embedded/VVD here/RB and/CC there/EX were/VBD fr
agments/NNS of/IN what/WP I/PP assumed/VVD were/VBD data/NNS ./SENT
7. 38769: ursos/NC ./FS Vamos/VLfin <allá/ADV> ./FS Desconecté/VLfin el
-->par_en_barbara: Here/RB goes/VVZ ./SENT "/`
```

Imagen 27 - Listado de concordancias de la palabra *allá* en CQP con las etiquetas visibles

Si nuestro corpus está etiquetado, para ver dicha información solo hay que introducir el comando `«show +pos»`. Recordemos que nosotros utilizamos ambas versiones del corpus, pues al principio queríamos asegurarnos de que el etiquetado se había realizado correctamente y de que se recogían todas las apariciones de los elementos en las búsquedas sobre el etiquetado y sobre el texto sin etiquetar, para poder utilizar uno u otro formato según nuestras necesidades. En la imagen 27 vemos el mismo texto de la imagen 26 con las etiquetas ahora visibles.

Tanto en la captura 26 como en la 27 vemos que, debajo de las concordancias de TRAD, aparece el fragmento de EN alineado. Para visualizar estas concordancias paralelas deberemos indicárselo a CQP cada vez que cambiemos de corpus mediante el comando `«show +nombre_del_corpus»`. Eso significa que, si estamos realizando las búsquedas en el corpus de traducciones, tendremos que escribir el nombre del corpus de originales. El contexto de la lengua opuesta siempre consta de una única frase, pero, dado que las búsquedas paralelas pueden hacerse en una dirección o la otra, si queremos ampliarlo, no tenemos más que realizar la búsqueda de una o dos palabras clave del mismo en el subcorpus opuesto.

Para guardar los resultados en un fichero de texto, utilizaremos el comando `«cat Last > “archivo.txt”»`, escribiendo, en lugar de *archivo*, el nombre del fichero que queremos crear. Una desventaja bastante importante con respecto a WordSmith es que no es posible eliminar casos irrelevantes antes de guardar los resultados (como es posible hacer en aquel programa con el comando *zap*). Esto nos obliga a revisar las concordancias con posterioridad y utilizar un procesador de textos para suprimirlas.

Corpus-analysis tools are not intelligent – they work using pattern matching techniques. Therefore, they will retrieve exactly what users ask them to retrieve, even if this is not necessarily what the users want to retrieve. Common problems include “silence” and “noise”. (Bowker 2002: 73)

El ruido es algo inevitable y por eso hemos realizado un cuidadoso cribado manual de concordancias en todos los casos. Por otra parte, una decisión que hemos tomado es la de no utilizar el parámetro `«%d»` en las búsquedas, que evitaría incluir las formas acentuadas en las expresiones regulares, con el objetivo de evitar el ruido que generarían los acentos en otras partes de la palabra: por ejemplo, `«*ando* %d»`

hallaría ambas terminaciones de gerundio, la acentuada y la no acentuada, pero también recuperaría casos como *mandó*.

Con respecto al silencio, nos ha pasado muchas veces no haber recuperado concordancias que nos interesaban por no haber introducido una expresión de búsqueda lo suficientemente amplia o simplemente por errores provocados por falta de atención, dado que una simple omisión (de un punto o una comilla) puede invalidar la expresión. En el caso de expresiones en las que nos faltaban elementos, muchas veces nos hemos dado cuenta a lo largo del análisis, cuando volvíamos al corpus para buscar ejemplos o ampliar contextos.

Por ejemplo, no fue hasta después de haber realizado prácticamente todas las búsquedas de perífrasis que nos dimos cuenta de que no habíamos obtenido ningún ejemplo precedido de un signo de apertura de interrogación o exclamación. Vimos, entonces, que, por algún motivo que quizás haya que analizar (dado que no sabemos si puede deberse a un problema de nuestro corpus o no), CQP interpreta que dichos signos aparecen pegados a las palabras y hay que incluirlos en las expresiones de búsqueda, lo que nos obligó a repetirlos. Posteriormente comprobamos que no sucede así con los de cierre, ni con otros signos como las comillas o las rayas de diálogo, afortunadamente. Otro caso de silencio han sido las búsquedas con etiquetas, como explicaremos en los capítulos de análisis de perífrasis y demostrativos. Como vemos, el análisis del corpus ha sido en muchas ocasiones un proceso circular repleto de revisiones.

3.5. Tratamiento estadístico de los resultados

Como hemos adelantado, los cómputos de concordancias obtenidos a partir de los distintos corpus pueden ser indicativos de tendencias pero, a menos que los normalicemos, no serán comparables, dado que el número de palabras de cada uno es distinto. Para remediar este problema, las frecuencias de aparición obtenidas en el análisis deben ser extrapoladas a un común denominador. Esto se consigue multiplicando el número de apariciones encontradas en un subcorpus determinado por un mismo número, que en lingüística de corpus suele ser mil o un millón —pero en realidad puede ser el que queramos—, y dividiendo el resultado por el número total de

palabras del subcorpus del que se ha extraído dicha frecuencia (Müller y Waibel 2009). Este procedimiento, denominado normalización, se realiza sobre las frecuencias de ambos subcorpus, por supuesto. McEnery, Xiao y Tono (2006: 53) nos advierten de que el denominador común debe ser una cifra comparable al tamaño de nuestro corpus. De lo contrario, los resultados podrían verse tergiversados, tanto magnificados como minimizados. Por este motivo hemos elegido normalizar las frecuencias obtenidas en el análisis por cien mil, una cifra más coherente con la cantidad de palabras de nuestro corpus que mil o un millón.

Sin embargo, la normalización de las frecuencias no es suficiente para validar nuestras hipótesis; es necesario realizar un test de significación estadística para comprobar que estos resultados no son producto del azar. Para saber si las diferencias en el número de apariciones observadas entre los distintos subcorpus pueden ser tenidas en cuenta desde el punto de vista estadístico, aplicaremos el test de log-likelihood (LL; también conocido como G^2), diseñado para la comparación de frecuencias entre corpus de distinto tamaño. Se trata de uno de los tests de significación más comunes, muy similar al de chi-cuadrado (χ^2)⁹⁴, pero más fiable cuando las frecuencias esperadas son pequeñas y las diferencias en el tamaño de los corpus, grandes (Rayson, Berridge y Francis 2004: 4). Además, no asume que los datos presentan una distribución normal, como otros tests (Oakes 1998: 189). Entre los estudios que han utilizado este test de significación estadística podemos nombrar a Botley y McEnery (2001), McEnery y Xiao (1999), Axelsson (2011), Gabrielatos y Sarmiento (2006) o Milton (1998).

Para los cálculos nos valdremos de una calculadora en línea diseñada por Paul Rayson, de la Universidad de Lancaster (Rayson 1993-2010). A continuación, podemos ver una imagen de la página, en donde también se explica cómo se realizan los cálculos y se ofrece más documentación sobre su utilización⁹⁵:

⁹⁴ El test de chi-cuadrado se describirá en el capítulo de análisis de los demostrativos, que es la única ocasión en la que lo utilizaremos.

⁹⁵ Estas capturas son de septiembre de 2015. Cuando utilizamos la calculadora, la página se llamaba simplemente *Log-likelihood calculator*. Al parecer el autor ha añadido también recientemente medidas de *effect size* (tamaño del efecto) que no hemos podido tener en cuenta para este trabajo (Rayson 1993-2014).

Log-likelihood and effect size calculator

To use this wizard, type in frequencies for one word and the corpus sizes and press the calculate button.

	Corpus 1	Corpus 2
Frequency of word	150	192
Corpus size	324431	520410

Notes:

1. Please enter plain numbers without commas (or other non-numeric characters) as they will confuse the calculator!
2. The LL wizard shows a plus or minus symbol before the log-likelihood value to indicate overuse or underuse respectively in corpus 1 relative to corpus 2.
3. The log-likelihood value itself is always a positive number. However, my script compares relative frequencies between the two corpora in order to insert an indicator for '+' overuse and '-' underuse of corpus 1 relative to corpus 2.

Imagen 28 – Calculadora de log-likelihood en línea de Paul Rayson

Hemos introducido los números de palabras de ES y TRAD (fase I) en la fila de abajo, y en la de arriba, un par de frecuencias. El cálculo tiene en cuenta los números de palabras totales de los subcorpus y, por tanto, no es necesario normalizar las frecuencias antes de introducirlas en las casillas. Los grados de libertad (d. f.) serán siempre los mismos (en el caso de nuestros análisis, 1) porque la tabla de contingencia es de 2x2, por lo que no los indicaremos a lo largo del trabajo cuando utilicemos este test.

Al resultado del test se le asigna un valor positivo o negativo según si la perífrasis aparece mayor o menor número de veces en el primer subcorpus de la comparación. El valor de p nos indica si aquellas perífrasis que poseen una frecuencia de uso menor en traducciones pueden ser tenidas en cuenta para validar la hipótesis de infrarrepresentación: si la diferencia es significativa, se indicará con qué nivel de confianza (es decir, probabilidad de fallar en nuestra estimación). Si no lo es, en nuestro análisis lo indicaremos mediante la abreviatura «n. s.». Cuanto más alto es el valor del test y más bajo el valor de p , más significativa es la diferencia; es decir, más baja la probabilidad de observar una diferencia como la que hemos observado incluso si las dos poblaciones fueran iguales, es decir, si la hipótesis nula fuera verdadera⁹⁶.

En realidad, nosotros establecemos un nivel de significación estadística, en cierta manera arbitrario, a partir del cual podemos rechazar la hipótesis nula de que no existe diferencia entre las frecuencias de ambos subcorpus⁹⁷, y lo que hacemos

⁹⁶ Para una recopilación de los malentendidos más habituales en torno a la interpretación del valor p , véase el siguiente recurso:

The most common misinterpretation of a P value (1995-2015). [en línea] Disponible en *GraphPad Software* http://www.graphpad.com/guides/prism/6/statistics/index.htm?common_misinterpretation_of_a_p_value.htm [última fecha de consulta: 27/09/2015].

⁹⁷ Baroni y Evert (2006: 14) lo explican así: «Since there is a considerable difference between f_1 and f_2 , we are tempted to reject H_0 . However, before we can do so, we must find out to what extent the difference can be explained by random variation, i.e., we have to calculate how likely it is that the two samples come from populations with the

mediante el test es aceptar o rechazar dicha hipótesis, aunque siempre hay que tener en cuenta que «even if the null hypothesis is not rejected, we cannot conclude that it is true. The cut-off value corresponding to the chosen degree of confidence may not be exceeded, but this only indicates there is not enough evidence to reject the null hypothesis» (Rayson, Berridge y Francis 2004: 3). Rechazar la hipótesis nula equivale a aceptar la hipótesis de que los resultados provienen de poblaciones diferentes, y «By convention, the general practice is that a hypothesis can be accepted only when the level of significance is less than 0.05 (i.e. $p < 0.05$). In other words, one must be more than 95 per cent confident that the observed differences have not arisen by chance» (McEnery, Xiao y Tono 2006: 55).

Log-likelihood calculator results

Key:
 O1 is observed frequency in Corpus 1
 O2 is observed frequency in Corpus 2
 %1 and %2 values show relative frequencies in the texts.
 + indicates overuse in O1 relative to O2,
 - indicates underuse in O1 relative to O2

Item	O1	%1	O2	%2	LL	%DIFF	Bayes
Word	150	0.05	192	0.04 +	4.24	25.32	-9.41

If you have technical problems please get in touch with [Paul Rayson](mailto:p.rayson@lancaster.ac.uk) (email: p.rayson@lancaster.ac.uk)

Imagen 29 – Pantalla de resultados de la calculadora de log-likelihood

En la captura de los resultados vemos que el valor del test (LL) es +4,24. El signo más (+) nos indica un mayor uso en el corpus de la primera columna (ES) y, si consultamos los valores críticos del test, comprobamos que esta cifra supera el primer umbral de significación y, por tanto, se rechaza la hipótesis nula con un 95 % de confianza y podemos afirmar que existe una infrarrepresentación en los textos traducidos del elemento cuya frecuencia estamos comparando. Los umbrales, que se encuentran disponibles en la misma página, son los siguientes:

- 95th percentile; 5% level; $p < 0.05$; critical value = 3.84
- 99th percentile; 1% level; $p < 0.01$; critical value = 6.63
- 99.9th percentile; 0.1% level; $p < 0.001$; critical value = 10.83
- 99.99th percentile; 0.01% level; $p < 0.0001$; critical value = 15.13

same proportion of passives, as stated by the null hypothesis (statistics textbooks will often phrase the null hypothesis directly as: the samples are from the same population)».

El problema con los tests de chi-cuadrado y log-likelihood es que, si alguna de las celdas de la tabla de valores esperados que se crean para calcularlos contiene un valor menor a cinco, los resultados del test no pueden considerarse fiables. Una solución a este problema es utilizar el Fisher's Exact Test (Kilgariff 2001: 105; Baroni y Evert 2006: 16; McEnery, Xiao y Tono 2006: 56). En aquellos pares de frecuencias en los que se dé esta circunstancia, realizaremos este test además del de log-likelihood para comprobar que los resultados de este último son correctos y apuntaremos el resultado con una nota a pie. Para realizar los cálculos, que al parecer requieren bastante capacidad de procesamiento, hemos tenido que probar varias calculadoras, puesto que con las frecuencias de nuestro corpus, algunas no eran capaces de proporcionar un resultado. Finalmente, hemos seleccionado la de la página *Social Science Statistics* (Stangroom 2015).

Saldanha (2009: 5) explica que existen diferentes posturas con respecto a la utilidad y la importancia de los tests de significación en lingüística de corpus, pues, aunque es importante demostrar que los fenómenos observados no se deben al azar, la naturaleza de los datos lingüísticos es diferente a la materia de estudio de otras ciencias:

For example, the most powerful tests used in the social sciences (parametric tests) assume that the data are normally distributed, which is often not true of linguistic data (Oakes, 1998: 11; McEnery and Wilson, 1996: 70). Non-parametric tests, such as chi-square, on the other hand, are unreliable with small frequencies. Besides, as Danielsson (2003) points out, statistical tests in many cases do not show anything that cannot be revealed by simply comparing raw frequencies. Danielsson argues that, if something is recurrent in a text, it is there for a reason, but it cannot be expected that the reason may be discovered in a simple calculation, because “the distribution of words in texts is far more complex than a mathematical formula can perceive” (ibid: 114).

Olohan (2004: 86) cita a varios otros autores que no están de acuerdo con que la utilización de técnicas de análisis estadístico sea un paso esencial; entre ellos, a Halliday, Leech y Short o Sinclair, quien afirma lo siguiente: «often it is enough to observe a broad tendency in the data for findings to be interesting and significant». Por otra parte, hay que tener en cuenta la opinión de lingüistas como Kilgariff, que dice

claramente: «The problem for empirical linguistics is that language is not random, so the null hypothesis is never true» (2005: 264). En su opinión, la hipótesis nula será rechazada siempre que tengamos suficiente cantidad de datos, por lo que estos tests en realidad no nos dicen gran cosa (aunque, a la vez, se requiere mucha menos cantidad de datos para rechazarla si dos fenómenos están fuertemente relacionados, lo que hace que la hipótesis nula tenga su utilidad en determinadas circunstancias, como podría ser el caso de nuestro corpus).

Hypothesis testing has been used to reach conclusions, where the difficulty in reaching a conclusion is caused by sparsity of data. But language data, in this age of information glut, is available in vast quantities. A better strategy will generally be to use more data. Then the difference between the motivated and the arbitrary will be evident without the use of compromised hypothesis testing. As Lord Rutherford put it: “If your experiment needs statistics, you ought to have done a better experiment.” (Kilgariff 2005: 273)

Sin embargo, los trabajos más recientes de esta área de conocimiento y nuestra propia experiencia con evaluadores de artículos nos dicen que la tendencia predominante hoy en día es la de apoyar las conclusiones de los estudios con datos estadísticos; especialmente tests de significación. De hecho, las objeciones de Kilgariff, han sido contestadas, por ejemplo, por Gries (2005), que aboga por complementar estos tests con otras técnicas como el cálculo del tamaño del efecto (*effect size*)⁹⁸. No obstante, los números no hablan por sí mismos ni cuentan toda la historia, por lo que, como ya hemos explicado en la introducción de este capítulo de metodología, en nuestro análisis no dejaremos de lado la evaluación cualitativa de los datos; los tests estadísticos servirán únicamente de apoyo a las conclusiones o para identificar posibles puntos de interés en los datos.

En los siguientes tres capítulos, que conforman la parte de análisis de la tesis, aplicaremos las pautas metodológicas descritas al estudio de los tres elementos que hemos seleccionado como posibles *unique items* españoles: un grupo de perífrasis

⁹⁸ Para una definición del concepto de *effect size* y de su importancia, véase Coe (2002). Hardie (2014b), por ejemplo, proporciona una explicación sencilla de la utilidad de una de las posibles medidas del tamaño del efecto, el Log Ratio (incluida también entre las medidas proporcionadas por la Log-likelihood Calculator de Rayson). No obstante, como hemos explicado, estos valores no estaban aún disponibles en el momento de redactar los capítulos de análisis de este trabajo, por lo que no hemos podido incluirlos.

verbales de infinitivo y gerundio, el sistema de pronombres, determinantes y adverbios locativos demostrativos y un conjunto de sufijos apreciativos aumentativos, diminutivos y peyorativos.

4. Perífrasis verbales

4.1. Perífrasis verbales: definición y caracterización

Las perífrasis verbales son elementos que diversos autores consideran una peculiaridad del sistema lingüístico español debido al elevado número de construcciones de este tipo que existen en nuestra lengua en comparación con otras como el inglés y a la riqueza de matices que aportan a la conjugación verbal. Yllera (1999: 3394) comenta que «las perífrasis de gerundio expresan valores aspectuales que en otras lenguas, en algunos casos incluso románicas, se obtienen por procedimientos léxicos (adverbios, locuciones adverbiales, etcétera)». Fernández de Castro (1999: 114) admite que el inglés posee construcciones similares pero cita a Lenz al hablar de la excepcional abundancia y complejidad de las perífrasis verbales en español: «Nuevas combinaciones de verbos principales con verbos auxiliares se han creado en masa durante la evolución analítica de los idiomas indoeuropeos modernos, y ninguna, quizá, ha desarrollado un sistema tan complicado como la lengua castellana».

Las perífrasis verbales son conjuntos de dos o más verbos que forman una unidad lingüística compleja que funciona como núcleo del predicado. El primero se denomina auxiliar y es un verbo conjugado; el segundo, o verbo principal, pues es el que aporta el contenido semántico, es una forma no personal y, según se trate de un infinitivo, un gerundio o un participio, se clasificará la perífrasis en uno de tres grupos: perífrasis de infinitivo, perífrasis de gerundio o perífrasis de participio. En ocasiones, no encontraremos ningún verbo conjugado, pues las perífrasis pueden formar parte de una oración de infinitivo o gerundio (p. ej. «Empezar a trabajar pronto era lo normal en aquella época») o de agrupaciones perifrásticas compuestas por una secuencia de dos o más auxiliares que aportan valores diferentes (p. ej. «Voy a dejar de correr» incluye las perífrasis «ir a + infinitivo y «dejar de + infinitivo»). A pesar de la unidad de sus componentes, entre el verbo auxiliar y el auxiliado de casi todas las perífrasis (excepto algunas como «haber de + infinitivo» o «deber de + infinitivo») pueden intercalarse

elementos como adverbios, complementos circunstanciales, sujetos e incisos de distinta naturaleza.

En español, las perífrasis complementan a la conjugación verbal de tiempos simples y compuestos aportando matices no representados en la misma de, por ejemplo, modo (obligación, probabilidad, necesidad, etc.) o aspecto (incoativo, durativo, puntual, etcétera.). Encontramos esta idea ampliada y sometida a pruebas empíricas en la obra de Fente Gómez (1971) *Estilística del verbo en inglés y en español*. En ella, el autor estudia fenómenos lingüísticos comparables en ambas lenguas (relacionados con el verbo) que no siempre manifiestan una identidad estructural. El capítulo III está dedicado a las construcciones perifrásticas y comienza con la advertencia de que el concepto de auxiliar varía enormemente entre el inglés y el español, especialmente en el ámbito de las frases verbales con verbos de movimiento en español y el de los auxiliares modales ingleses.

No pretendemos afirmar que tanto el inglés como el español carezcan por completo de procedimientos léxicos o sintácticos para expresar estas acciones verbales, pero sí que estos procedimientos, por causas lingüísticas o extralingüísticas que no es de nuestra incumbencia abordar aquí y que pertenecen al espíritu propio de cada lengua, se han desarrollado exageradamente o han quedado constreñidas a unos límites elementales cuando se las juzga con un criterio comparatista. (Fente Gómez, 1971: 121)

Si bien admite que algunas construcciones del inglés se acercan mucho a las españolas, el autor considera las perífrasis verbales con verbos de movimiento como uno de los rasgos más característicos del español, y se apoya en la opinión de autores como Amado Alonso:

Algunos de estos ejemplos tienen su equivalente en otras lenguas; pero, en conjunto, constituyen una manifestación de la específica forma interior del lenguaje del español, y uno de los rasgos más fisonómicos de nuestro estilo idiomático. Pues es singularidad del español la libertad, proliferación y variedad de matices de estas construcciones sintácticas y de los cambios semánticos que entrañan... (Alonso 1951 citado en Fente Gómez 1971: 123)

Fente Gómez señala que los resultados que arroja su estudio, basado en un corpus no demasiado amplio de novelas y ensayos contemporáneos en inglés y en español y sus respectivas traducciones, indican que la utilización de construcciones perifrásticas es un 10 % o 15 % superior en español que en inglés, y la diferencia es aún mayor en las que se construyen con verbos de movimiento. A pesar de sus limitaciones, creemos que este estudio constituye un interesante punto de partida para la consideración de ciertas perífrasis verbales como elementos típicos del español sin equivalentes directos en inglés. En este sentido, es reveladora la siguiente opinión del autor, que parece estar describiendo a la perfección el objeto de estudio de nuestra investigación:

Creemos que la parte más novedosa y trascendente del presente capítulo consiste en intentar una correspondencia o traducción inglesa de las principales perífrasis con verbos de movimiento o de reposo españoles. Los ejemplos recogidos no poseen, numéricamente, fuerza suficiente para asegurar una solución determinada. Son tan pocos, que nos hacen pensar (en el caso de originales ingleses traducidos al español) que éste es uno de los puntos donde la falta de estructuras originales paralelas a las nuestras influye de modo más directo en el traductor, que, inconscientemente, se siente arrastrado y encuentra más fácil la traducción literal por medio del apoyo de adverbios u otros medios sintácticos y léxicos. (Fente Gómez 1971: 130)

Todo lo que acabamos de comentar acerca del excepcional desarrollo del sistema perifrástico español no equivale a decir que el inglés no sea capaz de expresar la misma cantidad de significados y sentidos que el español, sino simplemente que lo hace por medio de otros recursos o mecanismos. Por lo tanto, las perífrasis de las que hablamos, como argumentábamos al proporcionar nuestra definición de *unique item*, no son intraducibles ni carecen de equivalentes. Simplemente no tienen un equivalente directo o inmediatamente accesible porque, generalmente, entre ellas y los elementos que les corresponden en inglés no hay identidad estructural. Esto las hace buenas candidatas al tipo de elementos que estamos buscando, ya que lo más probable es que en el inglés no aparezcan elementos que activen su utilización en el texto meta.

Por supuesto, no todas las perífrasis cumplen estas condiciones. Por el contacto entre lenguas y por el solapamiento de funciones, hay muchas que tienen un equivalente

fácilmente identificable y estandarizado en inglés. Por ejemplo, podemos mencionar «empezar a + infinitivo» que, como podríamos encontrar en cualquier diccionario bilingüe, se traduce normalmente por «*start/begin to* + infinitivo». La similitud formal es tal que sería difícil para el traductor no elegir esta perífrasis al enfrentarse a la traducción de estos dos verbos ingleses, aun cuando existen otras maneras de expresar este significado, como «ponerse a + infinitivo». Por lo tanto, no podemos esperar que todas las perífrasis se encuentren infrarrepresentadas.

Por esta razón, para los fines de este trabajo hemos intentado seleccionar todas las perífrasis que consideramos interesantes desde el punto de vista de su utilidad como objeto de investigación de la hipótesis de infrarrepresentación, tanto si a nuestro juicio presentan dificultades a la hora de establecer equivalencias con estructuras del inglés como si no, para así poder conformar tres grupos diferenciados en cuanto a los resultados que esperamos obtener. En uno de los grupos incluimos o bien perífrasis que codifican significados o formas de uso muy típicos del español e inexistentes en inglés («haber que + infinitivo»), o bien, dentro de las perífrasis más comunes, valores especiales que el inglés no comparte. Sin embargo, también incluimos perífrasis con equivalentes evidentes a modo de marco de comparación o «grupo control» (grupo I). El grupo III está conformado por aquellas perífrasis que no podemos clasificar en ninguno de los dos grupos anteriores, pues su utilización depende de muy variados factores y no estamos seguros de los resultados que pueden arrojar.

Como hemos dicho al principio, existen tres tipos de perífrasis: las de infinitivo, las de gerundio y las de participio. Muchas veces una misma construcción puede funcionar como perífrasis o no, dependiendo de diversos factores. Es el caso de «venir + gerundio», por ejemplo. Una expresión como «viene corriendo», dependiendo del contexto y de los elementos utilizados en el resto de la frase, puede significar que alguien se acerca rápidamente puesto que corre hacia nosotros o que repite la acción de correr desde hace un tiempo, por ejemplo, para entrenarse. En el primer caso no se consideraría perífrasis.

Entre algunas de las pruebas más comunes a las que suele someterse a las perífrasis para saber si pueden efectivamente ser consideradas como tales, se encuentran las siguientes:

- Pérdida de todo significado propio (gramaticalización) por parte del verbo auxiliar (algo bastante difícil de demostrar en casos como *empezar* o *deber*).
- Imposibilidad del verbo auxiliar para seleccionar sujetos y complementos propios. Los selecciona el verbo auxiliado, como en «Pedro acabó trabajando de manera gratuita»: no acabó gratuitamente, sino que trabajó de manera gratuita, y no podríamos sustituir a Pedro por, por ejemplo, una silla, salvo que cambiáramos el verbo principal.
- Imposibilidad de sustitución de la forma impersonal por elementos equivalentes (pronombres, sustantivos, adverbios...) sin cambio de significado: «No solemos vernos a menudo» no podría ser sustituida por «No lo solemos», pero «No prometemos dormirnos pronto» podría transformarse en «No lo prometemos».
- Imposibilidad de conmutación del gerundio o el participio por el adverbio interrogativo «cómo»: «Lleva cantando todo el día» puede transformarse en interrogación solo si añadimos el verbo «hacer»: «¿Qué lleva haciendo todo el día?», mientras que «Anda cojeando desde el accidente» responde a la pregunta «¿Cómo anda desde el accidente?». De la misma manera, si hay perífrasis de infinitivo, como en «No solemos vernos a menudo», no podrá transformarse en interrogación con la partícula «qué»: «¿Qué no solemos?».
- Posibilidad de anteposición y posposición de los clíticos. En caso de haber más de uno, ambos se situarían en la misma posición. «Llegó a mover la mesa» puede sustituirse por «La llegó a mover» o «Llegó a moverla». Sin embargo, con otras estructuras no perifrásticas esto no es posible: «Dejó volar la imaginación» no puede transformarse en «Dejó volarla».
- Combinación con verbos impersonales («Va a helar en el campo» es perífrasis, pero no «Va a correr al campo», frase en la que el verbo «ir» conserva su significado de movimiento).
- Formación de pasivas: los verbos auxiliares son intransitivos, por lo que lo que se pasiviza es el verbo principal, mientras que en estructuras no perifrásticas el verbo que se pasiviza es el primero, que tiene que ser transitivo: «Llegó a mover la mesa» es perífrasis, porque puede transformarse en «La mesa llegó a ser movida», pero no «Aprendió a arreglar vestidos», porque «Los vestidos aprendieron a ser arreglados» no tiene sentido.
- Formación de perífrasis de relativo: las perífrasis requieren el verbo «hacer» para focalizar el verbo principal y formar una perífrasis de relativo («Sigue

corriendo»: «Correr es lo que sigue haciendo») mientras que no sucede lo mismo con las estructuras no perifrásticas («Imagina saber volar»: «Saber volar es lo que imagina»).

Hay muchas más pruebas, algunas bastante más complejas, pero lo cierto es que ninguna es infalible: no todas las construcciones consideradas perifrásticas las cumplen, y algunas de las que no se consideran perifrásticas las pueden cumplir, por lo que hay que tener en cuenta muchos factores a la hora de clasificar una construcción como perifrástica, como veremos en este capítulo; sobre todo en lo que respecta a las perífrasis de gerundio.

La caracterización que hemos redactado acerca de cada una de las perífrasis que hemos utilizado en este estudio está basada en los trabajos de Fente Gómez (1971), Gómez Torrego (1988, 1999), Yllera (1999), Fernández de Castro (1999) y García Fernández (2012). Hemos intentado seguir a los expertos pero sin ser demasiado estrictos en cuanto a las pautas de inclusión, ya que lo que nos interesa no es tanto la definición y acotación del concepto de perífrasis, sino el uso que se da de estas construcciones verbales en textos originales y traducidos. Los criterios de inclusión y catalogación en grupos, por tanto, han sido algo subjetivos, pues nos hemos basado en la opinión de los autores acerca de las características particulares de determinadas perífrasis en español, y en información sobre posibles equivalentes recogida en trabajos contrastivos y los distintos diccionarios bilingües consultados (incluidos en la bibliografía), para después, a través de la información obtenida sobre la existencia o no de posibles elementos activadores en la lengua origen, clasificarlas según las posibilidades que pensamos que tendrán de estar más o menos representadas en traducciones.

4.2. Selección de perífrasis

La lista de perífrasis confeccionada se basa principalmente en la proporcionada por Gómez Torrego (1999). De este listado se omitieron siete perífrasis de infinitivo. Las cuatro primeras son perífrasis modales muy habituales en español, que cuentan con diversos elementos activadores evidentes en inglés. Por este motivo, consideramos que

arrojarían una cantidad excesiva de concordancias sin ofrecer resultados demasiado sorprendentes. Dos de ellas, «tener que + infinitivo» y «deber + infinitivo», expresan obligación y pueden traducirse por «*have to/must/should* + infinitivo». Las otras dos son construcciones con respecto a cuya naturaleza perifrástica no todos los gramáticos se ponen de acuerdo, dado que sus auxiliares no están del todo gramaticalizados. Se trata de «deber de + infinitivo» (que expresa probabilidad y puede traducirse por «*must* + infinitivo» o adverbios que expresan duda) y «poder + infinitivo» (que expresa posibilidad y puede traducirse por «*can/be able to* + infinitivo» o «*may/might/should* + infinitivo»). Las tres últimas son «ir a + infinitivo» («*to be going to* + infinitivo») y «terminar/acabar de + infinitivo» («*to finish* + gerundio», por ejemplo). Aunque algunos de sus valores semánticos pueden proporcionar resultados interesantes, pues no siempre tienen traducciones evidentes en inglés (por ejemplo, en enunciados negativos, «terminar/acabar de + infinitivo» no expresa valor perfectivo sino valores estilísticos de ansiedad o desazón: «No termino de entender por qué», *I just can't understand why*), las dejamos fuera de nuestra selección porque la mayor parte de las concordancias corresponden a sus otros valores más habituales y los casos que nos interesarían se perderían entre los demás ejemplos, mucho más numerosos.

Con respecto a las perífrasis de gerundio, hemos incluido todas las listadas por Gómez Torrego (1988) a excepción de «estar + gerundio», dado que en ocasiones no se la considera una construcción perifrástica sino parte del paradigma verbal (Yllera, 1999: 3402). Además, cuenta con una correspondencia directa y muy frecuente: «*to be* + gerundio». Hemos añadido, sin embargo, una de las construcciones que este autor considera semiperifrásticas, «acabar/terminar + gerundio» amparándonos en el hecho de que sus valores y equivalentes son idénticos a los de «acabar/terminar por + infinitivo», dos perífrasis incluidas en la selección.

Hemos decidido no incluir ninguna perífrasis de participio en este trabajo, puesto que por su particular naturaleza son objeto de las mayores discrepancias en cuanto a definición y a identificación. El participio adopta a veces función adjetival y ocasiona que el conjunto verbal no sea considerado perífrasis. Como admite Yllera (1999: 3425), los autores no se ponen de acuerdo, y en demasiadas ocasiones se contradicen, acerca de los criterios que deben utilizarse para saber, en primer lugar, qué conjuntos verbales funcionan a veces como perífrasis —por ejemplo, mientras que Gómez Torrego (1988) enumera hasta ocho perífrasis de participio, Fernández de Castro (1999) admite como perifrásticos únicamente dos auxiliares: *tener* y *llevar*— y, dentro

de estos casos, cuándo una expresión determinada puede ser o no considerada perífrasis. A veces, los criterios son tan restrictivos, que ningún ejemplo del corpus podría ser considerado perifrástico. Por lo tanto, sólo analizaremos algunas perífrasis de infinitivo y de gerundio, en torno a las cuales no parece haber tanta polémica y cuyos criterios identificadores están mucho más claros y unificados. Decimos «parece» porque, de hecho, el debate en torno a esta categoría lingüística no se limita a las perífrasis de participio, como resume Rabadán (2006: 299-300) en una de las notas de su artículo sobre el tema:

7. There is not agreement among experts concerning the number of periphrasis [*sic*] in Spanish, modal or otherwise (Fernández de Castro 1990:30–32). For different accounts of the inventory of this type of sequences in Spanish see Iglesias Bango (1988), Alarcos Llorach (1994:261–263), and Gutiérrez Ordóñez (1986:37ff). Morera (1991) denies the existence of periphrastic constructions on the sole basis of their syntactic behaviour. On the gradability and hierarchic nature of periphrastic constructions see Lamiroy (1994:72).

La lista final quedó configurada de la siguiente manera:

➤ Perífrasis de infinitivo

1. EMPEZAR A + INFINITIVO
2. COMENZAR A + INFINITIVO
3. VOLVER A + INFINITIVO
4. ACABAR POR + INFINITIVO
5. TERMINAR POR + INFINITIVO
6. HABER DE + INFINITIVO
7. HABER QUE + INFINITIVO
8. ACERTAR A + INFINITIVO
9. LLEGAR A + INFINITIVO
10. ALCANZAR A + INFINITIVO
11. VENIR A + INFINITIVO
12. QUERER + INFINITIVO
13. PONERSE A + INFINITIVO
14. ECHAR(SE) A + INFINITIVO
15. ROMPER A + INFINITIVO
16. ESTAR A PUNTO DE + INFINITIVO
17. DEJAR DE + INFINITIVO
18. PARAR DE + INFINITIVO
19. SOLER + INFINITIVO

➔ Perífrasis de gerundio:

1. SEGUIR + GERUNDIO
2. CONTINUAR + GERUNDIO
3. ACABAR + GERUNDIO
4. TERMINAR + GERUNDIO
5. ANDAR + GERUNDIO
6. IR + GERUNDIO
7. VENIR + GERUNDIO
8. LLEVAR + GERUNDIO

En nuestra caracterización, incluimos una descripción del significado de la perífrasis y los distintos valores que puede adoptar según el contexto, además de la traducción de estos valores al inglés y, siempre que las haya, opciones alternativas para trasladar estos equivalentes al español sin utilizar la perífrasis.

4.3. Corpus utilizado para el análisis

Este capítulo está basado en el trabajo realizado para la publicación del artículo *On the lower frequency of occurrence of Spanish verbal periphrases in translated texts as evidence for the unique items hypothesis* (Martínez Vilinsky 2012). En dicho artículo analizábamos las frecuencias de aparición de un conjunto de perífrasis verbales españolas en textos originales escritos en español y textos traducidos del inglés al español, con el objetivo de comprobar si se cumplía la hipótesis de la infrarrepresentación de elementos únicos del español en textos traducidos que hemos presentado y explicado en la introducción teórica. En aquella ocasión utilizamos el corpus comparable en su estado inicial (fase I) y no contábamos aún con el corpus paralelo para contrastar nuestras ideas acerca de la posible influencia de los originales ingleses en el texto traducido.

Los resultados fueron muy positivos en cuanto a la confirmación de la hipótesis de la que partíamos, es decir, que las perífrasis elegidas se encontrarían en gran medida infrarrepresentadas en traducciones. Tomando todas las perífrasis en conjunto, observamos una diferencia estadísticamente significativa entre textos traducidos del inglés al español y originales escritos en español, a pesar de que solo en el 63% de las

perífrasis (17 de 27) encontramos una infrarrepresentación en traducciones que pudiera ser considerada significativa en términos estadísticos. En cuanto a la división por grupos, en líneas generales (entraremos en detalles durante el análisis que sigue) hallamos que nuestras predicciones no se cumplieron en todos los casos. En dos de los grupos, el de perífrasis típicas de nuestra lengua que no cuentan con equivalentes formales directos en inglés, y el de las perífrasis sobre cuyo comportamiento no estábamos seguros —a pesar de contener perífrasis mejor representadas en traducciones o sin diferencias notables entre subcorpus—, se observó, en conjunto, una clara infrarrepresentación, mucho más marcada en el primero que en el segundo. En el caso del primer grupo, sin embargo, a pesar de que teóricamente incluía elementos con equivalentes fácilmente identificables en inglés, había perífrasis que, aun contando con equivalentes directos, se encontraban infrarrepresentadas. Tomando todas las perífrasis de este grupo en conjunto, obteníamos una infrarrepresentación significativa. Es decir, los resultados verificaban la hipótesis de la infrarrepresentación en los tres grupos.

Como hemos explicado en los apartados 3.2.2 y 3.3.1.6. del capítulo Metodología, quedaban muchas dudas por resolver en cuanto a qué influencia podían tener los cambios de variables o parámetros de compilación en un corpus y, dado que para esta tesis doctoral consideramos necesario cambiar algunos de los textos incluidos, aprovechamos esta oportunidad para repetir el análisis.

Las siguientes son las tablas-resumen correspondientes a la fase I (tabla 4) y la fase II (tabla 5) de compilación del corpus, creadas a partir de la síntesis de los resultados obtenidos mediante el análisis y cómputo de las concordancias. Para cada perífrasis hemos incluido el número de veces que aparece en el subcorpus de originales en español y, para facilitar la comparación de los resultados, la normalización del número de apariciones, es decir, la cantidad de veces que la perífrasis aparecería en un corpus de cien mil palabras. A continuación, se proporcionan los mismos datos para el subcorpus de traducciones al español. Finalmente, se compara cada par de frecuencias por medio del test de log-likelihood, al que se asigna un valor positivo o negativo según si la perífrasis aparece mayor número de veces en ES o en TRAD, respectivamente. El valor de p nos indica si aquellas perífrasis con valor positivo, es decir, las que poseen una frecuencia de uso menor en traducciones, pueden ser tenidas en cuenta para validar la hipótesis de infrarrepresentación: si la diferencia no es significativa, se indica mediante la abreviatura «n. s.»; si lo es, se indica con qué nivel de significación estadística.

Perífrasis	Total ES	Frecuencia ES normalizada	Total TRAD	Frecuencia TRAD normalizada	LL	p	Verificación de la hipótesis
Grupo I							
Empezar a + INF	129	39,8	265	50,9	-5,45	<0,05	NO
Comenzar a + INF	52	16	24	4,6	+28,00	<0,0001	SÍ
Volver a + INF	190	58,6	319	61,3	-0,25	n. s.	NO
Seguir + GER	150	46,2	192	36,9	+4,24	<0,05	SÍ
Continuar + GER	25	7,7	14	2,7	+10,50	<0,01	SÍ
Acabar por + INF	5	1,5	2	0,4	+3,13	n. s. ⁹⁹	NO
Terminar por + INF	7	2,2	0	0	+13,40	<0,001 ¹⁰⁰	SÍ
Acabar + GER	33	10,2	10	1,9	+26,22	<0,0001	SÍ
Terminar + GER	8	2,5	2	0,4	+7,24	<0,01 ¹⁰¹	SÍ
Subtotal Grupo I	599	184,6	828	159,1	+7,63	<0,01	SÍ
GRUPO II							
Haber de + INF	63	19,4	35	6,7	+26,76	<0,0001	SÍ
Haber que + INF	147	45,3	67	12,9	+80,29	<0,0001	SÍ
Acertar a + INF	6	1,8	0	0	+11,48	<0,001 ¹⁰²	SÍ
Llegar a + INF	50	15,4	46	8,8	+7,37	<0,01	SÍ
Alcanzar a + INF	0	0	6	1,1	-5,81	<0,05 ¹⁰³	NO
Andar + GER	35	10,8	24	4,6	+10,52	<0,01	SÍ
Ir + GER	127	39,1	30	5,8	+119,00	<0,0001	SÍ
Venir + GER	0	0	5	1	-4,85	<0,05 ¹⁰⁴	NO
Venir a + INF	1	0,3	0	0	+1,91	n. s. ¹⁰⁵	NO
Subtotal Grupo II	429	132,2	213	40,9	+211,69	<0,0001	SÍ
GRUPO III							
Querer + INF	22	6,8	1	0,2	+34,85	<0,0001	SÍ
Ponerse a + INF	42	12,9	33	6,3	+9,48	<0,01	SÍ
Echar(se) a + INF	31	9,5	77	14,8	-4,47	<0,05	NO
Romper a + INF	0	0	10	1,9	-9,69	<0,01 ¹⁰⁶	NO
Estar a punto de + INF	39	12	44	8,4	+2,53	n. s.	NO
Dejar de + INF	81	25	80	15,4	+9,38	<0,01	SÍ
Parar de + INF	11	3,4	6	1,1	+4,80	<0,05	SÍ
Soler + INF	76	23,4	88	16,9	+4,28	<0,05	SÍ
Llevar + GER	9	2,8	17	3,3	-0,16	n. s.	NO
Subtotal Grupo III	311	95,9	356	68,4	+18,67	<0,0001	SÍ
Total	1339	412,7	1397	268,4	+125,16	<0,0001	SÍ

Tabla 4 - Comparación de frecuencias entre subcorpus ES y TRAD y significación estadística de los resultados (FASE I)

Los diferentes colores que pueden observarse en la última columna provienen de las categorías que surgen al realizar el test de significación estadística de log-likelihood: además de las perífrasis claramente infrarrepresentadas (en verde) y las claramente sobrerrepresentadas (en rojo), se generan dos categorías intermedias, las que en principio presentan mayor frecuencia en originales o en traducciones pero no hay diferencias realmente significativas entre subcorpus que avalen esta clasificación en ninguno de los dos grupos (en amarillo y en fucsia, respectivamente). Los resultados del test exacto de Fisher para aquellas frecuencias que no llegan a sumar cinco en todas las

⁹⁹ Test exacto de Fisher: 0,115252 (n. s.)¹⁰⁰ Test exacto de Fisher: 0,001232 (p < 0,01)¹⁰¹ Test exacto de Fisher: 0,017128 (p < 0,05)¹⁰² Test exacto de Fisher: 0,003207 (p < 0,01)¹⁰³ Test exacto de Fisher: 0,0887 (n. s.)¹⁰⁴ Test exacto de Fisher: 0,164013 (n. s.)¹⁰⁵ Test exacto de Fisher: 0,2621 (n. s.)¹⁰⁶ Test exacto de Fisher: 0,009054 (p < 0,01)

celdas de la tabla de contingencia que se construye para calcular el test de log-likelihood —nueve casos en el corpus de la fase I y seis casos en el de la fase II— se recogen en notas a pie. En algunas ocasiones no coinciden con el valor p de este último test, pero la casilla se ha pintado del color que indica el test de Fisher.

Perífrasis	Total ES	Frecuencia ES normalizada	Total TRAD	Frecuencia TRAD normalizada	LL	P	Verificación de la hipótesis
Grupo I							
Empezar a + INF	200	54,5	312	63,9	-3,11	n. s.	NO
Comenzar a + INF	55	15	27	5,5	+19,42	< 0,0001	SÍ
Volver a + INF	199	54,2	378	77,4	-17,04	< 0,0001	NO
Seguir + GER	170	46,3	183	37,4	+3,95	< 0,05	SÍ
Continuar + GER	24	6,5	14	2,9	+6,30	< 0,05	SÍ
Acabar por + INF	7	1,9	5	1	+ 1,15	n. s.	NO
Terminar por + INF	7	1,9	0	0	+11,85	< 0,001 ¹⁰⁷	SÍ
Acabar + GER	38	10,3	22	4,5	+10,11	< 0,01	SÍ
Terminar + GER	9	2,4	4	0,8	+3,67	n. s.	NO
Subtotal Grupo I	709	193,1	945	193,4	-0,00	n. s.	NO
Grupo II							
Haber de + INF	63	17,2	131	26,8	-8,86	< 0,01	NO
Haber que + INF	159	43,3	75	15,3	+59,61	< 0,0001	SÍ
Acertar a + INF	7	1,9	0	0	+11,85	< 0,001 ¹⁰⁸	SÍ
Llegar a + INF	66	18	48	9,8	+10,32	< 0,01	SÍ
Alcanzar a + INF	1	0,3	10	2	-6,20	< 0,05 ¹⁰⁹	NO
Andar + GER	35	9,5	26	5,3	+5,15	< 0,05	SÍ
Ir + GER	148	40,3	80	16,4	+44,67	< 0,0001	SÍ
Venir + GER	1	0,3	7	1,4	-3,51	n. s. ¹¹⁰	NO
Venir a + INF	1	0,3	6	1,2	-2,68	n. s. ¹¹¹	NO
Subtotal Grupo II	481	131	383	78,4	+56,74	< 0,0001	SÍ
Grupo III							
Querer + INF	26	7,1	15	3,1	+6,97	< 0,01	SÍ
Ponerse a + INF	43	11,7	115	23,5	-16,69	< 0,0001	NO
Echar(se) a + INF	37	10,1	85	17,4	-8,17	< 0,01	NO
Romper a + INF	0	0	2	0,4	-2,24	n. s. ¹¹²	NO
Estar a punto de + INF	42	11,4	49	10	+0,39	n. s.	NO
Dejar de + INF	95	25,9	101	20,7	+2,46	n. s.	NO
Parar de + INF	12	3,3	9	1,8	+1,72	n. s.	NO
Soler + INF	77	21	102	20,9	+0,00	n. s.	NO
Llevar + GER	10	2,7	13	2,7	+0,00	n. s.	NO
Subtotal Grupo III	342	93,2	491	100,5	-1,16	n. s.	NO
Total ¹¹³	1532	417,3	1819	372,3	+10,79	< 0,01	SÍ

Tabla 5 - Comparación de frecuencias entre subcorpus ES y TRAD y significación estadística de los resultados (FASE II)

Como vemos, los resultados generales no varían drásticamente entre una fase y otra, dado que muchos de los textos, sobre todo en el subcorpus de originales, son los

¹⁰⁷ Test exacto de Fisher: 0,002675 ($p < 0,01$)

¹⁰⁸ Test exacto de Fisher: 0,002675 ($p < 0,01$)

¹⁰⁹ Test exacto de Fisher: 0,02973 ($p < 0,05$)

¹¹⁰ Test exacto de Fisher: 0,149489 (n. s.)

¹¹¹ Test exacto de Fisher: 0,25095 (n. s.)

¹¹² Test exacto de Fisher: 0,510077 (n. s.)

¹¹³ El aplicar el test de log-likelihood a los subtotales de los diferentes grupos y al grupo completo de perífrasis fue recomendación de un revisor anónimo en la primera versión del artículo de *Across*. Puede ser discutible por no tener en cuenta las cifras que aporta cada perífrasis o grupo, como veremos al final del capítulo, pero, siempre que analicemos también cada caso de manera individual, nos parece una idea interesante, ya que aporta elementos de análisis y diferentes perspectivas para valorar los resultados.

mismos, pero sí nos hemos encontrado con alguna diferencia fundamental al tomar cada grupo por separado, que explicaremos en profundidad a lo largo de este capítulo.

4.4. Resultados y discusión

A continuación, analizaremos las 27 perífrasis; por un lado, de manera individual, por otro, dentro del contexto del grupo al que pertenece cada una y, finalmente, de manera general como elemento único de análisis. Proporcionaremos en cada caso los datos de ambas fases del subcorpus comparable; es decir de originales españoles (ES) y de traducciones inglés-español (TRAD) para permitir una comparación más sencilla de los resultados y mostrar la manera en la que ciertos cambios en los parámetros de compilación pueden alterar los datos obtenidos.

Para la fase I del corpus se especifican las expresiones de búsqueda utilizadas en los campos de WordSmith Search Words (las palabras clave, generalmente las distintas flexiones del auxiliar) y Context Words (las requeridas en el contexto para formar la perífrasis, por lo general, las terminaciones de gerundio o infinitivo o el elemento intercalado entre auxiliar y verbo principal, es decir, la conjunción *que* o la preposición *a*). Los horizontes de búsqueda de contexto (Context Search Horizons), es decir, el límite de palabras a la derecha o a la izquierda en el que deben encontrarse las expresiones especificadas en el contexto, se establecieron para todas las perífrasis en OL a 4R, desde la palabra clave hasta cuatro palabras a la derecha, a excepción de perífrasis como «haber de + infinitivo» o «haber que + infinitivo», en las que no se intercalan elementos entre el auxiliar y la forma no personal. En estos casos, para acotar la búsqueda, se establecieron los horizontes del contexto en 0L a 1R (una palabra a la derecha). También se modificaron los horizontes en casos en que las palabras de contexto se situaban a la izquierda en lugar de la derecha.

La búsqueda de las perífrasis se realizó de la siguiente manera: con ayuda de una tabla de verbos españoles conjugados (Rubio Moraiz 2008), intentamos recoger todas las flexiones posibles de los auxiliares que forman parte de las perífrasis estudiadas en expresiones de extensión lo más reducida posible, dado que WordSmith sólo permite realizar búsquedas de hasta 15 alternativas (con un número máximo de 80 caracteres).

Aun así, con verbos como *ir*, cuya conjugación es particularmente compleja, fue necesario realizar la búsqueda en dos o más etapas. Las expresiones estaban conformadas por las raíces que más formas permitieran encontrar pero al mismo tiempo generaran la menor cantidad de resultados no deseados.

En el caso de las perífrasis de gerundio, como contexto se definió la expresión «*ANDO*/*IENDO*/*YENDO*/*ÁANDO*/*IÉENDO*/*YÉENDO*», para indicarle al programa que mostrara resultados que contuvieran terminaciones de gerundio y tener en cuenta, al mismo tiempo, los clíticos que pudieran acompañarlas, aunque, para recoger todos los casos, tuvimos que desechar muchas palabras que contenían esa combinación de letras pero no eran gerundios. No se hizo lo mismo en el caso de las perífrasis de infinitivo, pues en la sección de contexto, muchas veces había que colocar el nexo de unión del auxiliar y la forma no personal (*que, de o a*). En los casos en los que el ruido era excesivo, se intentó utilizar la siguiente fórmula: *AR*/*ER*/*IR*, aunque es evidente que tampoco arrojaba resultados demasiado precisos. Sin embargo, como hemos explicado, el corpus no era lo suficientemente grande para impedir el cribado manual de los resultados no deseados.

En las preferencias de búsqueda de WordSmith se definió el orden de aparición de las concordancias de la siguiente manera. Primero, ordenadas según el autor del texto en el que aparecen, para poder apreciar si existen diferencias significativas con respecto al número de veces que cada autor utiliza las expresiones analizadas. En segundo lugar, en orden alfabético por la primera letra de la palabra clave, para agrupar las distintas flexiones de los auxiliares y, en tercer lugar, por la primera palabra a la derecha de la palabra clave. Se marcaron los casos que no interesaban, se eliminaron mediante el comando *zap* y se copiaron las concordancias a un archivo de Microsoft Office Word como imagen (metarchivo de Windows).

Cuando tratemos la fase II del corpus también se indicará la fórmula utilizada en Corpus Workbench¹¹⁴, que hemos intentado que fuera la más eficaz a la hora de hallar todos los casos de la perífrasis correspondiente en el corpus sin recurrir, salvo en algún caso puntual, al etiquetado. Esta decisión se debe a que, a medida que íbamos realizando las búsquedas sobre el corpus etiquetado y comparábamos los resultados con los del corpus sin etiquetar, nos íbamos dando cuenta de que los números no coincidían.

¹¹⁴ La descripción del software y el método de realización de búsquedas en el corpus se ha explicado ya en el apartado 3.4.2 del capítulo Metodología.

Esto resultó un poco decepcionante, pues hay casos en los que fiarnos del etiquetado nos habría permitido afinar más la búsqueda. Por ejemplo, la expresión «[word="(¿|í)?acab.*" %c] []{0,5} "por" [pos=".*inf"]» proporciona la misma cantidad de concordancias válidas (7) que la misma expresión sin especificar que queremos que aparezca un infinitivo detrás («pos=".*inf»), es decir, simplemente «[word="acab.*" %c] []{0,5} "por"», con la ventaja de que nos ahorramos el tener que eliminar los casos en los que no se da esta condición. En el caso de las perífrasis de gerundio, en un primer momento probamos a utilizar la etiqueta correspondiente a las formas de gerundio. Por ejemplo, «[word="(¿|í)?contin(u|ú).*" %c] []{0,5} [pos=".*ger"]» arrojaba un resultado de 18 casos pertinentes de 21 concordancias. Sin embargo, pronto nos dimos cuenta de que no todas las formas de gerundio habían sido etiquetadas como tales, por lo que debimos utilizar una expresión más enrevesada que, a pesar de generar a veces bastante ruido, permitiera extraer todos los casos pertinentes, como «[word="(¿|í)?contin(u|ú).*" %c] []{0,5} [word=".*(a|á)ndo.*|.*i(e|é)ndo.*"]», que proporciona 25 casos válidos de 33 concordancias; es decir, 8 más (casi un 40 %).

Otra de las funciones que más útil nos hubiese resultado, sobre todo en este capítulo, en el que trabajamos con verbos, es la del lematizado. Utilizar la opción de búsqueda por lema (p. ej. «[lem="acabar.*"] []{0,5} "por"») nos habría ahorrado el tener que especificar todas las conjugaciones posibles del verbo mediante expresiones regulares. Sin embargo, estas búsquedas no arrojaron resultados completos en prácticamente ninguno de los casos. Hay excepciones, como «querer + infinitivo» o el verbo *haber*, que es el mejor lematizado de todos, pero, en general, hemos comprobado que existen algunos errores en la lematización que afectan a una gran parte de nuestras búsquedas, puesto que en muchas ocasiones quedarían casos sin salir a la luz si nos fiáramos de ella, lo que puede resultar aceptable en muchas ocasiones (cuando el corpus es muy grande, por ejemplo), pero, en este trabajo, en el que pretendemos realizar búsquedas exhaustivas, no lo sería.

Esto es una desventaja que es importante tener en cuenta a la hora de compilar un corpus y que hemos descubierto a base de realizar numerosas búsquedas. En nuestro caso decidimos que no valía la pena corregir el etiquetado, puesto que el tamaño reducido de nuestro corpus nos ha permitido ir haciendo un proceso de revisión sobre la marcha, pero si el corpus va a ser utilizado por otros investigadores y en varios proyectos distintos, lo más eficiente es realizar previamente una revisión minuciosa del etiquetado.

No obstante, hay que añadir que el etiquetado sí fue útil en fases posteriores, dado que, cuando ya habíamos comprobado que utilizarlo proporcionaba en ciertos casos el mismo número de resultados pertinentes, algunas expresiones regulares eliminaban el ruido excesivo. Un ejemplo es el de «`[word="(¿|¡)?ll(e|é)g.*" %c] []{0,5} "a" [pos=".*inf"]`» en el que la etiqueta «inf» nos evitó el tener que cribar más de 200 concordancias para extraer la distribución por autores.

Dentro de los resultados de la fase II también incluiremos tanto la distribución por textos proporcionada por Corpus Workbench como los resultados del análisis del corpus paralelo (EN), que no teníamos a nuestra disposición en la fase I. Los listados de concordancias de cada perífrasis y de los demás elementos analizados en los capítulos siguientes pueden consultarse en el apéndice 2, incluido en la versión electrónica de este trabajo. No se proporcionan en papel porque ocuparían varios cientos de páginas.

4.4.1. GRUPO I

En este grupo incluimos aquellas perífrasis que creemos que estarán bien representadas en el subcorpus de traducciones dado que cuentan con equivalentes fácilmente identificables en inglés (a pesar de que en algunos casos pueda no haber correspondencia formal o estructural) porque (a) constituyen la forma más habitual de expresar este valor o significado en español o (b) no existen estructuras equivalentes alternativas en español. Son las siguientes nueve perífrasis, en su mayoría de valor aspectual:

1. EMPEZAR A + INFINITIVO
2. COMENZAR A + INFINITIVO
3. VOLVER A + INFINITIVO
4. SEGUIR + GERUNDIO
5. CONTINUAR + GERUNDIO
6. ACABAR POR + INFINITIVO
7. TERMINAR POR + INFINITIVO
8. ACABAR + GERUNDIO
9. TERMINAR + GERUNDIO

4.4.1.1. Empezar a + infinitivo - Comenzar a + infinitivo

Se trata de dos perífrasis de significado equivalente y valor ¹¹⁵ aspectual incoativo o ingresivo que indican el comienzo de una acción. *Empezar* es más general y también más coloquial que *comenzar*. No incluimos «empezar + infinitivo» porque carezca de equivalentes directos en inglés, dado que no es el caso en absoluto, sino como marco de comparación para el resto de las perífrasis incoativas que veremos más adelante («comenzar a + infinitivo», «echar a + infinitivo», «romper a + infinitivo» y «ponerse a + infinitivo»). Se trata de la opción más neutra y habitual para traducir el significado incoativo de «*start/begin to* + infinitivo» o «*start/begin* + gerundio» en inglés, la que puede encontrarse en cualquier diccionario bilingüe. La similitud formal es tal que probablemente resulta difícil para el traductor no seleccionar esta perífrasis española cuando se encuentra con *start* o *begin*. Las otras maneras de expresar este significado pueden funcionar como opciones idiosincráticas o estilísticas del traductor o simplemente indicar una mayor independencia con respecto al texto original. Otras expresiones inglesas que pueden activar la utilización de estas perífrasis son: *set about*, *set off*, *set (out) to*, *break out* o *get +-ing*.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

EMPEZ*/EMPIEZ*/EMPIEC*/EMPIÉZ*/EMPIÉC*/EMPEC*
 COMENZ*/COMIENZ*/COMIÉNZ*/COMENC*/COMIENC*/COMIÉNC*
 Context words: A
 Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Empezar a + INF	129	39,8	265	50,9	-5,45	< 0,5	No
Comenzar a + INF	52	16	24	4,6	+28,00	< 0,0001	Sí

Como era de esperar, por los motivos que ya hemos alegado, en esta primera fase la perífrasis «empezar a + infinitivo» estaba bien representada en el subcorpus de

¹¹⁵ *Valor* es el término utilizado en la bibliografía sobre perífrasis verbales en lengua española para referirse a aquellos matices aspectuales o modales que no pueden expresarse a través de la conjugación verbal. Podrían denominarse también sentidos, significados o «valores semánticos».

traducciones y no verificaba la hipótesis de infrarrepresentación. De hecho, encontramos una abundancia de casos en TRAD que indica que se utiliza más incluso que en español original. Consideramos necesario analizar este resultado junto con el de «comenzar a + infinitivo» y las otras perífrasis incoativas estudiadas en este trabajo, dado que era posible que la mayor presencia de «empezar a + infinitivo» fuera en detrimento de estas otras perífrasis que, a su vez, podrían presentar mayor frecuencia de aparición en el subcorpus de originales en español. Eso indicaría que el español tiene más variedad de recursos para expresar la idea de comienzo, y no todos se activan en el proceso de traducir *start* o *begin*.

Esta teoría se sustenta con los resultados de «comenzar a + infinitivo», una variante menos habitual y más formal, y «ponerse a + infinitivo» (grupo III) que, en efecto, en el corpus de la fase I presentan una frecuencia de aparición significativamente menor en el subcorpus de traducciones, debido probablemente a que *start* y *begin* activan el uso de «empezar a + infinitivo». El traductor, por consiguiente, no siente la necesidad de buscar alternativas menos evidentes.

Sin embargo, no ocurre lo mismo con expresiones que tienen algún tipo de equivalente estilístico en inglés como «echarse a + infinitivo» o «romper a + infinitivo», que pueden traducirse como «*burst out* + -ing», «*burst into* + sustantivo», «*break into a* + sustantivo», *set off* (para andar y correr), etcétera.

Como nota aparte, «comenzar a + infinitivo», además de ser menos habitual en ambos subcorpus, aparece en el de traducciones fundamentalmente en solo dos de los textos, el de la novela de Hunter (traducido por Carlos Milla Soler) y el de Mosley (traducido por Rosa Corgatelli). Hay una única concordancia en el texto de Ellroy (Carlos Gardini).

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?emp(i?)(e|é)(z|c).*" %c] []{0,5} "a";

[word="(¿|i)?com(i?)(e|é)n(z|c).*" %c] []{0,5} "a";

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Empezar a + INF	200	54,5	312	63,9	-3,11	n. s.	No
Comenzar a + INF	55	15	27	5,5	+19,42	< 0,0001	Sí

Al cambiar los textos del corpus, comprobamos que sigue habiendo una frecuencia mayor de apariciones de «empezar + infinitivo» en las traducciones, pero esta diferencia entre subcorpus no parece ya tan grande, dado que no supera el umbral del valor crítico más pequeño del test de significación estadística, que es 3,84.

El comando «group Last match text_author» de CQP nos permite, tras realizar una búsqueda de las perífrasis existentes en nuestro corpus, agrupar estos resultados por autor. En la tabla siguiente podemos ver las frecuencias de uso absolutas en los distintos textos y, en la última columna, las frecuencias normalizadas, es decir, multiplicadas por 100 000 y divididas por el número de palabras de la obra en la que aparece la perífrasis. De esta manera, podemos establecer relaciones y comparar el uso que hacen de estas estructuras los distintos autores y traductores. Por ejemplo, a pesar de contar con el mismo número de casos cada uno en frecuencias absolutas, sabemos que Madrid hace un uso relativamente más abundante de esta perífrasis que Martín, puesto que supera en casi ocho a este último en la columna de frecuencias normalizadas. La razón es que su texto tiene casi quince mil palabras menos y, por tanto, le correspondería una frecuencia menor si hiciera un uso equivalente al de Martín.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Núñez	71	167,2	Cornwell	143	130,1
Giménez	64	75,4	Leonard	60	77,1
Vázquez	33	37,6	Highsmith	48	48,8
Madrid	12	29	Elroy	35	31,9
Martín	12	21,3	Grafton	26	28
Martínez	8	14,7			
Total	200	54,5		312	63,9

Tabla 6 - Distribución de «empezar a + INF» por texto

Como podemos ver en la tabla de distribución por textos, la novela de Núñez, la única que se ha añadido al subcorpus de originales en español, contiene un número de apariciones de esta perífrasis superior a todas las demás, lo que hace que la frecuencia en ES se eleve de 129 a 200 (un 35,5 % más que en la fase I). Este hecho no deja de resultar curioso, dado que en el caso de otras perífrasis, Núñez suele situarse al final de la clasificación, debido a que el número de palabras totales de la novela es el segundo menor. Si tomamos como válida nuestra suposición de que la utilización de «empezar a + infinitivo» en detrimento de otras perífrasis responde quizás a o un apego excesivo al original o a una falta de variedad de recursos del traductor, lo mismo podría decirse de

esta autora. En TRAD, por otro lado, solo ha habido un incremento del 17 %, a pesar de que una de las novelas nuevas, la de Cornwell, acapara el 46 % de los casos.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martínez	37	68,1	Grafton	26	28
Madrid	12	29	Cornwell	1	0,9
Núñez	2	4,7	Elroy	0	0
Vázquez	2	2,3	Highsmith	0	0
Giménez	1	1,2	Leonard	0	0
Martín	1	1,8			
Total	55	15		27	5,5

Tabla 7 - Distribución de «comenzar a + INF» por texto

En cuanto a «comenzar + infinitivo», se sigue confirmando su infrarrepresentación en traducciones. Un dato interesante es que todos los casos hallados en TRAD se encuentran en tiempo pasado, mientras que en ES, aunque escasos, hallamos ejemplos del uso de presente o subjuntivo (5). Un único autor (Martínez Reverte) abarca el 67 % de las apariciones en ES, mientras que en TRAD la utilización es aún más idiosincrática, puesto que solo dos traductores utilizan esta perífrasis, y uno de ellos (Moya) es responsable del 96 % de los casos.

Mientras que Juan Madrid utiliza cada forma un 50 % de las veces (12 y 12), hay autores que muestran clara preferencia por una estructura u otra, como Martínez Reverte, que es quien menos utiliza «empezar a + infinitivo» (en 8 ocasiones) y, a la vez, quien más utiliza «comenzar a + infinitivo» (en 37 ocasiones). Núñez, por su parte, es quien más utiliza «empezar» (71 casos), mientras que solo utiliza «comenzar» en 2 ocasiones. Aparte de Martínez Reverte y Madrid, en todo caso, la utilización de «comenzar» parece ser bastante escasa en español; un fenómeno que se confirma y se amplifica en las traducciones.

- CORPUS PARALELO¹¹⁶

¹¹⁶ Las categorías que incluimos en las tablas de los apartados denominados «corpus paralelo» han sido creadas *ad hoc* para cada perífrasis, aunque es evidente que muchas se repiten. Son categorías que agrupan el aspecto gramatical con el semántico, ya que hay ocasiones en las que no nos basta con saber que una expresión equivalente en inglés pertenece a determinada categoría gramatical para distinguirla de las demás. A veces necesitamos saber que las palabras o expresiones recogidas en dicha categoría expresan un significado concreto, como podrían ser los de duración o reiteración. Por ejemplo, en el caso de algunas perífrasis, podríamos contar con una categoría denominada «*phrasal verbs*» que expresan el mismo valor que dicha perífrasis y, a pesar de ello, incluir muchos otros *phrasal verbs* dentro de la categoría general de las reformulaciones, pues el significado que aportan a la oración no nos interesa como equivalente de la perífrasis en cuestión.

Elemento activador EN	Frecuencia	Porcentaje
Begin	118	37,8
Start	96	30,8
∅ ¹¹⁷	29	9,3
Link verbs (get/become/turn/go + ADJ)	18	5,8
Reformulación ¹¹⁸	13	4,2
Tiempos continuos	13	4,2
Phrasal verbs (come out, come on (2), go off, walk off, set about, set up)	7	2,2
Verbo seguido de preposición: (get to (2), come to, work off, work from, dawn on)	6	1,9
Present participle clause	4	1,3
Marcador temporal (now, finally)	4	1,3
Verbo <i>feel</i> seguido de infinitivo sin <i>to</i>	3	1
Commence	1	0,3
TOTAL	312	100

Tabla 8 - Equivalentes de «empezar a + INF» en EN

De los posibles equivalentes que habíamos adelantado mediante la búsqueda en diccionarios bilingües, se confirma, por un lado, la predominancia de *begin* y de *start* como desencadenantes principales de la utilización de esta perífrasis, dado que abarcan un 68,6 % de las concordancias, y, por otro, la presencia de *phrasal verbs* como *set about* (los demás que habíamos mencionado no han aparecido en nuestro subcorpus EN).

La suma de los porcentajes de los casos en los que no hallamos elemento activador y de los surgidos de reformulaciones¹¹⁹ nos da, por su parte, un 13,5 % de concordancias. En estos dos supuestos, debido a diversos factores como pueden ser las

¹¹⁷ Nos referimos a casos en los que no existe una fuente identificable en el texto original para la utilización de la perífrasis en la traducción. Puede estar motivada por razones estilísticas, de cohesión, o por una necesidad de explicitar determinadas relaciones sintácticas o semánticas, quizás porque en el original existe una indeterminación aspectual excesiva desde el punto de vista de nuestro idioma. En el texto traducido lo que veríamos es una adición de la perífrasis, mientras que el resto de los elementos de la frase se mantienen tal cual. Por ejemplo: «Mediado el descenso <empecé a> sentirme más ágil y caliente, y mis pensamientos volaron en libertad: *Halfway into the run I was limber and warm, thoughts flying free*» o «<Empecé a> vagar por toda la casa, gimiendo suavemente y retorciéndome las manos: *I wandered through my house, moaning quietly and wringing my hands*».

¹¹⁸ Nos referimos a la paráfrasis de la forma o el sentido de una palabra o expresión por necesidad, es decir, por la falta de un equivalente directo, o por cambios o reformulaciones que realiza el traductor o la traductora, que elige utilizar una expresión de significado diferente para transmitir la intención del original o para alterarla por algún motivo (falta de comprensión, razones estilísticas o adaptación al contexto meta, por ejemplo). Ejemplos: «justo donde <empezaba a> crecer el pelo: *just below the hairline*»; «cuando volvió a apagarse la luz, <empecé a> rociar las huellas borrosas de la pared: *when the lights were out again, I resumed spraying the smeared handprints on the wall*»; «<empezamos a> sacar algo en claro: *Now we're getting somewhere*».

¹¹⁹ En algunos casos podrá observarse que la inclusión en una u otra categoría es discutible, puesto que depende de si, desde nuestro punto de vista, existe algún elemento en el texto original que ha llevado al traductor a utilizar la perífrasis aunque no sea un equivalente estructural directo, o si el traductor la ha empleado sin ningún estímulo, por razones de preferencia personal o de ritmo o estilo. Por eso, muchas veces, ambos porcentajes se tienen en cuenta juntos.

Por otro lado, cuando una concordancia puede incluirse en dos categorías distintas, se prefiere la más particular o específica («*was turning gray*») podría pertenecer a la categoría de «tiempos continuos», pero además incluye un verbo de enlace (*link verb*) que posee un significado concreto que puede haber conducido al traductor a seleccionar la perífrasis en cuestión más que el aspecto estilístico o gramatical del tiempo continuo. Por supuesto, también se trata a veces de una cuestión de naturaleza subjetiva.

preferencias estilísticas del traductor, la explicitación de información, relaciones semánticas o elementos de cohesión que quedaban implícitos en el original o la necesidad de parafrasear o desglosar un concepto, se opta por la utilización de la perífrasis cuando no había en el original un elemento que, en principio, pudiera referirnos a ella directamente.

La estructura más repetida después de las formadas a partir de los verbos *start* y *begin* es la de los «link verbs» —o verbos que sirven de enlace entre el sujeto y el complemento, en este caso, un adjetivo— cuyo significado está relacionado con el cambiar o transformarse.

<Highsmith>¹²⁰: Por ejemplo , con la cantante de ópera noruega no creo que tuviera éxito , pero me acuerdo que ponía la televisión para verla , y decía que ella estaba <empezando a> temblar , que desfallecía .

-->par_en: I don't think he was successful with the Norwegian opera singer , for example , but he kept the TV on , watching her , I remember , and saying she's **becoming** shaky-failing .

Los tiempos continuos, cuando no pueden trasladarse directamente al español sin riesgo de generar estructuras extrañas, también son convertidos con cierta frecuencia en expresiones que incluyen la perífrasis incoativa:

<Grafton> La conversación <empezaba a> reventarme , pero no encontraba la forma de escurrir el bulto .

-->par_en: **I was hating** this conversation , but I could n't find any way to get out of it.

Sucede algo parecido con las oraciones subordinadas que incluyen el participio presente pero no conjunción. El traductor tiende a querer enlazar de algún modo la subordinada con la principal y, en este caso, se vale de la perífrasis:

<Cornwell> Otras voces se le unieron y hubo manos que aferraron las mallas de la cerca y <empezaron a> sacudirlas .

-->par_en: Other voices joined in and hands grabbed the chain-link fence , **shaking it** .

Sólo nos queda un 8 % de concordancias, que se distribuyen entre este último caso y los casos provocados, por ejemplo, por distintos *phrasal verbs* y verbos seguidos de preposición que pueden adoptar significado incoativo, además de las frases que

¹²⁰ No se proporcionan números de página para los ejemplos incluidos a lo largo del texto, dado que están copiados de las concordancias extraídas mediante los distintos programas de análisis, no de las novelas enteras que se digitalizaron para el corpus.

contienen un marcador temporal que, a nuestro juicio, activa la utilización de la perífrasis:

<Elroy>: Western Avenue : terreno de mirones , moteles de putas , ninguna dirección con la que <empezar a> trabajar .

-->par_en: Western Avenue-peeper turf-whore motels , no addresses **to work off** .

<Grafton>: <Empecé a> distinguir delante un suave resplandor rojizo que perfilaba la pared donde el pasillo torcía a la izquierda .

-->par_en: Ahead , I could **now** discern a soft reddish glow defining the wall where the corridor curved left

También encontramos entre estos últimos casos una categoría que podría considerarse un tipo específico de reformulación, la del verbo *feel* seguido de infinitivo sin «to»:

<Grafton>: <Empecé a> ponerme nerviosa y me pregunté si esperaba que le abonase el trayecto .

-->par_en: **I felt anxiety stir** , wondering if he expected payment for his services .

El 13,5 % de concordancias en las que no hay un elemento activador concreto y la gran variedad de estructuras que pueden dar pie a su utilización nos indican que esta perífrasis es un recurso habitual y accesible para los traductores, parte del inventario normal de estructuras a las que recurren para solucionar distintos problemas de traducción. Quizás de ahí surja esa pequeña diferencia en la cantidad de apariciones a favor del corpus de traducciones (que no llega de todas formas a ser significativa estadísticamente, como hemos mencionado).

Elemento activador EN	Frecuencia	Porcentaje
Begin	18	66,7
Verbo + preposición (burst into, head up, move down)	3	11,1
Start	2	7,4
Tiempos continuos	2	7,4
Link verbs (get/become/turn/go + ADJ)	1	3,7
Verbo <i>feel</i> seguido de infinitivo sin to	1	3,7
TOTAL	27	100

Tabla 9 - Equivalentes de «comenzar a + INF» en EN

En cuanto a «comenzar + infinitivo», *start*, junto con *burst into* y demás estructuras incluidas en la tabla, parecen elementos activadores de carácter marginal. El verbo que más encontramos traducido por «comenzar» es *begin* con un 66,7 % y los demás elementos ya los encontrábamos entre los equivalentes de «empezar». Podemos concluir que la elección de esta perífrasis en lugar de «empezar a + infinitivo» no es una

cuestión de influencia del texto original sino que depende de factores subjetivos relacionados con los procesos mentales del traductor.

4.4.1.2. Volver a + infinitivo

Es una perífrasis aspectual reiterativa —también denominada iterativa—, es decir, que expresa la repetición de una acción («Ha vuelto a beber»); la única del español, lo que la convierte en la traducción más habitual para las expresiones de significado similar en inglés, a pesar de que nuestra lengua cuenta también con adverbios y partículas que expresan repetición, como *de nuevo*, *otra vez* o el prefijo *re-* (*rellenar*).

En inglés, lo habitual es añadir al verbo adverbios como *again*, *once more*, *since*, *anew*, *over* o *back*, y también el prefijo *re-* (como en *refill*) para expresar el mismo significado. A veces es necesario recurrir al contexto para saber si nos encontramos o no ante una perífrasis. «Volvió a buscar su cartera» puede significar tanto «Buscó otra vez su cartera» como «Regresó allí a buscar su cartera». Las construcciones de este tipo son perifrásticas cuando, por el contexto, entre el auxiliar y el infinitivo no es posible introducir un complemento circunstancial locativo (Gómez Torrego 1999: 3376).

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: VOLV*/VUELV*/VUÉLV*/VUELTO

Context words: A

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Volver a + INF	190	58,6	319	61,3	-0,25	n. s.	NO

No hubo sorpresas con respecto a esta perífrasis, ya que esperábamos que no verificara la hipótesis por los motivos que ya hemos alegado en la descripción inicial. Teniendo en cuenta el elevado número de apariciones de esta perífrasis en el corpus, los resultados indican que «volver a + infinitivo» se utiliza con una frecuencia bastante similar en originales y traducciones (se aprecia una ligera sobrerrepresentación en

traducciones pero no estadísticamente significativa según nuestros valores umbral establecidos). Esto se debe probablemente a que es la única perífrasis reiterativa del español y su uso es muy habitual, por lo que, a pesar de ser formalmente muy diferente a sus posibles equivalentes en inglés, es la traducción estándar para las partículas que expresan este significado.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?volv.*| (¿|i)?vu(e|é)lv.*| (¿|i)?vuelto" %c] [0,5] "a"]

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Volver a + INF	199	54,2	378	77,4	-17,04	< 0,0001	NO

Dada nuestra idea de que es posible que los traductores españoles utilicen perífrasis verbales en menor medida que los argentinos, sorprende que en este caso hayamos encontrado todo lo contrario. La cantidad de casos aumenta seis veces más en el subcorpus de traducciones que en el de originales, lo que hace que, en esta ocasión, la perífrasis resulte claramente sobrerrepresentada en traducciones (esta vez, de manera estadísticamente muy significativa), en detrimento quizás de las alternativas de traducción mencionadas en la descripción de la perífrasis.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	62	73	Leonard	103	132,4
Martínez	53	97,5	Highsmith	80	81,4
Vázquez	32	36,4	Cornwell	80	72,8
Madrid	26	62,7	Grafton	79	85,1
Martín	16	28,5	Ellroy	36	32,8
Núñez	10	23,5			
Total	199	54,2		378	77,4

Tabla 10 - Distribución de «volver a + INF» por texto

La utilización de esta perífrasis en las traducciones es asombrosamente pareja en cuanto a distribución porcentual (tres autores con cerca del 21 % del total y uno con 27 %), y quizás esta regularidad pueda deberse a factores relacionados con los textos originales o más bien con la lengua de partida. Sin embargo, hay que tener en cuenta que las novelas estadounidenses tienen un número de palabras bastante semejante (en torno a 100 000 excepto Leonard, que está en casi 78 000, y aun así es quien mayor

porcentaje de «volver a + infinitivo» presenta), mientras que las españolas fluctúan desde las casi 88 000 de Vázquez Montalbán a las 41 400 de Madrid.

Solo se aparta de la media del 22 % Ellroy, que es el que la utiliza menos cantidad de veces (9,5 % del total), a pesar de ser su novela la segunda en extensión. Al respecto y, dado que esta contradicción se dará en numerosas ocasiones a lo largo de este capítulo, conviene apuntar que este autor utiliza técnicas bastante distintas a las de los otros autores, sobre todo en lo referente a los diálogos (Martínez Vilinsky, 2014: 11). En muchas ocasiones, este autor elimina los *verba dicendi*, dando lugar a acotaciones del narrador que parecen extraídas de guiones teatrales más que de novelas policíacas:

<Ellroy> Up yawning . " Results ? "

-->par_trad: Yo, con un bostezo: —¿Resultados?

De lo que puede deducirse de las entrevistas que ha concedido (Rich 2009), el peculiar estilo telegráfico de Ellroy empezó a desarrollarse cuando su editor le pidió que redujera la extensión de su novela *L. A. Confidential*. Él decidió que una manera de hacerlo sin eliminar ninguna parte fundamental de la obra era reducir la cantidad de palabras «innecesarias» de las frases. Esto dota a sus diálogos y personajes de un estilo directo, ágil y espontáneo, pero reduce la cantidad de verbos y perífrasis verbales que podemos encontrar en su texto.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Again	204	54
Back	54	14,3
∅	37	9,8
Prefijo re-	34	9
Marcadores temporales	18	4,8
Reformulación	12	3,2
Determinantes next, another y more	12	3,2
Over	7	1,8
TOTAL	378	100

Tabla 11 - Equivalentes de «volver a + INF» en EN

El principal activador de esta perífrasis es el adverbio de frecuencia *again* con un 54 % de concordancias, aunque no todos los *again*, como es lógico, se traducen por la perífrasis, puesto que las concordancias paralelas encontradas que incluyen este

equivalente solo representan un 41,5 % de las apariciones de este adverbio (491) en el corpus de originales ingleses (EN).

<Cornwell>: La ventanilla <volvió a> cerrarse .
 -->par_en: The window shut again .

Por orden de frecuencia, lo sigue el adverbio *back*, que expresa repetición o regreso a una condición anterior, con un 14,3 % de concordancias:

<Cornwell>: Parece ser que se casó en 1961 , se divorció dos años después y <volvió a> adoptar el apellido Deighton .
 -->par_en: Appears she was married in 1961 , got divorced two years later , and changed her name back to Deighton .

Con menos de un 10 % encontramos otros equivalentes habituales de esta perífrasis, como:

- el prefijo *re-* (y aquí también hemos decidido incluir palabras como *resumed*¹²¹ o *returned* que implican una repetición de la acción, aunque el *re-* que presentan no pueda ya considerarse prefijo en inglés moderno),

<Grafton>: Luego se <volvió a> casar y pareció olvidarse de Ray .
 -->par_en: Then she remarried and seemed to forget about Ray .

- sintagmas con los determinantes *next* (2), *another* (7) y *more* (3).

<Highsmith>: -¿Podéis <volver a> intentar averiguar de dónde ha sacado el nombre de Cynthia el tal David Pritchard ?.
 -->par_en: Can you and Ed make another effort to find out how this David Pritchard got hold of Cynthia's name ?

- marcadores temporales de duración y frecuencia como *now* (1), *next* (1), *since* (8), *still* (3), *ever* (2), *never* (2) y *sometimes* (1)

<Ellroy>: No sé cómo se llama , no la había visto nunca y no he <vuelto a> verla .
 -->par_en: I do n't know her name-I never saw her before and I have n't seen her since .

¹²¹ «Resume is from Latin *resumere* "to take up again, take back," from the prefix *re-* "again" plus *sumere* "to take up, take."» Fuente: *Vocabulary.com* <https://www.vocabulary.com/dictionary/resume> [fecha de consulta: 22/01/2014].

- y, finalmente, el adverbio *over*, con 7 casos.

<Leonard>: De modo que <volvió a> empezar .

-->par_en: So he started **over**

En el 3,2 % correspondiente a la categoría de reformulación hemos agrupado los casos en los que se utiliza la perífrasis para traducir una amplia variedad de partículas y estructuras, como pueden ser «*spin me fresh*», «*slid the padlock home*», «*unclench his fist*» o «*fix you up* (rellenar una bebida)».

Por último, observamos que casi un 10 % de los casos es fruto de una adición del traductor; es decir, no hay ningún estímulo lingüístico en el texto que lleve a pensar en la utilización de la perífrasis reiterativa.

<Cornwell>: La noticia determinaba que yo fuera a acostarme o <volviera a> meterme en el coche para regresar a la morgue .

-->par_en: The news determined whether I went on bed **or drove downtown to the morgue** .

Como vemos, no solo hay una amplísima variedad de estructuras que activan la aparición de la perífrasis, sino que, una vez más, nos encontramos con más del 9 % de las concordancias en las que el texto original no parece ser responsable de su utilización. Como adelantábamos en su descripción, «volver a + infinitivo» es el recurso más habitual para expresar el significado de repetición en español; tanto que se la utiliza también en ocasiones en las que en principio no haría falta, con fines estilísticos o cohesivos, por razones compensatorias o para explicitar cierta información que el texto da por sentado. Por ejemplo, en la siguiente concordancia, quizás podríamos averiguar por el contexto que el protagonista ha fruncido el ceño ya una vez o más de una vez antes del momento descrito. Sin embargo, si en la frase original no se menciona, cabría plantearse si realmente es imprescindible remarcar esa información en la traducción¹²², o si el hecho de que los traductores tiendan a hacerlo podría responder al fenómeno de explicitación descrito en el apartado sobre universales de traducción.

¹²² Hay que mencionar que algunos de estos casos pueden deberse a un uso superfluo de la perífrasis en español en ciertas ocasiones en las que lo que se pretende describir no es la repetición de una acción sino el retorno de un elemento a su estado inicial. Por ejemplo, se podría decir a alguien «vuelve a cerrar la puerta», aun cuando nuestro interlocutor solo la hubiese abierto esa única vez al entrar.

<Cornwell>: Wesley <volvió a> fruncir el ceño .

-->par_en: Wesley frowned .

Con respecto a la idea que aventurábamos sobre la posible menor utilización de alternativas de traducción de este significado en español debido al sobreuso de la perífrasis, vemos que, en principio, no se confirma, pues, al realizar una búsqueda de algunas de las expresiones sinónimas más habituales en ambos subcorpus, no encontramos prácticamente diferencia entre ellos. Por tanto, la frecuencia de «acciones reiteradas» sigue siendo mayor en TRAD debido exclusivamente al sobreuso de la perífrasis: si sumamos los casos de perífrasis y expresiones de la tabla de abajo (ES: 414; TRAD: 689), el resultado es LL: -13,14; $p < 0,001$.

- De nuevo:	ES 108	TRAD 144	LL -0,00	(n.s.)
- Otra vez:	ES 82	TRAD 140	LL -3,27	(n.s.)
- Una vez más:	ES 11	TRAD 15	LL: -0,0	(n.s.)
- Nunca más:	ES 14	TRAD 12	LL: + 1,26	(n.s.)

4.4.1.3. Seguir + gerundio - Continuar + gerundio

Estas dos perífrasis poseen valor durativo continuativo y expresan una acción que proviene de un estado anterior y se prolonga, por lo general, más allá de un momento en que se suponía, deseaba, temía o esperaba que no continuara desarrollándose. *Continuar* es bastante menos frecuente en la lengua actual que *seguir* y aporta un grado mayor de formalidad. El significado de esta perífrasis se expresa en inglés con las fórmulas «*keep on/go on/carry on + -ing*» (además de otros verbos como *drive* que, seguidos por la partícula «*on*» adquieren ese valor de continuidad) o con expresiones que incluyen *continue* o *still*: «Sigo pensando que no era así», *I still think it wasn't that way*. *Still* también puede traducirse por *todavía* o *aún*.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: *SEGUI*/SIGU*/SIGO/SEGUÍ*/SIGA*/SIGÁ*
CONTINU*/CONTINÚ**

Context words: **ANDO*/*IENDO*/*YENDO*/*ÁNDO*/*IÉNDO*/*YÉNDO**

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Seguir + GER	150	46,2	192	36,9	+4,24	< 0,05	SÍ
Continuar + GER	25	7,7	14	2,7	+10,50	< 0,01	SÍ

«Seguir + gerundio», como esperábamos, se encuentra bien representada en el subcorpus de traducciones, debido a su equivalencia bastante evidente con estructuras como «*keep/continue/go on + -ing*». Hay una diferencia significativa, sin embargo, a favor del subcorpus de originales. «Continuar + gerundio», a pesar de pertenecer al grupo con equivalentes fácilmente identificables, nos sorprende también al confirmar la hipótesis de la infrarrepresentación en el subcorpus de traducciones. Es bastante menos frecuente que «seguir + gerundio» en ambos subcorpus pero aparece casi tres veces más por cada cien mil palabras en los textos originales que en las traducciones.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?segu(i|i).*(¿|i)?sig(u|o|a|á).*" %c] []{0,5} [word=".*(a|á)ndo.*|(i|y)(e|é)ndo.*"];
[word="(¿|i)?contin(u|ú).*" %c] []{0,5} [word=".*(a|á)ndo.*|(i|y)(e|é)ndo.*"];

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Seguir + GER	170	46,3	183	37,4	+3,95	< 0,05	SÍ
Continuar + GER	24	6,5	14	2,9	+6,30	< 0,05	SÍ

En la segunda versión del corpus, los resultados se mantienen. Ambas perífrasis se encuentran infrarrepresentadas en traducciones.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	53	62,4	Cornwell	57	51,9
Martínez	34	62,6	Leonard	38	48,8
Vázquez	30	34,2	Highsmith	36	36,6
Núñez	24	56,5	Elroy	32	29,1
Martín	17	30,2	Grafton	20	21,5
Madrid	12	29			
Total	170	46,3		183	37,4

Tabla 12 - Distribución de «seguir + GER» por texto

Con esta perífrasis, así como con todas las de gerundio, hemos tenido alguna duda a la hora de decidir si considerar ciertos casos como verdadera perífrasis o no. La intercalación de elementos entre auxiliar y auxiliado (complementos, adverbios o incluso el sujeto, en casos como «Están los pájaros cantando») es habitual en este tipo de perífrasis, lo que a veces genera confusiones. La diferencia entre subcorpus pasaría a ser algo más significativa si tuviéramos en cuenta las concordancias siguientes, muy discutibles en cuanto a su naturaleza perifrástica. En TRAD encontramos:

<Leonard>: Max se inclinó hacia atrás para ver a Renee - que <segua allí picoteando>

-->par_en: Max leaned back to check on Renee - **still there**, nibbling -

<Leonard>: Ella <segua en una mesa tomando> alguna mierda griega con ese pan que llama « pita »

-->par_en: She was **still at the table** having some kind of Greek shit in that pita bread .

En ES:

<Vázquez>: - El jurado debe de <seguir reunido esperando> el veredicto .

<Giménez>: Mi compañero <segua imperturbable poniendo> cara de recogimiento conventual .

<Giménez> <Seguían fuera , intercambiando> consejos agrícolas y toqueteando las plantas que eran el emblema de mi nuevo hogar .

De tenerlas en cuenta, pasaríamos a contar con 173 casos en ES y 185 en TRAD, y el test de log-likelihood nos indicaría un valor de +4,26, significativo al nivel de $p < 0,05$, por lo que los resultados apenas variarían. Sin embargo, no creemos que en los dos primeros casos estemos ante verdaderas perífrasis, al menos si nos basamos en lo que Gómez Torrego (1988:136-138) explica acerca de la perífrasis «estar + gerundio». En una frase como «Estoy en Madrid estudiando medicina», el verbo nuclear es de carácter locativo, y se pretende focalizar tanto «estar en Madrid» como «estudiar medicina». Si la frase pasara a ser «Estoy estudiando medicina en Madrid», aunque el ejemplo podría parecerse idéntico, no lo es. Esta vez el sintagma preposicional afecta a todo el conjunto, y la focalización está en lo que se estudia (medicina en Madrid), por lo que el verbo es auxiliar y estaríamos ante una perífrasis. Con respecto a los tres casos de ES, creemos, como explica Gómez Torrego con otro ejemplo similar («Estamos en Madrid pasándolo muy bien») que solo podríamos considerarlas perífrasis si entendiéramos que existe omisión del segundo verbo «estar» («estamos en Madrid» y «estamos pasándolo muy bien») o «seguir», en nuestros ejemplos.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Madrid	9	21,7	Highsmith	6	6,1
Martín	8	14,2	Ellroy	5	4,5
Giménez	5	5,9	Cornwell	2	1,8
Vázquez	1	1,1	Leonard	1	1,3
Martínez	1	1,8	Grafton	0	0
Núñez	0	0			
Total	24	6,5		14	2,9

Tabla 13 - Distribución de «continuar + GER» por texto

La diferencia también es significativa con respecto a «continuar + gerundio», aunque con un nivel de significación un poco menor que en la configuración anterior del corpus. Lo más probable es que suceda lo mismo que con el par de perífrasis «empezar/comenzar + infinitivo»: los traductores tienden a utilizar la primera opción que viene a la cabeza, o la opción más comúnmente utilizada —estándar— para expresar el significado específico de estas perífrasis que, como comprobamos en el subcorpus de originales españoles, es «empezar» (o «seguir»), mientras que «comenzar» (o «continuar») se reserva para casos derivados de contextos más formales o usos lingüísticos más esmerados. Sin embargo, en este caso, ambas opciones están infrarrepresentadas en los textos traducidos, con lo que podríamos pensar que en los textos originales sucede algo parecido.

Por otra parte, las diferencias podrían deberse simplemente a cuestiones de estilo. Por ejemplo, Núñez presenta una frecuencia de uso algo mayor de la perífrasis formada con *seguir* que la que le correspondería por número de palabras total de su novela, mientras que no usa el auxiliar *continuar* ni una sola vez. Madrid, por su parte, es quien menos utiliza *seguir* y quien más utiliza *continuar*.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Still	49	26,8
Keep (on) + -ing	35	19,1
∅	32	17,5
Continue + -ing/infinitivo	28	15,3
Reformulación	13	7,1
Adverbios comparativos (more, longer, further)	9	4,9
Keep y continue + prep, adv, SN o nada	7	3,8
Link verb (stay, remain) + adj o SN	5	2,7
Verbo + partícula <i>on</i>	5	2,7
TOTAL	183	100

Tabla 14 - Equivalentes de «seguir + GER» en EN

Observamos que la utilización de la perífrasis «seguir + gerundio» procede en un 34,4 % de los casos (63) de la traducción de los verbos *keep* y *continue* a partes prácticamente iguales: *keep* + -ing (31 casos), *keep on* + -ing (4 casos), *continue* + -ing (12 casos) y *continue* + infinitivo (16 casos).

<Cornwell>: Lucy <seguía controlando> nuestro ordenador desde el terminal de mi casa para ver si alguien intentaba acceder de nuevo a mi directorio .

-->par_en: Lucy continued to monitor our computer from my terminal at home to see if anyone attempted to access my directory again .

Si a estos añadimos los casos en los que estos dos verbos son seguidos por una preposición (*continue down*) o un grupo nominal (*keep the same rate*), forman parte de una *phrasal verb* (*keep up*) o, en el caso de *continue*, aparecen solos con su significado original (*the current boom continues*) obtenemos un 38,2 % de concordancias que incluyen estos dos elementos.

<Grafton>: <Seguí andando> « normalmente » al doblar la esquina .

-->par_en: I kept my pace at " normal " as I turned the corner .

El adverbio de duración *still* es el elemento desencadenante de su utilización en un 26,8% de los casos y expresiones con adverbios comparativos como *further*, *longer* o *more*, que modifican al verbo y expresan prolongación de la actividad en el tiempo, son responsables de un 4,9 %.

<Cornwell>: <Seguía llevando> pajarita y chaleco y fumaba en pipa , y cuando me miró , sus ojos grises eran tan aptos para la disección como cualquier escalpelo .

-->par_en: He still wore bow ties and vests and smoked a pipe , and when he looked at me , his gray eyes were as capable of dissection as any scalpel .

<Leonard>: Se acabó la bebida y <siguió esperando> .

-->par_en: Finished the drink and waited some more .

Incluimos en esta categoría el único caso en el que *more* no funciona como adverbio sino como adjetivo, por formar parte de una frase hecha («*give some thought to something*»):

<Leonard> Él asintió , y <siguió pensando> .

-->par_en: He nodded , giving it some more thought .

Las últimas dos categorías, verbos de enlace seguidos de adjetivo o sintagma nominal y verbos seguidos de la partícula adverbial «on» para expresar progreso o avance, cuentan con un 2,7 % cada una.

El 24,6 % restante, es decir, casi un cuarto de las concordancias entre adiciones y reformulaciones, corresponde a una utilización, podríamos decir, libre o espontánea por parte del traductor, lo que parece indicar, una vez más, que nos encontramos ante una perífrasis que constituye un recurso habitual del español y no necesita forzosamente un elemento desencadenante en la lengua original del texto (aunque sí es cierto que encontramos una presencia importante de tiempos continuos o participios presentes). Como acabamos de comprobar, no obstante, ni esta condición ni el número de elementos detonadores en inglés son suficientes para asegurar una representación de la perífrasis en traducciones similar a la que puede encontrarse en nuestro subcorpus de textos escritos originalmente en español.

Elemento activador EN	Frecuencia	Porcentaje
Continue + -ing/infinitive	4	28,6
Keep (on) + -ing	3	21,4
∅	3	21,4
Continue + prep, adv, SN o nada	2	14,3
Reformulación	1	7,1
Go on + infinitivo	1	7,1
TOTAL	14	100

Tabla 15 - Equivalentes de «continuar + GER» en EN

El elemento activador más habitual de la perífrasis «continuar + gerundio» es, previsiblemente, la palabra *continue*, que aparece, ya sea sola o seguida de gerundio, en seis de las catorce concordancias y, junto con *keep* (3) y *go on* (1), conforma el 71,4 % de las concordancias. Sin embargo, el 28,5 % de los casos corresponde a una única reformulación y tres adiciones, un porcentaje algo más alto que el de «seguir + gerundio».

<Ellroy>: Yo <continuaré protestando> de esta burda maniobra como ciudadano privado y.. . »

-->par_en: I will continue to protest this travesty as a private citizen . I.. . "

Salvo en este caso que hemos transcrito, en las traducciones no parece reservarse a contextos más formales o lenguaje más cuidado, sino más bien ser activada por su equivalente literal *continue* o utilizada según el contexto a discreción del traductor:

<Leonard>: Jackie asintió y Ramona <continuó hablando> - . Me estaba preguntando si te pagarían bien .
 -->par_en: Jackie nodded and Ramona said , " What I was wondering , they pay pretty good ?"

4.4.1.4. Acabar por + infinitivo - Terminar por + infinitivo

Estas dos perífrasis de valor aspectual perfectivo que equivalen a «acabar/terminar + gerundio» indican la culminación de un proceso o acción que muchas veces acarrea connotaciones negativas de impaciencia o reprobación para el sujeto («Mi madre acabó por hartarse»). Es posible la intercalación del adverbio *no* entre la preposición y el infinitivo para formar un enunciado negativo: «Siempre termina por no hacer nada». Hay que tener en cuenta que no se utilizan para expresar la finalización de un proceso, pues «terminó comiendo» o «acabó por comer» no significan lo mismo que «terminó de comer», por lo que las expresiones que encontraremos en inglés en cada caso serán distintas («*He ended up eating*» y «*He finished eating*»). Para expresar el significado que nos ocupa, en inglés se suele utilizar las expresiones *end up*, *end by* o *in the end* o *finally*, que pueden traducirse al español también por fórmulas más literales como *finalmente*, *al final* o *a la larga*.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: ACAB*|ACÁB* Y TERMIN*|TERMÍN*

Context words: POR

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Acabar por + INF	5	1,5	2	0,4	+3,13	n. s. ¹²³	NO
Terminar por + INF	7	22	0	0	+13,40	< 0,001 ¹²⁴	SÍ

Los ejemplos son escasos, pero podemos comprobar que mientras que «acabar por + infinitivo» se utiliza más que «terminar por + infinitivo» en el subcorpus de traducciones, en el de originales sucede lo contrario.

¹²³ Test exacto de Fisher: 0,115252 (n. s.)

¹²⁴ Test exacto de Fisher: 0,001232 (p < 0,01)

Estas dos perífrasis pertenecen al grupo I, es decir, de las que en principio cuentan con equivalentes evidentes o inmediatos en inglés. Por este motivo, pensábamos que estarían representadas en el subcorpus de traducciones de manera similar a la de los originales y nos sorprendió comprobar que no sucedía así en el caso de «terminar por + infinitivo», que resultó infrarrepresentada. De hecho, más tarde comprobamos que, en esta primera fase del corpus, de las cuatro perífrasis de significado equivalente que mencionamos en la descripción inicial, la única que no resulta infrarrepresentada en traducciones es «acabar por + infinitivo». De todas formas, hay que tener en cuenta que, de hecho, son más los casos de «acabar por + infinitivo» en originales españoles que los hallados en traducciones, y la perífrasis estuvo cerca de pasar el test de significación estadística.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿ | ¡)?ac(a | á)b.*" %c] []{0,5} "por";
 [word="(¿ | ¡)?term(i | í)n.*" %c] []{0,5} "por";

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Acabar por + INF	7	1,9	5	1	+1,15	n. s.	NO
Terminar por + INF	7	1,9	0	0	+11,85	< 0,001 ¹²⁵	SÍ

No hay cambios importantes en las cifras obtenidas en referencia a estas perífrasis, pues la mayor presencia de «acabar por + infinitivo» en originales (presente en cuatro de los seis autores del subcorpus ES y solo dos de cinco del subcorpus TRAD) sigue sin ser significativa en términos estadísticos, mientras que sí lo es en el caso de «terminar por + infinitivo», que, a pesar de estar presente solo en dos de los seis textos españoles, no aparece en ninguna de las traducciones.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	3	3,5	Grafton	4	4,3
Núñez	2	4,7	Ellroy	1	0,9
Martínez	1	1,8	Cornwell	0	0
Martín	1	1,8	Highsmith	0	0
Madrid	0	0	Leonard	0	0
Vázquez	0	0			
Total	7	1,9		5	1

Tabla 16 - Distribución de «acabar por + INF» por texto

¹²⁵ Test exacto de Fisher: 0,002675 (p < 0,01)

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	5	8,9	Cornwell	0	0
Madrid	2	4,8	Ellroy	0	0
Giménez	0	0	Grafton	0	0
Martínez	0	0	Highsmith	0	0
Núñez	0	0	Leonard	0	0
Vázquez	0	0			
Total	7	1,9		0	0

Tabla 17 - Distribución de «terminar por + INF» por texto

En el corpus de originales españoles, el uso de una u otra perífrasis parece casi mutuamente excluyente, puesto que los tres autores que utilizan el auxiliar *acabar* no se valen de *terminar* para expresar el mismo significado en ninguna ocasión y lo mismo sucede pero a la inversa con los dos autores que utilizan esta última, salvo en el caso de Andreu Martín, que utiliza una vez «acabar por» (contra cinco veces en las que se decanta por «terminar por»). Vázquez Montalbán no se vale de estas expresiones en su texto.

«Terminar por + infinitivo» es una perífrasis que debería estudiarse con mayor profundidad en otros corpus de mayor tamaño o distintos parámetros de compilación para determinar su idoneidad como elemento único del español, pues no hemos encontrado ni un caso de esta expresión en nuestro subcorpus de traducciones.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
∅	3	60
Finally	2	40
TOTAL	5	100

Tabla 18 - Equivalentes de «acabar por + INF» en EN

En dos ocasiones el detonante es explícito, el adverbio *finally*:

<Grafton> ¿Sabes lo que he <acabado por> aprender ?

-->par_en: " You know what I 've **finally** learned ?

Sin embargo, en los otros tres casos, la utilización de la perífrasis parece cuestión de preferencia estilística del traductor en cuestión. Dos de los casos están en futuro y uno en pasado:

<Ellroy>: <Acabaréis por> decírmelo , por decírselo a alguien : esta vez cuento con un respaldo muy sólido .

-->par_en: You **II** tell me , you **II** tell somebody-I 'm partnered up strong on my ride down .

<Grafton>: Desde el fallecimiento de mi tía , ocurrido hace unos diez años , he <acabado por> firmar un acuerdo con mi condición solitaria .

-->par_en: **Since** my aunt's death some ten years ago , **I 'd made** my peace with my solitary state

<Grafton>: <Acabarían por> despedirla si seguía hablando mal de Asociados .

-->par_en: She **was going to** get herself fired if she continued to mouth off at Corporate .

Al parecer, «acabar por + infinitivo» se encuentra en el repertorio de recursos de algunos traductores (dos, en este caso) que lo utilizan sin necesidad de verse empujados a ello por estructuras del inglés (lo cierto es que lo hacen en escasas ocasiones, pero tampoco parece ser excesivamente frecuente en originales españoles). Sin embargo, no sucede así con «terminar por + infinitivo», que no es tenida en cuenta en ningún caso.

Con respecto a las otras expresiones que podrían activar la utilización de la perífrasis en traducciones, encontramos 25 apariciones de *end up* en TRAD y solo una de *end by*, que se traducen al español en su mayoría por *terminar* o *acabar* a secas o como parte de la perífrasis correspondiente de gerundio, como veremos más adelante, o simplemente se dejan sin traducir. *In the end* (5 casos) se traduce por las siguientes expresiones: *a la larga*, *a fin de cuentas*, *al final*, *en último término* y *en los últimos momentos*. Además, vemos que existen 73 casos más de utilización del adverbio *finally* que no se traducen por ninguna perífrasis sino por fórmulas como *por fin*, *al fin*, *finalmente* o *principalmente*. Las perífrasis que nos ocupan no parecen ser un recurso frecuente para expresar el valor perfectivo culminativo en español.

4.4.1.5. Acabar + gerundio - Terminar + gerundio

Estamos ante dos construcciones de dudosa interpretación perifrástica, según algunos autores (p. ej. Yllera 1999: 3423), que a veces se denominan semiperifrásticas (Gómez Torrego 1988: 171) y cuyos valores y equivalentes son idénticos a los de las perífrasis de infinitivo del apartado anterior: «acabar/terminar por + infinitivo». Como hemos comentado en la descripción de estas últimas, presentan la acción como culminación de un proceso con valor puntual pero no la terminación de la acción, y muchas veces transmiten matices de impaciencia o acción reprobada o inesperada.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: ACAB*/TERMIN*/ACÁB*/TERMÍN*

Context words: *ANDO*/*IENDO*/*ÁNDO*/*IÉNDO*/*YENDO*/*YÉNDO*

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Acabar + GER	33	10,2	10	1,9	+26,22	< 0,0001	SÍ
Terminar + GER	8	2,5	2	0,4	+7,24	< 0,01 ¹²⁶	SÍ

Son, estas, dos perífrasis que pensábamos que estarían bien representadas en el corpus de traducciones y, sin embargo, resultaron infrarrepresentadas, al igual que «terminar por + infinitivo». En ambos subcorpus, «acabar + gerundio» es bastante más frecuente que «terminar + gerundio».

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?ac(a|á)b.*" %c] []{0,5} [word=".*(a|á)ndo.*|(i|y)(e|é)ndo.*"];
 [word="(¿|i)?term(i|f)n.*" %c] []{0,5} [word=".*(a|á)ndo.*|(i|y)(e|é)ndo.*"];

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Acabar + GER	38	10,3	22	4,5	+10,11	< 0,01	SÍ
Terminar + GER	9	2,4	4	0,8	+3,67	n. s.	NO

Acabar es, también en esta segunda fase, el auxiliar más utilizado de los dos. Hay una diferencia importante con respecto a los resultados anteriores: mientras que «acabar + gerundio» ha resultado, una vez más, infrarrepresentada en el corpus de traducciones, esta vez no ha sido así en el caso de «terminar + gerundio» (auxiliar que, por cierto, esta vez sí está representado en el subcorpus de traducciones, a diferencia de lo que sucedía en el caso de la perífrasis «terminar por + infinitivo»). Su resultado en el test de log-likelihood no llega a ser estadísticamente significativo, a pesar de haberse quedado casi en el umbral de significación, establecido en 3,84, debido al aumento de casos en el subcorpus de originales y probablemente también a la mayor equiparación del número de palabras entre ambos subcorpus.

¹²⁶ Test exacto de Fisher: 0,017128 (p < 0,05)

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martínez	15	27,6	Cornwell	8	7,3
Giménez	10	11,8	Highsmith	5	5,1
Núñez	5	11,8	Grafton	5	5,4
Vázquez	5	5,7	Leonard	2	2,6
Martín	3	5,3	Ellroy	2	1,8
Madrid	0	0			
Total	38	10,3		22	4,5

Tabla 19 - Distribución de «acabar + GER» por texto

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	7	12,4	Ellroy	3	2,7
Vázquez	1	1,1	Grafton	1	1,1
Núñez	1	2,3	Cornwell	0	0
Giménez	0	0	Highsmith	0	0
Madrid	0	0	Leonard	0	0
Martínez	0	0			
Total	9	2,4		4	0,8

Tabla 20 - Distribución de «terminar + GER» por texto

En el subcorpus ES encontramos una curiosa relación inversa entre los autores que utilizan uno u otro auxiliar: los tres autores que menos usan «acabar + gerundio» (Núñez, Vázquez Montalbán y Martín, en orden de mayor a menor uso), son los únicos tres que usan «terminar + gerundio» (en orden de uso inverso), aunque Vázquez Montalbán y Núñez solo la utilizan una vez cada uno. Juan Madrid no parece utilizar ninguna de las dos perífrasis, así como tampoco se vale de «acabar por + infinitivo». Sin embargo, es uno de los únicos dos autores que utilizan «terminar por + infinitivo».

En TRAD, la elección de uno u otro auxiliar también podría pensarse un aspecto idiosincrático del traductor. Sabaté Vargas (texto de Ellroy), por ejemplo, es el que menos utiliza «acabar» y el que más utiliza «terminar» (de solo dos traductores que lo hacen), aunque solo sea en dos y tres ocasiones respectivamente.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
∅	11	50
End up / by	5	22,7
Reformulación	3	13,6
Marcadores temporales (finally, ever, eventually)	3	13,6
TOTAL	22	100

Tabla 21 - Equivalentes de «acabar + GER» en EN

Elemento activador EN	Frecuencia	Porcentaje
End up /end in	2	50
Reformulación	1	25
∅	1	25
TOTAL	4	100

Tabla 22 - Equivalentes de «terminar + GER» en EN

Como vimos que sucedía con «acabar por + infinitivo», más de la mitad de las concordancias de estas dos perífrasis de gerundio no corresponden a un elemento específico en inglés que pueda decirse que activa la utilización de la perífrasis de manera directa. Probablemente este es el motivo por el cual la perífrasis «acabar + gerundio» está infrarrepresentada en TRAD. Los traductores la utilizan poco, ya que no encuentran estímulo en el texto original, y cuando lo hacen, se sirven de ella para trasladar estructuras muy diversas; de ahí que más del 60 % de los equivalentes hallados en las concordancias lo ocupen adiciones y reformulaciones. En el caso de «terminar + gerundio», la mitad corresponde a una reformulación y una adición.

<Highsmith>: Tom siempre se prometía que no iba a hacer ninguno más , y luego <acababa aceptando> .

-->par_en: Tom often vowed he would not do another , then **found himself** accepting .

<Ellroy>: Es una lástima que la película <termine pasándose> en los autocines de pueblo .

-->par_en: Too bad the movie **'s headed for** drive-ins in Dogdick , Arkansas .

La otra mitad de las concordancias de «terminar + gerundio» son traducción de una expresión con *end*, y algo parecido sucede en el caso de «acabar + gerundio», en el que la estructura más representada entre las 8 concordancias que no corresponden ni a adiciones ni a reformulaciones es la de «*end up/by* + -ing/sustantivo», con un 22,7 %.

<Highsmith>: Intentó sentarse en una caja de herramientas , pero quedaba demasiado baja , y <acabó apoyándose> contra el guardabarros delantero con la cabeza gacha .

-->par_en: Tom tried to sit on a tool box , which was too low , and **ended by** leaning against a front fender with his head bent low .

<Ellroy>: Le disparó y <terminó usando> la navaja del propio tipo

-->par_en: She shot him , and she **ended up taking** his knife away from him .

Otro elemento que activa su utilización son los marcadores temporales *finally*, *eventually* y *ever*:

<Grafton>: - Supuse que <acabaríamos conociéndonos> .

-->par_en: " I figured we 'd meet eventually .

4.4.1.6. Recapitulación de resultados: grupo I

- CORPUS COMPARABLE – FASE I

Subtotal Grupo I	599	184,6	828	159,1	+7,63	< 0,01	SÍ
-------------------------	------------	-------	------------	-------	--------------	--------	-----------

Los resultados para aquellas perífrasis que contaban con equivalentes evidentes o inmediatos en inglés confirmaron nuestras expectativas —es decir, no resultaron infrarrepresentados en el subcorpus de traducciones— solo en tres de nueve casos: «acabar por + infinitivo», «empezar a + infinitivo» (que, de hecho, se encuentra sobrerrepresentada) y «volver a + infinitivo». Un hecho interesante que no habíamos anticipado al diseñar el estudio es que seis de las perífrasis de este grupo, que pensábamos que estarían bien representadas en traducciones, o al menos en la misma medida que en originales en español, resultaron infrarrepresentadas. Esto provocó que la balanza se inclinara notablemente hacia el subcorpus de originales y que en este grupo, contra toda expectativa, registráramos una infrarrepresentación en textos traducidos de estas perífrasis que en principio no tendría por qué ocurrir.

- CORPUS COMPARABLE – FASE II

Subtotal Grupo I	709	193	944	193,2	-0,00	n. s.	NO
-------------------------	------------	-----	------------	-------	--------------	-------	-----------

En esta segunda fase de compilación del corpus observamos que los casos hallados se han incrementado bastante en ambos subcorpus pero de manera que, al normalizar las cifras, apenas hay diferencia entre ellos; de hecho, salta a la vista que existe mucha menos diferencia que en la fase anterior, dado que las cifras totales normalizadas son prácticamente iguales. En este caso, la conclusión es más afín a muestras predicciones iniciales: no parece existir una diferencia significativa en la manera en la que los traductores del inglés y los autores que escriben en español sin pasar por un texto original utilizan las perífrasis verbales que tienen equivalentes fácilmente accesibles en la lengua de partida.

Por supuesto, si entramos en detalle, volvemos a observar que hay perífrasis que son la excepción a esta conclusión. Resultaron no infrarrepresentadas las mismas perífrasis que en la fase anterior del corpus («acabar por + infinitivo», «empezar a + infinitivo» y «volver a + infinitivo»), con la diferencia de que «empezar a + infinitivo» ha dejado de estar sobrerrepresentada para pasar a estarlo «volver a + infinitivo» y que se ha añadido al grupo la perífrasis «terminar + gerundio», con lo que en lugar de tres son cuatro las que hacen honor al grupo en que se encuentran.

Sin embargo, cinco de nueve perífrasis están infrarrepresentadas aun en un grupo de estas características, y solo dos de los cuatro *no* podríamos decir que son verdaderamente rotundos (los otros dos indican una falta de significación estadística de la diferencia entre subcorpus, pero no mayor presencia en traducciones sino al contrario) y, justamente, son los que cuentan con mayor número de apariciones. Si eliminamos estos «no» rotundos que desequilibran el conjunto, los de «empezar a + infinitivo» y «volver a + infinitivo» del cómputo, el grupo presentaría una infrarrepresentación general muy significativa (LL +32,59 $p < 0,0001$).

Esto puede hacer dudar de la conveniencia de sumar totales generales, como habíamos apuntado al presentar la tabla de resultados, pero, sin embargo, si realizáramos el mismo procedimiento en los otros dos grupos —es decir, eliminar del cómputo las dos perífrasis con mayor número de concordancias («haber que + infinitivo» e «ir + gerundio» en el II y «dejar de + infinitivo» y «soler + infinitivo» en el III)—, los resultados parciales de cada grupo se verían alterados (la diferencia dejaría de ser significativa en el segundo y habría una infrarrepresentación general en el tercero), y, curiosamente, aunque con un significación estadística algo menor, el total de perífrasis de la tabla seguiría estando infrarrepresentado (LL +5,18 $p < 0,05$).

- CORPUS PARALELO

Si nos esforzáramos por buscar una tendencia general en el subcorpus de originales, podríamos observar que hay un patrón que se repite prácticamente en casi todas las perífrasis: en torno a un 70 % de las concordancias paralelas corresponde a los dos o tres elementos que hemos denominado «detonadores» principales, es decir, los equivalentes más accesibles y habituales que ya habíamos predicho que encontraríamos bien representados en EN: *begin/start* para «empezar a + infinitivo» (68,6 %) y para «comenzar a + infinitivo» (74,1 %), *again/back* para «volver a + infinitivo» (68,3 %),

keep/continue/still para «seguir + gerundio» (64,8 %) y *keep/continue/go on* para «continuar + gerundio» (71,4 %).

No encontramos, sin embargo, patrones claros para la parte correspondiente a las adiciones y reformulaciones, es decir, al uso libre de la perífrasis por parte del traductor, sin estímulos directos de estructuras concretas de la lengua original. Este puede variar desde casi un 30 % para «continuar + gerundio» hasta un 0 % para «comenzar a + infinitivo», ambas perífrasis claramente infrarrepresentadas en traducciones, pasando por un 13 % en el caso de las dos perífrasis mejor representadas en traducciones («empezar a + infinitivo» y «volver + infinitivo»).

En las tres perífrasis perfectivas con casos computables no se cumple este patrón, puesto que la mitad de sus concordancias o más corresponden a adiciones y reformulaciones: 60 % en el caso de «acabar por + infinitivo» y «acabar + gerundio» y 50 % en el de «terminar + gerundio». En las cuatro hay una utilización menor y menos coherente en traducciones, puesto que traducen muy diversas estructuras y expresiones.

Parece claro que la perífrasis no se utilizará en traducciones de manera similar a cuando un escritor se expresa espontáneamente en español si la estructura no tiene claros equivalentes estructurales en la lengua de partida (como vemos que sucede en el caso de «terminar por infinitivo» o «acabar + gerundio»), o si existe una estructura mucho más habitual para expresar ese significado que funcione en detrimento de su alternativa menos frecuente (como en el caso de «comenzar a + infinitivo» o «continuar + gerundio»). Cuanta más variedad de estructuras encontramos en los textos originales, más representada está la perífrasis en traducciones, puesto que este hecho significa que la perífrasis en cuestión es un recurso habitual en la paleta mental del traductor, que se vale de ella para solucionar problemas de variada naturaleza y eso provoca, a veces, una utilización más frecuente en traducciones que en originales, como hemos visto en «empezar a + infinitivo» y «volver a + infinitivo». Además de la influencia del texto original, pues, existe otro factor que tiene que ver, como hemos visto en la introducción teórica, con lo que Halverson denominaba *cognitive salience*: el grado de prominencia de un elemento de la lengua meta en las redes semánticas del traductor que lo conectan con las estructuras del texto original.

The basic idea is straightforward: in a translation task, a semantic network is activated by lexical and grammatical structures in the

ST. Within this activated network, which also includes nodes for TL words and grammatical structures, highly salient structures will exert a gravitational pull, resulting in an overrepresentation in translation of the specific TL lexical and grammatical structures that correspond to those salient nodes and configurations in the schematic network. (Halverson 2003: 218)

De todas formas, la influencia del original no puede obviarse. «Seguir + gerundio» cuenta con diversos equivalentes en inglés y es un elemento habitual en español y, sin embargo, sorprendentemente, se encuentra infrarrepresentada. La única razón que se nos ocurre es que dichas estructuras equivalentes en inglés no sean de uso habitual en esta lengua y los casos en los que el traductor utiliza esta perífrasis sin estímulo en la lengua de partida no sean suficientes para compensar esta escasez. Otra alternativa es que los sinónimos de esta estructura en español (p. ej. *aún*) se activen con mayor facilidad por ser más similares a las estructuras inglesas.

4.4.2. GRUPO II

En este grupo incluimos aquellas perífrasis que son consideradas características del español y no cuentan con elementos activadores evidentes en inglés pues no existen en esta lengua expresiones paralelas en cuanto a forma o estructura que puedan transmitir el mismo significado, aunque puedan existir para las maneras alternativas de expresarlo en español. Los valores que poseen son de naturaleza muy variada (más que la de los otros dos grupos), tanto modales como aspectuales y durativos.

10. HABER DE + INFINITIVO
11. HABER QUE + INFINITIVO
12. ACERTAR A + INFINITIVO
13. LLEGAR A + INFINITIVO
14. ALCANZAR A + INFINITIVO
15. ANDAR + GERUNDIO
16. IR + GERUNDIO
17. VENIR + GERUNDIO
18. VENIR A + INFINITIVO

4.4.2.1. Haber de + infinitivo

Esta perífrasis posee dos valores: modal de obligación o necesidad (en este caso «Has de comértelo» podría traducirse como *You must/have to eat it*) y temporal futuro («Si vienes, lo has de pasar bien» se traduciría entonces por *If you come, you will have a good time*). Sin embargo, cada vez se utiliza menos en su segundo valor (en beneficio de «ir a + infinitivo» y el futuro simple), y, en el primero, va quedando relegada a la lengua literaria, culta, arcaica o afectada (Gómez Torrego 1999: 3355; Gili Gaya 1981: 112), por lo que resulta interesante comprobar si hay diferencias de frecuencia de uso entre textos originales y traducidos.

Quizás este fenómeno sea menos marcado en regiones en las que el español se encuentra en contacto con el catalán, ya que, por influencia de esta lengua, esta perífrasis parece utilizarse de manera más habitual que en otras zonas hispanohablantes (Martínez Díaz 2003), por lo que cabría esperar encontrar más casos de esta perífrasis en los textos del corpus escritos por autores cuya variedad de español estuviera en contacto directo con la lengua catalana.

También expresa probabilidad o suposición, como en «Ha de ser tarde» (*It must be late*) y, como «ir a + infinitivo», se usa en enunciados exclamativos para expresar sorpresa o indignación, como en «¡Qué he de saber yo!» (*How should I know?*).

A pesar de todos los equivalentes que acabamos de proporcionar a través de los ejemplos, en los diccionarios que hemos consultado, «haber de + infinitivo» no aparece entre las traducciones posibles de *have to* y *must* para expresar obligación, puesto que hay otras traducciones más evidentes o habituales en español para estas palabras (p.ej. *tener que*). Sí que figura, por otro lado, entre los equivalentes de *must* y de *will* cuando expresan probabilidad. Con respecto a su uso en exclamaciones, «haber de + infinitivo» aparece únicamente entre los equivalentes de *should* (*Why should he lie?: «¿Por qué ha de mentir?»*).

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: HAY*/HABR*/HUB*/HABÍ*/HAN/HABÉIS/HEMOS/HAS/
HA/HE/HABID*/HABIENDO/HABIÉNDO*/HABE*

Context words: DE

Context Search Horizons: OL to 1R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Haber de + INF	63	19,4	35	6,7	+26,76	< 0,0001	SÍ

En este caso, no hay necesidad de calcular porcentajes ni normalizar las cifras para saber que en las traducciones del inglés «haber de + infinitivo» se encuentra claramente infrarrepresentada, pues, a pesar de ser el corpus de traducciones bastante más grande que el de originales, en este último encontramos casi el doble de apariciones de esta perífrasis. Antes habíamos apuntado que el valor de futuridad era cada vez menos frecuente y que tanto este valor como el de obligación se usaban cada vez menos por sentirse propios de una lengua culta y anticuada. En nuestro corpus hemos encontrado que de las 63 apariciones en el subcorpus de originales, sólo 8, es decir, un 12,7 %, tienen valor de temporalidad futura y, de las 35 del subcorpus de traducciones, ninguna.

La frecuencia de aparición más elevada en textos originales que en traducidos también podría explicarse, como hemos mencionado en el apartado de caracterización de la perífrasis, por la influencia del catalán en escritores que están en contacto con esa lengua. Si nos fijamos en los textos en los que aparece, vemos que 44 de las 63 (un 69,8 %, cifra elevadísima) pertenecen a un único autor, Vázquez Montalbán, escritor catalán que vivió prácticamente toda su vida en Barcelona. De las restantes apariciones, 12 son de Andreu Martín, otro escritor nacido y criado en Barcelona, y sólo 7 de Giménez Bartlett (todas las de esta última poseen valor obligatorio), que, a pesar de no ser catalana, estudió en Valencia y reside en Barcelona. Los dos autores que no aparecen en las concordancias son un madrileño, Jorge Martínez Reverte, y un malagueño que vive en Madrid (Juan Madrid). En cuanto a las traducciones, solo aparece en las de Hunter y de Rendell, hechas ambas por traductores barceloneses, Milla y Camps.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

```
[word="(¿|i)?ha(n|s)?|(¿|i)?hay(a|á).*(¿|i)?habr.*(¿|i)?hub.*(¿|i)?hab(i|í)(d|e|é|a).*(¿|i)?hab(e|é).*(¿|i)?he|(¿|i)?hemos" %c] "de";
```

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Haber de + INF	63	17,2	131	26,8	-8,86	< 0,01	NO

Vemos en esta segunda fase que el número de apariciones en ES no ha variado, dado que la autora vasca Soledad Núñez no aporta ningún caso al cómputo, mientras que el de TRAD se ha cuadruplicado, lo que ha hecho que esta perífrasis pase de estar infrarrepresentada en la primera fase del corpus a sobrerrepresentada.

Esta importantísima diferencia podría estar apoyando la hipótesis que apuntábamos más arriba del contacto entre lenguas como principal factor de utilización en español, también traducido, de esta perífrasis cada vez más en desuso: al eliminar a los traductores argentinos del subcorpus de traducciones, los que quedan son de origen catalán o han vivido en zonas de habla catalana o valenciana la mayor parte de sus vidas, lo que incrementa la utilización de esta perífrasis enormemente.

Otra hipótesis podría ser la percepción o las ideas preconcebidas que tiene el mismo traductor del lenguaje de una traducción o de la obra que traduce, como poseedora de un grado de literariedad que le empuja a utilizar una expresión que, como hemos explicado, hoy en día ostenta ciertas connotaciones que la hacen parecer más propia de un lenguaje culto y cuidado.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Vázquez	44	50,1	Leonard	59	75,8
Martín	12	21,3	Cornwell	53	48,2
Giménez	7	8,2	Grafton	14	15,1
Madrid	0	0	Ellroy	5	4,5
Martínez	0	0	Highsmith	0	0
Núñez	0	0			
Total	63	17,2		131	26,8

Tabla 23 - Distribución de «haber de + INF» por texto

En cuanto a su distribución, volvemos a comprobar que la utilizan solo tres de seis autores y existe una marcada diferencia en el uso de esta perífrasis. Además de lo mencionado anteriormente con respecto a la hipotética influencia del contacto entre lenguas, es necesario mencionar que el mayor porcentaje de uso (un 70 % del total) corresponde a Vázquez, una de cuyas particularidades es la utilización de un lenguaje esmerado en sus novelas. De hecho, esta es una de las pocas perífrasis en las que Vázquez Montalbán va a la cabeza, ya que, a pesar de que su texto contiene el mayor número de palabras, suele estar en las últimas posiciones o a veces, intermedias,

probablemente debido a que la mayor parte de su novela está escrita desde el punto de vista del narrador y, por tanto, cuenta con una menor proporción de diálogo en su texto que los demás autores del subcorpus (Martínez Vilinsky 2014:5), lo que equivale a un menor grado de oralidad. Los otros dos autores que la utilizan cuentan solo con 12 y 7 casos respectivamente, por lo que el uso de la perífrasis no deja de ser limitado.

En TRAD vuelve a resultar curioso que el autor con menos palabras (Leonard) sea el que más utiliza la perífrasis —pero aquí ya no podemos aludir a la procedencia geográfica como causa principal, pues todos los traductores de la segunda fase se mueven o movían en un entorno catalanoparlante—, y Highsmith, que suele estar a la cabeza de la distribución no cuenta con ningún caso.

De las 63 apariciones en el subcorpus de originales, sólo 6, es decir, un 9,5 % tiene valor de temporalidad futura (casi todas en el texto de Andreu Martín) y, de las 131 del subcorpus de traducciones, únicamente 4 (3 %), distribuidas entre las de Grafton, Cornwell y Leonard.

<Martín>: Era Papa Ogro , el padre de Avelino , el que <había de> arruinar a toda la familia .

Sin embargo, es preciso apuntar que estos casos son los que consideramos «puramente» temporales, pues a veces los valores de temporalidad futura y obligación están muy ligados y no es posible distinguirlos, como en:

<Cornwell> - Sólo si no <has de> conducir esta noche .

-->par_en: " Only if you 're not driving anywhere tonight . "

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
(Have) (got) to	51	38,9
Need to	19	14,5
∅	12	9,2
Reformulación	10	7,6
Futuro (will/be going to/present continuous)	10	7,6
Infinitive wh- clause	7	5,3
Would	7	5,3
Should	6	4,6
Suposición (suspect/think/be supposed to)	6	4,6
Must	3	2,3
TOTAL	131	100

Tabla 24 - Equivalentes de «haber de + INF» en EN

Un 70,9 % de las concordancias corresponde a valores de obligación o necesidad: 93 casos provienen de la traducción de los modales y semimodales *need to*,

would, should, must, have to y have got to (a veces con omisión de *have*, lo que añade coloquialidad a la frase y hace que resulte más extraña la utilización de esta perífrasis, más propia de otros registros), y también a oraciones subordinadas formadas con un pronombre o adverbio interrogativo y un infinitivo —las llamadas «infinitive wh-clauses» (Close 1985: 97)—.

<Leonard>: - Si <he de> hacerlo , sí .

-->par_en: " If I have to . "

<Cornwell>: <He de> hablar con ella .

-->par_en: I got to talk to her . "

<Cornwell>: - No sé si <he de> alegrarme de verlo o no - le dije al abrir la puerta .

-->par_en: " I 'm not sure if I should be glad to see you or not , " I said when I opened the door

<Leonard>: Luego ya le diré qué <ha de> hacer » .

-->par_en: I will tell you what to do next .

Todos los casos de *would* hallados corresponden al supuesto explicado más arriba: enunciados exclamativos e interrogativos que indican sorpresa, indignación o desacuerdo. De hecho, todos empiezan con la partícula interrogativa *why*:

<Cornwell>: ¿Por qué <habría de> mentir Susan ?

-->par_en: Why would Susan lie ?

Quizás la similitud de *have to* con «haber de» hace que sea esta la estructura que en mayor proporción activa su utilización (más de la mitad de este casi 71 %), pero hay que tener en cuenta que solo 46 de 540 concordancias de *have to* en EN se traducen por la perífrasis «haber de + infinitivo». Probablemente se traduzca mucho más por la más habitual «tener que + infinitivo» (una búsqueda rápida de «tener que» en sus diferentes conjugaciones arroja más de 700 concordancias).

Una pequeña parte de las concordancias corresponde al tercer valor de la perífrasis, la suposición o hipótesis: se utiliza como traducción de frases con los verbos *suspect* y *think* y con la expresión *be supposed to*, aun acompañada también de «se supone» para remarcar la duda.

<Cornwell>: - Se supone que los guardias no <han de> pararse por nada cuando conducen a un condenado a muerte , pero si se trataba de una conspiración pudo ocurrir cualquier cosa .

-->par_en: " Guards **are n't supposed** to stop for anything while transporting a death row inmate . But if some conspiracy were involved , anything could have happened .

El 7,6 % corresponde al uso de auxiliares y tiempos verbales utilizados para expresar futuro (*will, be going to y present continuous*). Hay más concordancias clasificadas con este valor en inglés (10) que en español (4), como vemos, porque al traducir este futuro por *haber de*, la frase en español adquiere connotaciones de obligatoriedad o necesidad y no creemos que pueda considerarse un caso «puramente» temporal, como decíamos arriba.

<Leonard>: - El que tengo allí y el que <ha de> llegar .

-->par_en: " What's there now and what **will be coming in** . "

La mayor parte de reformulaciones (7 de 10) viene dada por la utilización de expresiones idiomáticas utilizadas en inglés para captar la atención del interlocutor o enfatizar lo dicho, como *you know, by the way o to be honest*. Con respecto a lo que apuntábamos de aumentar el grado de literariedad percibida del texto original, encontramos que este uso de la perífrasis para traducir dichas expresiones modifica, a nuestro juicio, el grado de familiaridad o coloquialidad del diálogo, ya que estas fórmulas conversacionales suelen utilizarse en contextos informales o neutros, y el «ha de saber» o «he de ser sincero» que vemos en las traducciones eleva el registro y el grado de formalidad del enunciado:

<Cornwell>: <Ha de> saber que Waddell ya había tenido un par de hemorragias nasales anteriormente , y subidas de presión

-->par_en: **You know** , Waddell 'd had a couple of nosebleeds in the past , and high blood pressure .

<Grafton>: El asunto me dolió , si <he de> serte sincero .

-->par_en: The whole deal kind of soured me , **to tell you the truth** .

Estas reformulaciones, junto con las adiciones, conforman el 16,8 % restante.

4.4.2.2. Haber que + infinitivo

Perífrasis de valor modal de obligación o necesidad. Es impersonal, se conjuga sólo en la tercera persona del singular, implica un actor animado y no tiene una

correspondencia estructural exacta en inglés pues éste necesita un sujeto que realice la acción además del auxiliar equivalente (*must, have to, should* o *need*): «Hubo que contárselo» se convierte en «*We (one/they) had to tell him*». Existen también algunas estructuras impersonales que pueden trasladar el significado de esta perífrasis, como *to be necessary to* o *it should/must/have to/ought to be* + participio, aunque no remiten formalmente a la perífrasis. Entre las alternativas a esta perífrasis que se encuentran disponibles en español se incluyen las expresiones: *ser necesario, deber* o *tener que*.

En general, es posible insertar diversos elementos (p. ej. sujetos o complementos) entre el auxiliar y el verbo principal de una perífrasis. En este caso, a pesar de que, a primera vista no parece posible, sí que encontramos ocasiones en las que sucede. Hemos hallado, por ejemplo, un caso de intercalación de la conjunción *pues*, y el adverbio *entonces* también podría ser una posibilidad, aunque no hayamos encontrado ningún caso en nuestro corpus. Sin embargo, los casos se cuadriplican si tenemos en cuenta estas intercalaciones, y la práctica totalidad de las concordancias se vuelve irrelevante, por lo que hemos preferido mantener la suposición de que esto no ocurre con frecuencia significativa. Se acercan más a las locuciones las siguientes expresiones que no tendremos en cuenta: «Hay que ver», «Hay que fastidiarse/joderse».

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: HAY*/HABR*/HUB*/HABÍ*/HAN/HABÉIS/HEMOS/HAS/HA/HE/HABID*/HABIENDO/HABIÉNDO*/HABE*

Context words: QUE

Context Search Horizons: OL to 1R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Haber que + INF	147	45,3	67	12,9	+80,29	< 0,0001	SÍ

Como cabía esperar del carácter especial de esta perífrasis, su falta de estructuras similares en inglés y la existencia de otros equivalentes españoles más cercanos a la lengua origen (por no ser estructuras impersonales y requerir sujeto), su frecuencia de aparición en textos originales de esta primera fase del corpus resultó ser más de tres veces mayor que en traducciones.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?ha(n|s)?|(¿|i)?hay.*|(¿|i)?habr.*|(¿|i)?hub.*|(¿|i)?hab(i|í)(d|a|e|é).*(¿|i)?hab(e|é).*(¿|i)?he|(¿|i)?hemos" %c] "que";

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Haber que + INF	159	43,3	75	15,3	+59,61	<0,0001	SÍ

En esta segunda fase vemos que hay un ligero incremento en la frecuencia de aparición de la perífrasis en ambos subcorpus (13 casos más en ES y 24 en TRAD). Los casos son bastante numerosos y este incremento no afecta a la relación entre subcorpus: la perífrasis sigue estando claramente infrarrepresentada en traducciones.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	62	73	Grafton	26	28
Martínez	27	49,7	Cornwell	19	17,3
Vázquez	26	29,6	Leonard	16	20,6
Martín	25	44,5	Highsmith	11	11,2
Núñez	12	28,2	Ellroy	3	2,7
Madrid	7	16,9			
Total	159	43,3		75	15,3

Tabla 25 - Distribución de «haber que + INF» por texto

Todos los autores de ambos subcorpus utilizan esta perífrasis en mayor o menor medida. En TRAD llama la atención la escasez en Ellroy, aunque ya hemos explicado el motivo, y la abundancia en Grafton, que es el cuarto texto en número de palabras.

Los tiempos verbales se distribuyen de manera similar. El auxiliar en presente es el que más abunda, con un 62 % en ES y un 65 % en TRAD, seguido del pretérito imperfecto, con un 21 % en ES y un 25 % en TRAD. Sí apreciamos menor variedad en las formas que quedan: en ambos subcorpus encontramos futuros, condicionales, subjuntivos y pretéritos perfectos simples. El infinitivo (en combinación con otra perífrasis) y el pretérito imperfecto del subjuntivo, sin embargo, solo aparecen en ES.

<Martínez>: - Va a <haber que hacer> algo .

<Giménez>: Quizás <hubiera que> ofrecer la otra mejilla , o aparecer de una vez por todas en televisión, jurando que los policías también sabíamos llorar y arrepentimos.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
(Have) (got) to	27	36
Reformulación	22	29,3
∅	8	10,7
Infinitivo con to	6	8
Need + -ing/infinitive	5	6,7
(Had) better	4	5,3
Futuro	2	2,7
Should	1	1,3
TOTAL	75	100

Tabla 26 - Equivalentes de «haber que + INF» en EN

Un 57,3 % de las concordancias corresponde a valores de obligación o necesidad, principalmente a «*have to + infinitivo*» y «*(have) got to + infinitivo*» (36 %) pero también hay algunos casos de *need to y need -ing, should y (had) better*:

<Cornwell>: Mientras Lucy lo contemplaba en silencio y con curiosidad , pontificó sobre la acumulación de residuos que había dejado el disolvente que yo utilizaba y me anunció que mi Ruger probablemente tenía « espolones » que <habría que> limar .

-->par_en: While Lucy watched him in curious silence , he pontificated on the residue buildup left by the solvent I used and informed me that my Ruger probably had " spurs " that **needed filing** .

<Grafton>: - <Hay que> esperar a la policía - dije .

-->par_en: " **You better** wait for the cops , " I said .

También encontramos, dentro de este valor de necesidad, infinitivos a secas, reemplazando oraciones de relativo o en las ya mencionadas «infinitive wh- clauses» (pronombre o adverbio interrogativo + infinitivo):

<Highsmith>: Tom le comunicó las buenas noticias : llegaba un huésped , alguien para quien cocinar y para el que <habría que> preparar la habitación de invitados .

-->par_en: Tom imparted his good news : **someone** to cook for , **to prepare the guest room for** .

<Grafton>: Nos explicó también qué <había que> hacer si corríamos peligro de estrellarnos , hacernos papilla y carbonizarnos por caer hacia la corteza terrestre a velocidad supersónica desde ochocientos metros de altura .

-->par_en: She also explained **what to do** if we were on the verge of being smashed , crushed , and burned by hurtling at high velocity from our flying altitude of twenty-six thousand feet straight down through the earth 's crust .

Por otro lado, también observamos en inglés valores de futuridad, a veces mezclado con obligatoriedad, como en estos casos que incluyen la construcción «*be + infinitivo*» y un futuro hipotético en oración condicional:

<Grafton>: Si <había que> dar media vuelta , era el momento indicado .
 -->par_en: If I was going to bail out , now was the time to do it .
 <Ellroy>: Tres hombres que <hay que> detener en la comisaría de Wilshire .
 -->par_en: Three men to be detained at Wilshire Station .

El resto de concordancias (un 40 %) corresponde a adiciones y reformulaciones. Un buen número de estas últimas (7) surge de la reexpresión de sentencias o verdades establecidas (al menos en opinión de los personajes que las pronuncian) en inglés, que en español pueden pasar perfectamente a tener un sujeto impersonal puesto que se refieren a cuestiones generales y no a un sujeto específico, a pesar de que el inglés precise la explicitación del mismo:

<Cornwell>: No <había que> ofender a los vecinos tendiendo la colada a la vista ni presentándose sin avisar .
 -->par_en: You did not offend your neighbors by putting up clotheslines or dropping by unannounced .
 <Leonard>: - Es lo que <hay que> hacer , Max , llegar al final cuando se empieza .
 -->par_en: " It's what you do , Max , to go through with it once you start .

Muchas de las concordancias que incluyen *have to*, poseen sujetos inclusivos o generales; especialmente forma de pronombres indefinidos —*you, we y one*—, y esto es lo que probablemente active la utilización de la perífrasis impersonal en español. También funciona como equivalente en casos de sujeto *it* o de cosa y en oraciones pasivas.

<Highsmith>: <Había que> rellenar formularios y aceptar que te colocaran un aparato de escucha que , naturalmente , escuchaba también todo lo demás .
 -->par_en: One had to fill out forms and submit to a monitoring device , which would of course monitor everything else .
 <Cornwell>: Pero la alfombra <hay que> quitarla .
 -->par_en: " But the rug 's got to go .
 <Leonard>: Ahora , Louis entendió que al tipo le gustaba exhibirse ; para que su cerebro se pusiera en marcha <había que> retarlo de algún modo .

-->par_en: The way Louis saw him now , the guy liked being on the muscle ; he had to be challenged in some way to get his head to work

Sin embargo, aunque son los menos, también encontramos sujetos personales animados, como en:

<Grafton>: - Era evidente que <había que> decirle algo - .

-->par_en: I realized I d better give him one .

Es evidente que en todos estos casos, las alternativas de traducción (por estructuras más parecidas a las del original, por ejemplo) son numerosas, y la utilización de esta perífrasis no es imprescindible; depende en gran medida del criterio del traductor. Es probablemente por este motivo por el que no encontramos tantos casos en TRAD como en ES.

4.4.2.3. Acertar a + infinitivo

Su naturaleza como perífrasis es defendida por Gómez Torrego, aunque no viene incluida en casi ninguna otra clasificación (apenas se menciona en García Fernández 2012: 47 como perífrasis modal). Se utiliza para referirse a algo que ha ocurrido por casualidad o sorpresa, como en «Acertó a pasar por ahí una mujer (*A woman happened to pass by*), o con acierto, como en «Acertó a atajar el gol» (*He managed/was able to catch the ball*). Otro uso bastante frecuente es el que hallamos en enunciados negativos como «No acierto a comprender de qué habla» (*I fail to / just can't understand what he's talking about*). Sin embargo, «acertar a + infinitivo» no aparece en los diccionarios como equivalente de las expresiones inglesas que acabamos de reproducir. Un ejemplo como «*I fail to see*» se traduce sencillamente por «no veo».

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: ACERT*/ACIERT*/ACIÉRT*

Context words: A

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Acertar a + INF	6	1,8	0	0	+11,48	< 0,001 ¹²⁷	SÍ

Como podemos observar en la tabla, el subcorpus de traducciones no incluye ningún ejemplo de esta perífrasis, mientras que en el de originales encontramos una frecuencia de 1,8 apariciones por cien mil palabras.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?ac(i)?(e|é)rt.*" %c] []{0,5} "a";

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Acertar a + INF	7	1,9	0	0	+11,85	< 0,001 ¹²⁸	SÍ

Comprobamos que, tras cambiar cuatro de las cinco obras incluidas en el subcorpus de traducciones, las cifras continúan siendo prácticamente iguales y la perífrasis sigue estando infrarrepresentada.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martínez	3	5,5	Cornwell	0	0
Vázquez	1	1,1	Grafton	0	0
Núñez	1	2,3	Elroy	0	0
Martín	1	1,8	Highsmith	0	0
Giménez	1	1,2	Leonard	0	0
Madrid	0	0			
Total	7	1,9		0	0

Tabla 27 - Distribución de «acertar a + INF» por texto

A pesar de no ser en absoluto frecuente, encontramos la expresión «acertar a + infinitivo» en todos menos uno de los textos españoles, lo que indica que, aunque poco, se utiliza en español. Sin embargo, no hallamos en ninguna de las obras de esta fase del subcorpus de traducciones ni de la anterior un ejemplo de esta perífrasis. De los siete ejemplos de ES, cuatro poseen valor de logro o acierto, uno de casualidad y dos son enunciados negativos que transmiten falta de comprensión o de capacidad por parte del hablante:

¹²⁷ Test exacto de Fisher: 0,003207 (p < 0,01)

¹²⁸ Test exacto de Fisher: 0,002675 (p < 0,01)

<Vázquez>: Al vendedor se le había paralizado la sonrisa , la palabra , la gesticulación y por fin <acertó a> balbucir : - Pero hombre.. . [logro]

<Martín>: Y como <acierte a> pasar otra puta , o la procesión de Corpus Christi , ya me explicarás , van apañadas . [casualidad]

<Núñez>: No <acierto a> descifrar este enigma . [enunciado negativo]

- CORPUS PARALELO

De la falta de ejemplos en TRAD podemos deducir que, probablemente, los valores que posee la perífrasis «acertar a + infinitivo» puedan expresarse en español mediante otras fórmulas más accesibles en el lexicón mental bilingüe del traductor. Para comprobarlo, realizamos una búsqueda de algunos de sus equivalentes ingleses en el corpus paralelo. Una búsqueda de «"fail.*" "to"» arroja ocho concordancias. Tres de ellas recurren a la simple negación, como habíamos visto que también sucedía en los diccionarios bilingües:

<Highsmith>: I <fail> to see - "

-->par_trad: No entiendo por qué.. .

<Leonard>: The guy at Glades Mutual told Max to " use him for heavy work , " like picking up guys that <failed> to appear . -

->par_trad: El tipo de la Glades le dijo a Max que lo « usara para trabajos duros » , como ir a buscar a los que no se presentaban .

En un caso se complementa la negación con el verbo *conseguir*, que añade la sensación de que hay una dificultad para alcanzar aquello que no se logra, por lo que se acerca más al sentido de la perífrasis:

<Leonard>: He did tell one - how he drove all the way to Van Horn , Texas , to return a defendant who'd skipped on a five-hundred-dollar bond - and they didn't get it , <failed> to understand the street value of what that kind of dedication meant .

-->par_trad: Sí que solía contar una historia - la de cómo había ido en coche hasta Van Horn , Texas , para recoger a un cliente que se había saltado la condicional con una fianza de quinientos dólares - , pero no lo entendían , no conseguían apreciar el valor que ese tipo de dedicación tenía en la calle .

El resto recurre a la perífrasis «dejar de + infinitivo» y a verbos como *escapar* u *olvidar*, que denotan una imposibilidad de conseguir lo que se pretende:

<Cornwell>: It was a fact I <failed> to understand , for I had not known it was possible to criticize plants into good health .

-->par_trad: Era algo que escapaba a mi comprensión , pues nunca había sabido que se pudiera dar buena salud a las plantas a fuerza de críticas ..

<Grafton>: " I might have <failed> to mention a couple things . "

-->par_trad: Puede que haya dejado de mencionar un par de cosas .

Del verbo *to manage* encontramos bastantes más concordancias (47), con el sentido casi siempre de «alcanzar a» o «lograr». Los traductores eligen utilizar, en el 34 % de los casos, el verbo *conseguir*:

<Leonard>: he was looking at her and <managed> to smile .

-->par_trad: Él se la quedó mirando y consiguió sonreír .

En el 23 % de los casos se opta por «arreglárselas para hacer algo» y sus variantes «ingeniárselas», «apañárselas» y «hacérselo»:

<Grafton>: I vacuumed in that direction , <managing> to get a glimpse of her when she let him in.

-->par_trad: Avancé en aquella dirección arreglándomelas para ver a Laura cuando hizo pasar al mozo .

El 21 % corresponde a adiciones:

<Grafton>: I 'd have to get out the instructor 's manual if I ever <managed> to get laid again .

-->par_trad: Tendría que sacar el manual de instrucciones si volvía a meterme en la cama con alguien .

En cuanto al 22 % restante, con uno, dos o tres casos únicamente cada una, encontramos «lograr», «poder», «llegar», «esforzarse por» y «por fin».

<Highsmith>: Tom <managed> a smile .

-->par_trad: Tom se esforzó por sonreír .

<Grafton>: It took a fair amount of huffing and puffing before they <managed> to set it down on the floor in the middle of the room .

-->par_trad: Les costó muchos tirones y bufidos hasta que por fin la depositaron en el suelo , en el centro de la habitación .

Para el último valor, el de ocurrir algo por casualidad o sorpresa, hemos seleccionado la expresión *happen to* (29 casos). En muy pocas ocasiones se traslada al

español esa connotación de suceso fortuito o sorprendente; encontramos omisiones en un 62,1 % de las concordancias. Los traductores de Highsmith y Cornwell, Núñez y Mustieles, la omiten en prácticamente todas las ocasiones.

<Highsmith>: Both Marie with her big red smile and balding Georges <happened to> be behind the bar just as Tom pushed his cup and saucer away . "

-->par_trad: Marie , con su gran sonrisa de labios rojos , y Georges , con su cabeza casi calva , **estaban detrás de la barra** cuando Tom dejó la taza y el platillo.

El traductor de Leonard, de Hériz, es quien más intenta transmitirla, sobre todo a través de la expresión «por casualidad» (20,7 % de las concordancias; también utilizada en el texto de Cornwell y el de Grafton):

<Leonard>: Like if during deer season they were out and <happened to> run into a couple of friendly girls in a bar ..

-->par_trad: Como cuando estaban cazando y coincidían **por casualidad** con un par de chicas simpáticas en un bar .

También encontramos expresiones con el verbo *dar*, «darle a alguien por hacer algo» y «darse el caso» (3):

<Cornwell>: Not a sound , except on rare occasion when the security guard <happens to> wander through .

-->par_trad: No hay ruidos , excepto en las raras ocasiones en que al guardia de seguridad **le da por** hacer la ronda .

<Cornwell>: If I , this guard , also <happened to> be the contact for Ben Stevens , the person forking over cash in exchange for making sure Waddell 's body was n't printed at the morgue , for example , then I 'd know exactly where to turn . "

-->par_trad: Si además **se diera el caso** de que yo , el guardia , era el contacto de Ben Stevens , la persona que le pagaba para que no le tomaran las huellas a Waddell en la morgue , por ejemplo... . , si fuera esa persona sabría exactamente a quién recurrir.

Por último, encontramos dos casos en los que se intenta dar alguna marca de énfasis a la frase mediante el adverbio *además* y la perífrasis «acabar + gerundio»:

<Leonard>: Guy that <happened to> be her husband at the time .

-->par_trad: Un tipo que en aquella época era **además** su marido .

<Leonard>: It was because he was pacing the room , Raynelle nodding on the sofa for what company she was , and <happened to> look outside as he came to the window , he saw Max Cherry on the sidewalk .

-->par_trad: Si vio a Max Cherry en la acera fue porque se estaba paseando por la habitación y la única compañía que proporcionaba Raynelle era algún que otro gesto con cabeza, y acabó mirando por la ventana.

Como vemos, los valores que comporta la utilización de la perífrasis «acertar a + infinitivo» muchas veces no se reproducen en el texto traducido o, si los encontramos, es a través de otras expresiones españolas más habituales.

4.4.2.4. Llegar a + infinitivo - Alcanzar a + infinitivo

Estas dos perífrasis de significado equivalente poseen valor perfectivo e indican la culminación de un proceso. A veces pueden adquirir connotaciones negativas, como en «Llegó a tener que pedir limosna» (*He even had to beg*) o también de logro, como en «Llegarás a ser un hombre de bien» (*You will grow up to become a good man*). *Llegar* es mucho más frecuente como auxiliar que *alcanzar*, y en los dos casos anteriores, no podría intercambiarse por este último. En frases negativas, indican que no se ha conseguido algo a pesar de haber estado a punto de hacerlo («No alcanzó a terminar el trabajo a tiempo»). En inglés se utiliza una amplia variedad de fórmulas para transmitir el mismo matiz, como «*to go so far as to*», «*to grow/get to + infinitive*», «*to manage to*», «*to become/grow/get + -ed*», *even* o *finally*, que, por supuesto, tienen todas equivalentes más literales en español.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: LLEG*/LLÉG* y ALCANZ*/ALCÁNZ*/ALCANC*/ALCÁNC*

Context words: A

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Llegar a + INF	50	15,4	46	8,8	+7,37	< 0,01	SÍ
Alcanzar a + INF	0	0	6	1,1	-5,81	< 0,05 ¹²⁹	NO

¹²⁹ Test exacto de Fisher: 0,0887 (n. s.)

Confirmamos que «alcanzar a + infinitivo» es menos frecuente que «llegar a + infinitivo». En el subcorpus de traducciones las apariciones son escasas y en el de originales ni siquiera aparece una vez. Por lo tanto, la hipótesis no se confirma para esta perífrasis pero sí para «llegar a + infinitivo», con una diferencia de 66 apariciones por millón de palabras a favor del subcorpus de originales.

Es la primera vez que nos encontramos con una divergencia entre los resultados del test de log-likelihood y el de Fisher. En casos anteriores, hemos visto que lo que difería era el umbral a partir del cual consideramos significativas las diferencias de frecuencias. Por ejemplo, el test de log-likelihood decía que la infrarrepresentación de la perífrasis «terminar por + infinitivo» era significativa al nivel de $p < 0,001$, mientras que el test de Fisher solo lo hacía al de $p < 0,01$. Sin embargo, en este caso, el test exacto de Fisher nos dice, a diferencia del de log-likelihood, que no hay diferencias significativas entre ambos subcorpus. Mantenemos en la tabla el valor del test original, pero, como hemos explicado en la introducción, el color de la celda¹³⁰ de la columna final refleja el resultado del Test de Fisher, que en principio se considera más preciso con frecuencias bajas como las que se presentan en esta ocasión.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

```
[word="(¿|i)?ll(e|é)g.*" %c] []{0,5} "a"
[word="(¿|i)?alc(a|á)n(z|c).*" %c] []{0,5} "a";
```

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Llegar a + INF	66	18	48	9,8	+10,32	< 0,01	SÍ
Alcanzar a + INF	1	0,3	10	2	-6,20	< 0,05 ¹³¹	NO

A pesar de que en esta segunda fase encontramos más casos de cada una de las cuatro perífrasis, los resultados apenas se diferencian: «llegar a + infinitivo» continúa encontrándose infrarrepresentada en traducciones mientras que la sobrerrepresentación de «alcanzar a + infinitivo» ahora es significativa. Está claro que esperábamos encontrar una infrarrepresentación en el corpus de traducciones de ambas perífrasis (puesto que

¹³⁰ Recordemos que el rojo se reservaba para los casos en los que la frecuencia era mayor en traducciones y, además, esta diferencia se consideraba significativa, mientras que el morado representaba los casos en los que dicha diferencia a favor de las traducciones no podía considerarse estadísticamente significativa.

¹³¹ Test exacto de Fisher: 0,02973 ($p < 0,05$)

las hemos clasificado en el grupo II), pero es posible que esto no haya sido así para «alcanzar a + infinitivo» debido al hecho de que esta perífrasis es muy poco frecuente en español estándar actual, al igual que otras como «venir a + infinitivo» o «venir + gerundio», según los resultados de nuestro corpus de originales y el de Fente Gómez (1971: 135, 141).

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	20	23,6	Highsmith	19	19,3
Martínez	16	29,4	Cornwell	16	14,6
Núñez	12	28,2	Leonard	6	7,7
Vázquez	9	10,2	Ellroy	4	3,6
Martín	5	8,9	Grafton	3	3,2
Madrid	4	9,6			
Total	66	18		48	9,8

Tabla 28 - Distribución de «llegar a + INF» por texto

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Núñez	1	2,3	Cornwell	7	6,4
Giménez	0	0	Grafton	1	1,1
Madrid	0	0	Ellroy	1	0,9
Martín	0	0	Highsmith	1	1
Martínez-	0	0	Leonard	0	0
Vázquez	0	0			
Total	1	0,3		10	2

Tabla 29 - Distribución de «alcanzar a + INF» por texto

«Llegar a + infinitivo» está bien representada en ambos subcorpus y distribuida de manera similar entre los autores y traductores. Aunque es evidente que no llega a alterar el resultado final, en el nuevo texto del subcorpus de originales finalmente aparece un caso de «alcanzar a + infinitivo» en sentido negativo:

<Núñez>: No <alcanzaba a> ver la gravedad de la situación .

En traducciones encontramos 10 casos, 7 de los cuales pertenecen al texto de Mustieles, el traductor de Cornwell, que también es de los dos que más utiliza «llegar a + infinitivo», lo que puede apuntar a una abundancia de estructuras equivalentes en el texto original en inglés o a un rasgo estilístico del traductor. Lo que está claro es que en los textos traducidos se hace un uso mayor de esta perífrasis del que en principio sugerirían los textos originales españoles.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
∅	7	70
Reformulación	3	30
TOTAL	10	100

Tabla 30 - Equivalentes de «alcanzar a + INF» en EN

No hay un patrón claro en los equivalentes de «alcanzar a + infinitivo», pues en el 70 % de las concordancias el original se limita a decir que el personaje puede o no puede realizar cierta acción:

<Cornwell>: Desde donde me hallaba no <alcanzaba a> verle el rostro .

-->par_en: From where I stood , I could not see her face .

<Leonard>: Unos pocos kilómetros de polvo y de piedras que golpeaban los bajos de la furgoneta y por fin Louis <alcanzó a> ver la silueta de una granja entre los árboles (...)

-->par_en: A few miles of dust and stones hitting the underside of the van and Louis could see a farm layout through the trees (...)

Sin embargo, en todos los ejemplos de adición se aprecia una dificultad para llevar a cabo la acción que logra realizarse o no, y de ahí proviene la necesidad del traductor de expresar ese matiz a través de la perífrasis que en español tenemos disponible y, al parecer, bastante presente en el repertorio de recursos de traducción, pues no se precisa de un elemento activador en inglés para utilizarla, al contrario de lo que habíamos supuesto.

Por supuesto, en los tres casos de reformulación, todos de Mustieles, las expresiones utilizadas dan claros indicios de esta connotación:

<Cornwell>: Vi en la pantalla del monitor a los cuatro chiquillos que esperaban en el porche y , tras descolgar el auricular , me disponía a decirles que abría enseguida cuando <alcancé a> oír lo que estaban comentando :

-->par_en: I could see in my video monitor the four trick-or-treaters on my porch , and picking up the handset , I started to tell them that I would be right there when I overheard what they were saying to each other .

<Cornwell>: <Alcancé a> vislumbrar una mejilla pálida y una nariz aquilina antes de que el espacio volviera a quedar vacío .

-->par_en: I caught a flash of a pale cheek and aquiline nose before the space was empty again .

<Cornwell>: En el piso de arriba sonó una voz fuerte , una protesta amortiguada que no <alcancé a> entender .

-->par_en: A loud voice sounded from upstairs , a muffled protest **could not make out** A chair scraped across the floor .

Elemento activador EN	Frecuencia	Porcentaje
∅	14	29,2
Adverbios temporales (ever, never)	13	27,1
Reformulación	9	18,7
Get/come to	6	12,5
Intensificadores (at all, really, rather, even)	4	8,3
As far as	1	2,1
About	1	2,1
TOTAL	48	100

Tabla 31 - Equivalentes de «llegar a + INF» en EN

Con respecto a «llegar a + infinitivo», un 27,1 % surge de frases en las que se incluyen los adverbios temporales de frecuencia *ever* y *never*.

<Highsmith>: Y averigüad si Cynthia ha <llegado a> conocerle .

-->par_en: And-did Cynthia **ever** meet him ?

<Ellroy>: Y yo no <llegué a> conocer en persona a la chica , ni nadie me volvió a hablar de ella nunca más.

-->par_en: **And I never met that** girl myself , and nobody else ever brought her up to me

El resto de elementos activadores tienen poca representación: con 3 casos cada uno tenemos a *get to* y *come to*.

<Leonard>: De nuevo tiró de él , pero esta vez no <llegó a> sonar el percutor .

-->par_en: Racked it again , but **didn't get to** click it this time

<Leonard>: ¿Has <llegado a> usarla ?

-->par_en: " You **get to** use it ?

De cada uno de los demás elementos solo hallamos un caso:

<Cornwell>: El caso contra Norring llevaría mucho tiempo , si es que alguna vez <llegaba a> probarse , y me pareció comprender , hasta cierto punto , lo que debía de haber sentido Waddell .

-->par_en: The case against Norring would take a very long time , if it could be proven **at all** , and I thought I understood , to a degree , how Waddell must have felt

<Cornwell>: Su verdadero objetivo , casi había <llegado a> creérmelo , consistía en impugnar mi competencia profesional y , en general , hacerme sentir como una estúpida .

-->par_en: His true agenda , I was **about to** believe , was to impeach my professional competence and in general make me feel stupid .

<Highsmith>: Tom se detuvo , con la horrible certeza de que los Pritchard le estaban investigando sólo a él , y que habían <llegado a> enterarse incluso de lo de Dickie Greenleaf .

-->par_en: Tom stopped , horribly aware that Pritchard or the Pritchards had been boning up on him , and as far back as Dickie Greenleaf

Casi un 48 % corresponde a adiciones y reformulaciones. En estas últimas, 3 de 8 son adjetivos que comienzan con el prefijo *un-* (*unfinished, undisturbed y unheard*).

<Highsmith>: En el momento de descubrir los cadáveres , decía el periódico , la casa estaba iluminada , con las puertas abiertas y en la sala había bebidas que apenas habían <llegado a> probar .

-->par_en: The house lights had been on , door open , and there were unfinished drinks in the living room , when the two corpses had been discovered .

<Cornwell>: Le he dicho muy claramente , aunque de un modo sutil , por supuesto , que sería una verdadera lástima que <llegara a> divulgarse algo acerca del maletín robado de casa de Robyn Naismith .

-->par_en: " I 've been clear - in a subtle way , of course - that it would be a damn shame if something got leaked about the briefcase stolen from Robyn Naismith 's house .

<Ellroy>: - No creo que yo <llegara a> conocerla .

-->par_en: " I do n't think I knew her .

No puede decirse que exista un patrón o elemento activador claro en la traducción de estas dos perífrasis, motivo por el cual depende de la variedad del repertorio de recursos del traductor el utilizarlas o no. Vemos que en la anterior versión de TRAD apenas se utilizaba «alcanzar a + inf». Sin embargo, esta vez, nos hemos encontrado con que hay al menos un caso en cuatro de los textos y que el traductor de Cornwell ha elegido utilizarla más de lo que observamos habitual en el subcorpus de español original, y sin ningún estímulo textual aparente, por lo que quizás podríamos apuntar a un cierto grado de afectación en el lenguaje o literariedad intencionada por parte del mismo, que lo aleja de la norma de los originales españoles, pero no podemos hablar en este caso de interferencia.

4.4.2.5. Andar + gerundio

«Andar + gerundio» es una perífrasis con valor durativo reiterativo que suele expresar movimiento reiterado sin dirección fija o de manera irreflexiva («Anda

buscando un trabajo») o matices de ansiedad, desprecio, reprobación, etcétera. Su significado en inglés puede ser parafraseado utilizando los tiempos continuos, expresiones como *go around/about*, adverbios temporales de frecuencia como *always*, *continually* o *constantly* o con expresiones de indeterminación como *here and there*. Se utiliza sobre todo en lenguaje informal o coloquial.

A veces, como sucede con varias de las perífrasis de gerundio, pueden encontrarse entre el auxiliar y el gerundio elementos como el sujeto o un complemento circunstancial. Otras, es difícil determinar si una combinación funciona o no como perífrasis y es imprescindible recurrir al contexto, que, aun así, puede no ser del todo esclarecedor, como en el caso de «*andan mirando escaparates*», que puede significar tanto que están mirando escaparates desde hace un tiempo de manera reiterada, como simplemente que cuando andan miran escaparates.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: *ANDA*/ANDE*/ANDO/ANDÁ*/ANDUV*/ANDÉ**

Context words: **ANDO*/*IENDO*/*ÁNDO*/*IÉNDO*/*YENDO*/*YÉNDO**

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Andar + GER	35	10,8	24	4,6	+10,52	< 0,01	SÍ

Existe diferencia entre el número de apariciones en originales y traducciones, y la normalización de las cifras da como resultado 6,2 apariciones más por cien mil palabras en el subcorpus de originales en español. El resultado es el esperado, dado que el inglés carece de correspondencia directa y se vale de adverbios que cuentan con equivalentes más literales en español para expresar los mismos matices. La utilización de esta perífrasis en el caso de las traducciones parece ser una cuestión de preferencia personal o idiolecto del traductor, ya que todas las apariciones se encuentran, una vez más, en las novelas de Hunter y Mosley, menos un único caso en la de Ellroy.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?(a|á)nd(a|á|e|é).*(¿|i)?ando|(¿|i)?anduv.*" %c] []{0,5}
 [word=".*(a|á)ndo.*.*(i|y)(e|é)ndo.*"]

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Andar + GER	35	9,5	26	5,3	+5,15	< 0,05	SÍ

Volvemos a encontrar una infrarrepresentación significativa en TRAD. Esta perífrasis aparece acompañada de la expresión deíctica «por ahí» en numerosas ocasiones, y, en el subcorpus de traducciones, a veces se encuentra intercalada entre auxiliar y verbo principal. Como ocurría con «seguir + gerundio», hemos tenido muchas dudas a la hora de tener en cuenta estos casos o no:

<Ellroy>: Unidad 3-B71 : - Lucille <anda por Chinatown haciendo> preguntas .

-->par_en: Unit 3-B71 : " Lucille 's walking around Chinatown asking questions .

<Ellroy>: Mierda , está más sucio de lo que nadie podría.. . ¡mierda , está inyectándose droga y <anda por ahí extorsionando> a los vendedores !

-->par_en: " Fuck ... he 's dirty past your wildest ... luck , he 's shooting dope , he 's shaking down pushers .

<Ellroy>: Así pues , ese chalado maricón de Junior <andaba por ahí proclamando> su rollo de que iba a hacerse el amo del barrio negro , y estoy seguro de que Dudley se lo habría cargado personalmente de no haber sido por esa sobredosis , fuera cosa suya o de los Kafesjian .

-->par_en: So , this nutty faggot Junior , he 's talking up his own Niggertown takeover stuff , and in my opinion Dud would have clipped him himself , if he had n't of OD 'd or got snuffed by the Kafesjians . "

Es verdad que encontramos esta forma deíctica en los ejemplos de los distintos autores que tratan las perífrasis de gerundio, pero siempre aparece detrás de la perífrasis (Yllera 1999: 3418). De hecho, lo mismo sucede en el subcorpus de originales españoles:

<Giménez>: Nadie <anda marcando> chicas por ahí .

<Martín>: José <andaré diciendo> por ahí que Soto ya sabía que había tres muertas antes de que apareciera el tercer cadáver .

Gómez Torrego (1988: 147) se vale de los dos siguientes ejemplos para explicar por qué cree que no estaríamos ante verdaderas perífrasis en los casos reproducidos más arriba. En el caso de «Andan diciendo por ahí que ha dimitido el presidente», Gómez Torrego argumenta que el sintagma preposicional «por ahí» complementa a toda la

perífrasis y no solo al auxiliar (se focaliza el verbo *decir*, como sería el caso en «dicen por ahí»); por lo tanto, se trata de una perífrasis verbal. Sin embargo, del mismo ejemplo pero con el sintagma preposicional intercalado entre verbo auxiliar y gerundio, es decir «Andan por ahí diciendo que ha dimitido el presidente», opina lo contrario, pues considera que el complemento circunstancial incide sobre el primer verbo, debilitando el carácter perifrástico. Estilísticamente parece destacarse más la idea de desplazamiento y se focalizan tanto *decir* como *andar*.

En el primer ejemplo, la expresión que motiva la supuesta perífrasis es «*walking around Chinatown asking questions*». Vemos que en inglés se focalizan efectivamente dos cosas distintas, como habíamos pensado en español: el andar por Chinatown y el hacer preguntas, por lo que no nos parece que se pueda hablar de perífrasis en este caso. En los otros dos casos el «por ahí» es adición del traductor, ya que en inglés solo se utiliza el tiempo continuo, sin complementos.

Si decidiéramos tener en cuenta estos casos dudosos, las diferencias entre subcorpus dejarían de ser estadísticamente significativas (35 casos en ES y 29 en TRAD; LL 3,58). Pero, así como no aceptamos los casos dudosos en la perífrasis «seguir + gerundio», tampoco lo hacemos esta vez, basándonos una vez más en las explicaciones de Gómez Torrego —que, de los lingüistas consultados, es el único que profundiza tanto en este tipo de detalles—. La decisión de no contar estos ejemplos hace que se confirme la infrarrepresentación.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	16	18,8	Ellroy	13	11,8
Martín	12	21,3	Cornwell	10	9,1
Madrid	5	12,1	Grafton	2	2,1
Martínez	2	3,7	Leonard	1	1,3
Núñez	0	0	Highsmith	0	0
Vázquez	0	0			
Total	35	9,5		26	5,3

Tabla 32 - Distribución de «andar + GER» por texto

Ahora que todos los traductores del corpus son españoles, encontramos todavía más apariciones de esta perífrasis, lo que contradice nuestra idea inicial de que determinadas perífrasis eran de uso más habitual en la variante argentina. Su uso podría depender de factores de interferencia, lo que comprobaremos consultando el corpus paralelo, o del idiolecto de determinados traductores.

En ES, Vázquez Montalbán y Núñez no la utilizan ni una sola vez, mientras que Giménez y Martín abarcan un 46 y un 34 por ciento de las concordancias respectivamente. En TRAD, Highsmith, que encabeza varias de las listas de distribución de perífrasis, no cuenta ni con un solo caso, mientras que el traductor de Ellroy, que suele estar a la cola en la utilización de las expresiones estudiadas por la falta intencionada de verbos en su texto, sorprende al colocarse en el primer puesto. Una posible explicación es que esta perífrasis es capaz de aportar ese toque de coloquialidad que va muy bien al registro general de la novela de Ellroy, debido a la gran cantidad de texto dialogado que incluye y a los ambientes marginales en los que transcurre la acción.

<Ellroy>: Le ha contado que Steve le había vendido a Junior esa buena cantidad y que el policía <andaba proclamando> esa chifladura de que será el próximo rey de la droga .

-->par_en: He said that Steve sold Junior this big stash , and that the Junior policeman was talking up all this dope-kingpin jive .

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Tiempos continuos	23	88,5
∅	2	7,7
Keep + -ing	1	3,8
TOTAL	26	100

Tabla 33 - Equivalentes de «andar + GER» en EN

Entre las concordancias hallamos dos casos de adición y uno del verbo *keep* seguido de un participio de presente:

<Cornwell>: - ¿Qué <anda buscando> ? - pregunté al fin .

-->par_en: " What are you after ? " I finally asked .

<Ellroy>: Un tipo con un gran tupé que <anda tocándose> todo el rato .

-->par_en: He had this big pom-po-dour that he kept playin ' with . "

Sin embargo, prácticamente todas las frases originales se encuentran conjugadas en tiempos continuos simples o compuestos, ya sea en *present continuous* (8 casos):

<Ellroy>: Y la policía local de Malibú también ha estado aquí tras los pasos de mi ex estrella , Glenda Bledsoe , con la que tengo entendido que <andas jugando> a romances .

-->par_en: The Malibu Sheriff 's were looking for my erstwhile star Glenda Bledsoe , who I understand **you are playing** Bury the Brisket with

Past continuous (6 casos):

<Cornwell>: - ¿Cree que esa persona encontró lo que <andaba buscando> ?

-->par_en: " Do you think this person found **what he was looking for** ? "

O present perfect continuous (9 casos):

<Ellroy>: Señor Klein , he oído que los muchachos del señor Smith me <andan buscando> porque les ha llegado el absurdo rumor de que fui yo quien se cargó al jodido Wardell , y me huelo que la han tomado conmigo para engrosar su apresurada estadística .

-->par_en: Mr. Klein , I heard them Mr. Smith boys **been lookin ' for me** on account of some bogus rumor that I snuffed fuckin ' Wardell , and it looks to me like they be measurin ' me for a quick statistic .

A veces se añaden adverbios temporales de duración (*always, since y still*) para dotarlas de un cierto sentido de hartazgo o fastidio:

<Leonard>: Aún lo <andaba buscando> , pues ya llevaba cinco días con el Rolex falso que no era feo y marcaba bien la hora , pero no dejaba de costar sólo unos cientos de pavos .

-->par_en: **Still trying to** find him , five days now with the fake Rolex that wasn't bad-looking , kept the right time , but still wasn't worth a thousand bucks .

<Ellroy>: Siempre <anda hablando> con ese tipo , Wylie Bullock , el cámara , que al menos tiene un lugar donde dormir , y no como la mayoría de los vagabundos que ha contratado Mickey .

-->par_en: He **'s always talking** to that guy Wylie Bullock , the cameraman , who at least has got a place to live , unlike most of the bums Mickey 's hired .

<Cornwell> - Un puñado de lunáticos religiosos <anda merodeando> por la calle Spring desde esta mañana

-->par_en: " A bunch of religious nuts **have been hanging out** at Spring Street **since** this morning .

En todas ellas el sentido general indica una acción que se sigue repitiendo a lo largo de un período de tiempo, con la utilización de verbos como *looking for*, sobre todo (11), pero también *chasing, talking up* o *trying*. En todas, asimismo, podría haberse evitado la utilización de «andar + gerundio» traduciendo simplemente por el verbo «estar» conjugado y otro verbo en gerundio (lo que conformaría la perífrasis «estar +

gerundio»¹³²), o a través de complementos circunstanciales que expresaran la reiteración de la acción, y el traductor ha optado por ir un paso más allá.

4.4.2.6. Ir + gerundio

Esta perífrasis de gerundio expresa una acción en curso que se desarrolla gradualmente hacia un momento posterior y puede presentar diversos valores aspectuales durativos: reiterativo, cuando expresa la repetición de una acción sin dirección fija, como en «Iba contando mentiras a todo el mundo» (*He was going around telling lies to everyone*); progresivo, cuando la acción se desarrolla de manera gradual, como por ejemplo en la frase «Los precios van aumentando año tras año» (*Prices keep increasing every year*); o incoativo, cuando expresa el inicio de una acción que se prolonga en el tiempo, como vemos en «Ya va entendiendo» (*She is beginning to understand*). «Ir tirando» es locución y no la tendremos en cuenta.

Para reproducir estos valores, el inglés se vale de los tiempos continuos con adverbios y expresiones de progresión temporal como *little by little*, *gradually*, *as* o «*grow* + adjetivo», todas las cuales pueden ser traducidas mediante expresiones formalmente más similares en español (p.ej. «poco a poco», «gradualmente», «paulatinamente», «volverse», etcétera.) que incluso se utilizan en este idioma para complementar a la perífrasis en algunos casos.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words:

IR/VAY*/VÁY*/IRÍ*/IRÁ*/IRÉ*/FUERE*/FUESE*/FUÉR*/FUÉS*/FUI*/IBA*/ÍBA*

IDO/VAN/VA/VE/VOY/VAS/VAIS/VAMOS/IREMOS/FUÉ*/FUE/FUERA*

Context words: *ANDO*/*IENDO*/*ÁNDO*/*IÉNDO*/*YENDO*/*YÉNDO*

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Ir + GER	127	39,1	30	5,8	+119,00	< 0,0001	SÍ

¹³² No hemos incluido esta perífrasis en el análisis ya que es la perífrasis de gerundio «neutra» por excelencia. Solo expresa valor durativo y su estructura es demasiado cercana a la del presente continuo en inglés. Es la más utilizada y algunos autores la consideran parte del paradigma verbal (Yllera 1999: 3402), por lo que generaría demasiadas concordancias y no creemos que su uso entrañe dificultades de interés para este estudio.

En el caso de esta perífrasis, la diferencia entre subcorpus es muy considerable (33,3 apariciones más por cada cien mil palabras en el subcorpus de originales). Además, las concordancias se distribuyen de manera más o menos equitativa entre los distintos autores españoles, lo que indica que no es un rasgo idiosincrático de un autor o una autora en particular.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

```
[word="(¿|i)?ido?(¿|i)?y(e|é)ndo.*(¿|i)?(i|í)rse.*(¿|i)?voy|(¿|i)?vet?e?(¿|i)?v(a|á)(y|i).*(¿|i)?v(a|á)mo.*(¿|i)?van?(¿|i)?ir(é|á|í)?.*(¿|i)?(i|í)ba.*(¿|i)?fu(e|é|i)(r|s|m)?(a|e|o|t)?.*"%c] [] [0,5] [word=".*(a|á)ndo.*(i|y)(e|é)ndo.*"]
```

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Ir + GER	148	40,3	80	16,4	+44,67	< 0,0001	SÍ

En la nueva versión del subcorpus de traducciones hallamos casi el triple de apariciones de la perífrasis que en la anterior. Sin embargo, sigue encontrándose infrarrepresentada, puesto que el número de apariciones en originales también ha aumentado —aunque no tanto— y continúa siendo muy superior.

Nuevamente hemos tenido dudas acerca de la naturaleza perifrástica de los casos en los que entre auxiliar y verbo principal se insertan adverbios de lugar y complementos diversos, pero consideramos este caso idéntico al de «andar + gerundio», por lo tanto no tuvimos en cuenta los casos dudosos:

<Martín>: ¿Has oído hablar de ese Hombre de la Navaja que <va por ahí cargándose> putas ?

<Núñez>: Y una señora no <va por ahí soltando> basteces .

<Ellroy>: Lester , sin aliento : - <Van por ahí intentando> resolver muertes de morenos a manos de otros morenos .

-->par_en: Breathless : " They goin ' aroun ' trying ' to solve colored-on-colored killins , word is to forestall all this potential good Federal investigation publicity .

De haberlos tenido en cuenta, las cifras no habrían alterado los resultados de las perífrasis ni el cómputo total del grupo (150 en ES, 81 en TRAD; LL +45,36, p < 0,0001).

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Vázquez	37	42,1	Cornwell	35	31,8
Martínez	31	57	Leonard	19	24,4
Giménez	23	27,1	Grafton	11	11,8
Martín	22	39,1	Highsmith	10	10,2
Madrid	18	43,4	Ellroy	5	4,5
Núñez	17	40			
Total	148	40,3		80	16,4

Tabla 34 - Distribución de «ir + GER» por texto

En cuanto a la distribución por autores, en ES todos están en torno a una frecuencia de 40 apariciones por cien mil palabras, aunque despunta un poco Martínez Reverte y se queda algo por debajo Giménez Bartlett. No sucede lo mismo en TRAD, donde Cornwell abarca casi el 44 % de las concordancias. Es interesante que el segundo puesto por número de apariciones pertenezca a Leonard, cuyo texto es el que menos palabras contiene.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Tiempos continuos e <i>-ing participle</i>	18	22,5
Subordinadas con <i>as / while</i>	9	11,2
Verbo + <i>through</i>	7	8,7
∅	7	8,7
Reformulación	7	8,7
Secuencias (each + sust; sust + <i>after/and</i> + sust)	7	8,7
Link verbs (<i>get/become/grow</i> + ADJ)	6	7,5
Verbo + <i>around</i>	6	7,5
Marcador temporal (frecuencia y duración)	6	7,5
Gerundio	5	6,2
Verbo (<i>tick, go</i>) + <i>by</i>	2	2,5
TOTAL	80	100

Tabla 35 - Equivalentes de «ir + GER» en EN

Los elementos de las distintas categorías encontradas en el subcorpus de originales en inglés se combinan mucho más frecuentemente que en otros casos para dar lugar a la utilización de esta perífrasis durativa. Por ejemplo, tenemos concordancias en las que las conjunciones *while* y *as*, que introducen una oración temporal subordinada, se combinan con varios de los elementos de otras categorías, como la de los verbos de enlace *become* y *get* (dos casos), la de «verbo + *through*» (dos casos) o con secuencias (dos casos):

<Cornwell>: A medida que la sangre envejece , se <va oxidando> cada vez más .

-->par_en: As blood ages , it becomes more strongly oxidized . "

<Cornwell>: Debajo había un sobre grande lleno de fotografías que Marino <fue pasando> una por una mientras yo miraba en silencio .

-->par_en: Beneath it was a large envelope of photographs that Marino went through while I silently looked on .

<Highsmith>: Mientras Nick <iba pasándolos> uno tras otro , todos con su funda plástica , Tom se dio cuenta de que no podía distinguir los Derwatts de los Bernard Tufts , por lo menos no con toda seguridad .

-->par_en: As Nick lifted one after the other , all in their plastic , Tom realized that he could not tell the Derwatts from the Bernard Tufts , not with total confidence , anyway .

En estos casos, para poder clasificar la concordancia en alguna de nuestras categorías, hemos tenido que decidir cuál de los elementos aportaba en mayor medida a la frase el valor durativo, progresivo o reiterativo expresado por la perífrasis «ir + gerundio». Dado que las conjunciones eran más numerosas y aparecían combinadas con más categorías, decidimos dejar estos seis ejemplos como casos pertenecientes a estas otras categorías. Es decir que, en realidad, podríamos haber contado catorce casos de *as* y *while* en lugar de nueve.

Lo mismo podemos decir de la combinación de *get*, *become* y *grow* con marcadores temporales. Podríamos haber contado siete casos de esta última categoría en lugar de seis pero privilegiamos la clasificación en la anterior:

<Cornwell>: Con el paso del tiempo , los matices y sutilezas de interminables análisis sólo <fueron haciéndose> más complicados .

-->par_en: The nuances and twists in the endless analyses got only more complicated with time .

Aunque todos estos elementos aparecen sobre todo cuando se utilizan tiempos verbales simples como el presente o el pasado, también encontramos combinaciones con la categoría de tiempos continuos y participios presentes. Esto ocurre en siete casos que decidimos incluir en las otras categorías con las que se combinan, ya que consideramos que en este caso un tiempo verbal o un participio presente son elementos más generales que el resto de categorías, y estas últimas quizás remarquen más la idea de que la acción es continuada en el tiempo e influyan en mayor grado sobre el traductor para tomar la decisión de utilizar la perífrasis durativa.

<Grafton>: Acerqué la silla del escritorio a la caja de cartón , tomé asiento y <fui sacando> libros , sujetándolos por el lomo mientras pasaba las páginas .

-->par_en: I dragged the desk chair closer to the box and sat down , pulling out **book after book** , holding **each** by the spine while I riffled through the pages .

<Cornwell>: Y al final cogimos al cabrón porque <iba circulando> por Carolina del Norte con una pegatina de inspección técnica caducada -

-->par_en: Then we catch the bastard because he 's driving **around** North Carolina with an expired inspection sticker . "

<Leonard>: Tyler tenía ante sí la bolsa abierta e <iba tanteando> entre sus enseres - una blusa estampada , una falda de uniforme - hasta que sacó un sobre marrón hinchado , de nueve por doce pulgadas .

-->par_en: Tyler had the bag open now and **was feeling through** her things , a soiled blouse , uniform skirt , bringing out a manila envelope , a fat one , nine by twelve .

Tres de los casos de reformulación corresponden al verbo *fade* y el adjetivo *faded*, que se traducen por la perífrasis progresiva debido a su significado de algo que pierde color o se desvanece gradualmente:

<Grafton>: El estilo de la vivienda era difícil de clasificar , una típica casa californiana de una sola planta , con las paredes desconchadas y un tejado de tejas rojas que habían <ido desapareciendo> con el tiempo .

-->par_en: The style of the house was hard to classify , a vintage California cottage with a flaking stucco exterior and a **faded** red-tile roof .

<Cornwell>: Eché a andar con paso enérgico y seguro , pues había estado allí el suficiente número de veces para conocer el camino , y el parloteo de las agudas voces infantiles se <fue difuminando> como si se lo llevara el viento .

-->par_en: The babble of shrill young voices **faded** as if carried away from me by the wind , my step brisk and directed , for I had been here enough times to know the way .

Las adiciones espontáneas del traductor son menos numerosas, proporcionalmente, que en otras perífrasis, pero aun así ocupan, junto con las reformulaciones, más del 17 %. Los elementos que activan la perífrasis en traducciones son numerosos pero no tienen una forma o estructura clara; entre ellos encontramos tiempos continuos, adverbios de frecuencia y duración, conjunciones y, en general, indicios aspectuales en la frase que hacen pensar que esta describe una acción que se desarrolla a lo largo del tiempo. Al no contar con un elemento equivalente claro y diferenciado en inglés, la utilización por parte de los traductores es menor que en lenguaje espontáneo. Es probable que, en su lugar, se utilice el pasado simple o imperfecto junto a expresiones temporales más cercanas al inglés. Por ejemplo, si

buscamos los otros casos en los que *gradually* no se traduce por la perífrasis en cuestión, obtenemos las tres siguientes concordancias:

<Cornwell>: We sat like this for an hour , words <gradually> getting sharper , one shade of gray contrasting with another , molecule by molecule , bit by bit .

-->par_trad: Seguimos así durante una hora , viendo cómo las palabras se hacían cada vez más nítidas , un tono de gris contrastado con otro , milímetro a milímetro , molécula a molécula .

<Highsmith>: Only <gradually> did they return to the subject of tomorrow , the next hotel .

-->par_trad: Sólo gradualmente volvieron al tema del día siguiente , del cambio de hotel .

<Leonard>: Max rang the bell and waited , hearing the faint sound of music he <gradually> identified as vintage Mo-town , the sound familiar , but couldn't name the vocal group or the number they were doing .

-->par_trad: Max llamó al timbre y esperó , oyendo el débil sonido de la música que poco a poco identificó como de la cosecha Motown , un sonido familiar , aunque no logró adivinar qué grupo era ni qué cantaban .

4.4.2.7. Venir + gerundio

«Venir + gerundio» es una perífrasis durativa progresiva que puede presentar valor de aproximación (Viene costando lo mismo: «*It amounts to more or less the same*»), como «venir a + infinitivo», o frecuentativo (Hemos venido reivindicando nuestros derechos durante dos décadas: *We have been fighting for our rights for two decades*). En este caso, se diferencia de «ir + gerundio» porque expresa una acción en curso que parte de un tiempo anterior y se desarrolla gradualmente hasta el momento expresado por el auxiliar.

Este valor semántico se expresa en inglés por medio de la combinación de presente o pasado continuo y expresiones temporales. En estos casos, la perífrasis española no es siempre la opción de traducción más evidente.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words:

VENI*/VENÍ*/VINI*/VENID*/VEN/VENG*/VÉNG*/VIENE*/VINE/VINO/VENDR*

Context words: *ANDO*/*IENDO*/*ÁNDO*/*IÉNDO*/*YENDO*/*YÉNDO*

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Venir + GER	0	0	5	1	-4,85	< 0,05 ¹³³	NO

Este es un caso curioso porque, a pesar de ser una perífrasis teóricamente característica del español (aunque infrecuente), no aparece en el subcorpus de originales, mientras que en el de traducciones se dan cinco casos. Es esta una cifra no muy representativa en un corpus de más de 500 000 palabras, lo que nos sugiere que sería conveniente estudiarla en uno de mayores dimensiones. Por segunda vez, el test de Fisher invalida el resultado del test de log-likelihood.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?ven(¿|i)?v(e|é)n(i|í|g).*(¿|i)?vendr.*(¿|i)?vini(e|é|s|m).*(¿|i)?vin(e|o)|(¿|i)?viene.*" %c & pos="V.*" []{0,5} [word=".*(a|á)ndo.*.*(i|y)(e|é)ndo.*"];

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Venir + GER	1	0,3	7	1,4	-3,51	n. s. ¹³⁴	NO

La nueva versión del corpus deja claro que la frecuencia de esta perífrasis en las traducciones de la fase I no fue producto de la casualidad. Cuatro de los cinco textos de TRAD han cambiado y el número de palabras del subcorpus ha disminuido y, sin embargo, el número de apariciones aumenta de cinco a siete. La novela que añadimos al subcorpus ES aporta el único caso hallado de esta perífrasis en originales españoles.

<Núñez>: Pero cuando me quedaban unos pocos metros para doblar la esquina de la calle sentí que me <venían siguiendo> .

Es probable que estuviéramos equivocados al pensar que la perífrasis no era la opción evidente al traducir la combinación de los tiempos verbales *past* y *present perfect continuous* + expresiones temporales. Fente Gómez (1971: 136), de hecho, afirma: «Es quizá ésta una de las perífrasis con traducción más fija y permanente en

¹³³ Test exacto de Fisher: 0,164013 (n. s.)

¹³⁴ Test exacto de Fisher: 0,149489 (n. s.)

inglés». Puede que suceda lo mismo en la dirección inversa. De esta manera, parece producirse una representación de la perífrasis en traducciones mayor de lo habitual si tomamos como referencia los textos escritos en español originalmente, o como mínimo similar, si tenemos en cuenta que los resultados de los tests estadísticos no respaldan la hipótesis de que existen diferencias entre subcorpus.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Núñez	1	2,3	Grafton	5	5,4
Giménez	0	0	Cornwell	2	1,8
Madrid	0	0	Ellroy	0	0
Martín	0	0	Highsmith	0	0
Martínez	0	0	Leonard	0	0
Vázquez	0	0			
Total	1	0,3		7	1,4

Tabla 36 - Distribución de «venir + GER» por texto

Por otro lado, es necesario tener en cuenta que solo dos traductores son responsables de todas las apariciones del subcorpus TRAD, uno con más del doble de la frecuencia del otro, lo que puede indicar una preferencia idiolectal o una mayor abundancia en el original de la estructura inglesa, cuestión que comprobaremos mediante el corpus paralelo.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Tiempos perfectos continuos + marcador temporal	4	57,1
Tiempos perfectos + marcador temporal	1	14,3
Tiempos perfectos	1	14,3
Participio presente	1	14,3
TOTAL	7	100

Tabla 37 - Equivalentes de «venir + GER» en EN

Los tiempos perfectos continuos son el elemento que detona la utilización de la perífrasis en las traducciones en mayor medida (3 casos de *present perfect continuous* y 1 de *past perfect continuous*).

<Cornwell>: Bien , parece ser que desde hace algún tiempo <viene padeciendo> problemas de salud , y ha llegado a la conclusión de que le convendría dimitir y buscar un empleo en el sector privado
 -->par_en: " Well , it seems he 's been suffering health problems of late and has decided it would be wise to resign and look for something less stressful in the private sector .

<Cornwell>: En los últimos tiempos <venía apareciendo> en la pantalla varias veces por semana

-->par_en: It had been appearing on my screen several times a week of late when the caller would hang up without leaving a message .

Sin embargo, también encontramos un caso de participio presente, uno de *present perfect* y otro de *past perfect*, por lo que, contra nuestros pronósticos, la presencia del aspecto progresivo de los dos primeros tiempos verbales no parece ser imprescindible a la hora de activar esta perífrasis en el proceso traductor.

<Grafton>: Sufrí al pensar que Laura <venía pagando> los gastos cotidianos con billetes de banco por los que un coleccionista serio pagaría una burrada .

-->par_en: It grieved me to think of Laura covering day-to-day expenses with bank notes that a serious collector would have paid dearly for .

<Grafton>: <Vengo comiendo> lo mismo cada cumpleaños durante los últimos ochenta y seis años , desde que cumplí dos - dijo - .

-->par_en: I 've had this same meal every birthday for the last eighty-six years , since I was two , " he said .

<Grafton>: Desde su punto de vista , <venían sucediéndole> cosas inexplicables .

-->par_en: From her point of view , she 'd been subjected to inexplicable occurrences .

De hecho, de más de 150 casos de «*have/has/had* (y sus respectivas contracciones) + *been* +ing» que hemos encontrado en el corpus, solo estos cuatro se han traducido por la perífrasis: las demás concordancias incluyen verbos en presente (*tiene*), pretérito imperfecto (*estaban*), pretérito perfecto (*han intentado*) o pretérito pluscuamperfecto (*había estado espionando*) y distintas perífrasis y paráfrasis.

Sin embargo, más allá de los tiempos verbales utilizados, vemos que en más del 70 % de estos siete casos hallados en el corpus hay presente una expresión temporal y en los dos casos en los que esto no sucede, el contexto nos aclara el aspecto continuado de la acción en cuestión. Si tenemos en cuenta solo las concordancias obtenidas al buscar «*have/has/had* + *been* +ing» que también incluyen algún tipo de marcador temporal, nos quedamos con 28 casos únicamente. Además de los cuatro de «venir + gerundio» (14,3 %), el resto se traduce por las perífrasis «andar + gerundio», «dejar de + infinitivo», «tener + participio», «llevar + gerundio» (4 casos), «seguir + gerundio» y «estar + gerundio» (46 % de las concordancias) y los distintos tiempos verbales.

Aunque vemos que está más presente que otras perífrasis —o igual de presente en el caso de «llevar + gerundio»— a la hora de traducir la combinación de los tiempos continuos y un marcador temporal, no podemos afirmar, con solo un 14,3 % de los 28

casos, que la estructura inglesa sea responsable de la activación de la perífrasis, por lo que su presencia en traducciones puede ser una cuestión de estilo de estos dos traductores en concreto. Sin embargo, sí podemos decir que la combinación de los tiempos continuos y un marcador temporal hace que existan más probabilidades de que los traductores se alejen de las opciones más directas y recurran a estructuras que aportan algún tipo de particularidad aspectual al verbo, como las perífrasis que acabamos de mencionar. Sin descartar la explicación idiolectal a la que parece apuntar la distribución por autores, este podría ser un caso de influencia del factor 2 de la Gravitational Pull Hypothesis (Hareide 2013: 35-40), es decir, un ejemplo de la fuerza de atracción que ejerce la estructura de los tiempos continuos, muy prominente en la lengua de partida (puesto que estos forman parte del paradigma verbal). Podría aventurarse que los traductores se encuentran de manera tan frecuente con esta estructura en el texto original que sienten la necesidad de introducir variaciones para no traducirla constantemente por la perífrasis «estar + gerundio», lo cual acaba afectando a la aparición de la perífrasis de la que hablamos y provocando una frecuencia algo mayor de la esperable en traducciones (aunque no alcance a ser significativa).

4.4.2.8. Venir a + infinitivo

A pesar de que, como perífrasis, no es una construcción muy utilizada en español, creemos interesante su estudio porque el inglés no parece tener procedimientos similares para expresar el valor perfectivo, de culminación o logro, contenido, por ejemplo, en la frase «Sus promesas vinieron a tranquilizarla» (*He managed to calm her down with his promises*) y de aproximación o falta de precisión, como en «Todo junto viene a costar unos 10 euros» (*It all amounts to about 10 euros*) o en «Vino a decir eso» (*He more or less said so*). A veces coincide en sus valores con los de la perífrasis «venir + gerundio» (Yllera 1999: 3394).

Como vemos, en inglés se suele utilizar muy diversas expresiones para marcar dichos valores (*come to, amount to, around, about, end up, in the end, finally*, etcétera.) que pueden traducirse al español de muy diversas maneras (p. ej. *aproximadamente* o *al final*). Consideraremos locución las expresiones «venir a parar» y «venga a + infinitivo».

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words:

VENI*/VENÍ*/VINI*/VENID*/VEN/VENG*/VÉNG*/VIENE*/VINE/VINO/VENDR*

Context words: A

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Venir a + INF	1	0,3	0	0	+1,91	n. s. ¹³⁵	NO

A pesar de ser una construcción perífrástica sin duda típica del español, ya habíamos apuntado antes que no es en absoluto habitual. Para obtener resultados significativos, habría que contar con un corpus de dimensiones bastante mayores. Aun así, encontramos un ejemplo de valor aproximativo en el corpus de originales en español, mientras que en el de traducciones no se encuentra ningún caso aproximativo ni perfectivo.

<Martín>: asesinó a treinta y tres jovencitos con quienes había mantenido relaciones homosexuales , o sea , jovencitos que <vendrían a ser> la parte débil , vulnerable , femenina , de la pareja , digo : si prescindimos de este ejemplo por confuso , el porcentaje se nos eleva a un 77,5 % de víctimas mujeres.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?ven|(¿|i)?v(e|é)n(i|í|g).*(¿|i)?vendr.*(¿|i)?vini(e|é|s|m).*(¿|i)?vin(e|o)|(¿|i)?viene.*" %c & pos="V.*" & lem!="ver" & lem!="vengar"] []{0,5} "a"

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Venir a + INF	1	0,3	6	1,2	-2,68	n. s. ¹³⁶	NO

La modificación de los textos traducidos ha cambiado mucho las cosas. La incorporación del texto de Cornwell y el de Grafton ha provocado que la diferencia entre subcorpus sea ahora a la inversa: hay bastantes más apariciones de esta perífrasis

¹³⁵ Test exacto de Fisher: 0,2621 (n. s.)

¹³⁶ Test exacto de Fisher: 0,25095 (n. s.)

en traducciones que en originales. Todos los casos son aproximativos, aunque el valor del siguiente es dudoso:

<Cornwell>: -En otras palabras , <viene a> confirmar lo que ha dicho la prensa .
 -->par_en: In other words , he 's reinforcing what 's been in the news . "

Sin embargo, las diferencias entre subcorpus siguen sin ser significativas estadísticamente. Podría pensarse que el motivo es que estamos manejando cifras muy pequeñas, pero no sucede lo mismo cuando la cifra menor está en la casilla del subcorpus de traducciones, como hemos visto en el caso de «terminar por + infinitivo» y «acertar a + infinitivo».

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	1	1,8	Cornwell	5	4,5
Giménez	0	0	Grafton	1	1,1
Madrid	0	0	Ellroy	0	0
Martínez	0	0	Highsmith	0	0
Núñez	0	0	Leonard	0	0
Vázquez	0	0			
Total	1	0,3		6	1,2

Tabla 38 - Distribución de «venir a + INF» por texto

La utilización de esta perífrasis sería, entonces, exagerada en traducciones, dado que es un recurso bastante menos utilizado en el lenguaje de las novelas españolas actuales, según los datos de nuestro subcorpus de originales. La diferencia entre las dos fases del corpus demuestra una de las debilidades de la metodología de corpus: depende mucho del tamaño del corpus y de los textos concretos seleccionados, por lo que las generalizaciones son arriesgadas. Cinco de los seis casos se encuentran en la traducción de Mustieles. Si la perífrasis hubiera pasado el test de significación estadística, habríamos comprobado cómo una preferencia estilística o hábito particular de un autor o traductor puede afectar a los resultados de nuestra investigación.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Ø	4	66,7
Reformulación	1	16,7
About	1	16,7
TOTAL	6	100

Tabla 39 - Equivalentes de «venir a + INF» en EN

De todos los posibles equivalentes en inglés que mencionábamos en la introducción, el único que hemos encontrado es «about»:

<Cornwell>: <Viene a> ser el doble de grande que un sobre normal y está hecho de un papel negro duro y brillante .

-->par_en: " It 's about twice as wide as a legal-size envelope of stiff , shiny black paper .

Los casos son muy pocos para extraer conclusiones válidas, pero en los cinco casos restantes, la utilización de la perífrasis no es imprescindible, y no parece estar motivada por factores textuales sino más bien por preferencias del traductor, puesto que no hay nada en el texto que impida la omisión de la perífrasis. Tenemos que suponer, entonces, que son razones estilísticas o cuestiones idiosincráticas de cada traductor las que lo han llevado a utilizarla.

<Cornwell>: Las barbas son los filamentos que confieren a la pluma su apariencia « plumosa » , y cuando las vea al microscopio comprobará que en realidad <vienen a> ser como unas mini plumas que crecen en el astil -.

-->par_en: The barbs are the filaments that result in the feather 's feathery appearance , and when they 're magnified you 'll find they 're actually like minifeathers coming out of the shaft . "

4.4.2.9. Recapitulación de resultados: grupo II

- CORPUS COMPARABLE – FASE I

Subtotal Grupo II	429	132,2	213	40,9	+211,69	< 0,0001	SÍ
------------------------------	------------	-------	------------	------	----------------	--------------------	-----------

En este segundo grupo, la mayor parte de las nueve perífrasis, que en principio no contaban con equivalentes formales evidentes en inglés, confirmaron la hipótesis de la infrarrepresentación en traducciones, excepto en tres casos («venir a + infinitivo», «alcanzar a + infinitivo» y «venir + gerundio»). Este hecho no nos resultó excesivamente sorprendente, puesto que dichas perífrasis son muy poco habituales en el español escrito estándar de hoy en día, lo que viene a ser confirmado por el hecho de que su frecuencia de aparición en nuestro subcorpus de originales españoles es muy

baja. De hecho, «venir + gerundio» no aparece ni una vez. Sin embargo, resulta curioso que sí aparezca cinco veces en traducciones, aunque es necesario tener en cuenta que este subcorpus es bastante más grande que el de originales. El resultado general de este grupo muestra, como esperábamos, una clara infrarrepresentación estadísticamente significativa en el subcorpus de traducciones.

- CORPUS COMPARABLE – FASE II

Subtotal Grupo II	481	131	383	78,4	+56,74	< 0,0001	SÍ
------------------------------	------------	-----	------------	------	---------------	--------------------	-----------

El resultado del grupo de perífrasis tomado en su totalidad en esta segunda fase no ha cambiado; sigue estando significativamente infrarrepresentado en traducciones, aunque con menor diferencia que en la fase anterior del corpus. Algunos de los resultados individuales han variado notablemente, sin embargo. Ha disminuido el número de casos no significativos de tres a dos («venir a + infinitivo» y «venir + gerundio»), pero en dos perífrasis hallamos el caso contrario al que predecíamos: se encuentran sobrerrepresentadas en traducciones («haber de + infinitivo» y «alcanzar a + infinitivo»). Por lo tanto, de seis perífrasis claramente infrarrepresentadas en la fase I, pasamos a tener cinco. En el caso de «haber de + infinitivo», pasamos sorpresivamente de un rotundo *sí* ($p < 0,0001$) a un casi igual de rotundo *no* ($p < 0,01$). El corpus paralelo no nos permite saber con certeza la razón de este fenómeno, pero sospechamos que en este caso sí que influye la variedad lingüística utilizada por los traductores, aunque en el resto de perífrasis no hayamos comprobado que esta variable afecte a los resultados.

- CORPUS PARALELO

Las perífrasis de este grupo son de naturaleza muy diversa, y eso se refleja también en el corpus paralelo. De seis perífrasis de infinitivo, tres están infrarrepresentadas, «haber que + infinitivo», «llegar a + infinitivo» y «acertar a + infinitivo», pero esta última no cuenta con concordancias paralelas, puesto que no aparece en el subcorpus de traducciones, por lo que nos referiremos a las dos primeras. Estas dos perífrasis no cuentan con elementos detonantes en inglés claramente marcados, pues las estructuras que desencadenan su utilización son variadas y un alto

porcentaje de concordancias (40 % y 48 % respectivamente) corresponde a adiciones y reformulaciones; es decir, a un uso espontáneo por parte del traductor.

De las tres perífrasis de gerundio, se encuentran infrarrepresentadas dos, «andar + gerundio» e «ir + gerundio». En su caso, las adiciones y reformulaciones son menos —un 7,7 % y 17,4 % respectivamente—, en el primer caso debido a que hemos considerado necesario tener en cuenta determinados tiempos verbales como elementos activadores, lo que no había sucedido en el caso de las perífrasis de infinitivo, y en el segundo, porque no hay una o dos estructuras equivalentes que motiven su uso, sino que la perífrasis se elige para traducir muchas expresiones diferentes, que hemos podido agrupar en numerosas y diversas categorías. Vuelve a suceder lo mismo: en gran medida queda al arbitrio del traductor la interpretación del sentido y las connotaciones de la frase y la decisión de si utilizar la perífrasis o valerse de las numerosas opciones que en español tenemos disponible para expresarlas.

En cuanto a las perífrasis que resultaron sobrerrepresentadas, «haber de + infinitivo» y «alcanzar a + infinitivo», nos enfrentamos a dos situaciones completamente opuestas. Es evidente que este fenómeno puede darse tanto si hay estructuras claras en la lengua de origen que puedan servir de equivalente, ya que «haber de + infinitivo» solo posee un 16,8 % de adiciones y reformulaciones y el resto de categorías poseen valores muy similares de necesidad y obligación (quizás nos equivocamos al incluirla en este grupo, o quizás la sobrerrepresentación venga dada por otros factores, como ya explicamos en el apartado correspondiente), como si el material textual no tiene nada que ver con la decisión de utilizar la perífrasis o no, como sucede con «alcanzar a + infinitivo», que cuenta con un 100 % de adiciones y reformulaciones.

Nos quedan los dos casos no significativos, «venir a + infinitivo» y «venir + gerundio». Ambas cuentan con más casos en traducciones y comparten algunos de sus valores. Aunque en una hemos clasificado el 83,4 % de las concordancias como adiciones y reformulaciones, y en la otra todas las categorías están ocupadas por distintos tiempos verbales, lo cierto es que en ambas es el contexto o razones de tipo estilístico lo que lleva al traductor a utilizarlas, puesto que en el texto no hay elementos que puedan considerarse detonantes claros.

A modo de recapitulación, podríamos decir que en este grupo es el traductor quien tiene, aún más si cabe, la última palabra. Está claro, dada la infrarrepresentación general hallada en el grupo, que estas perífrasis son estructuras que el traductor puede

tener en cuenta en determinados casos —dado que hemos visto como en numerosas ocasiones no hay nada que motive su utilización en ciertas concordancias—, pero que, en general, no se utilizan con la misma frecuencia que en textos originales. Esto parece deberse a la falta de equivalentes textuales claros en inglés; es probable que estas perífrasis se utilicen mucho menos en traducciones al español de lo que sería habitual en originales españoles, a menos que cuenten con un significado útil que pueda servir para parafrasear diversas estructuras inglesas (como «haber que + infinitivo» y «llegar a + infinitivo», que, aun así no se utilizan en la misma medida en traducciones puesto que no cuentan con equivalentes claros en inglés), o que, por algún motivo, formen parte del estilo o de los recursos de traducción preferidos de un determinado traductor (como podría ser el caso en «alcanzar a + infinitivo» o en «haber de + infinitivo»).

4.4.3. GRUPO III

Incluimos aquí las perífrasis que se encuentran en una situación intermedia entre los dos grupos anteriores, ya sea porque existen distintas alternativas en español con los mismos valores o significados, unas más evidentes que otras, y el que el traductor se decante por la perífrasis o no depende de diversas variables, o por un abanico de razones distintas, como en el caso de «soler + infinitivo», en el que el grado de similitud cambia según el tiempo verbal (pasado o presente) en que se use la perífrasis. Son las siguientes nueve perífrasis, en su mayoría, como en el primer grupo, de valor aspectual:

19. QUERER + INFINITIVO
20. PONERSE A +INFINITIVO
21. ECHAR(SE) A + INFINITIVO
22. ROMPER A + INFINITIVO
23. ESTAR A PUNTO DE + INFINITIVO
24. DEJAR DE + INFINITIVO
25. PARAR DE + INFINITIVO
26. SOLER+ INFINITIVO
27. LLEVAR + GERUNDIO

4.4.3.1. Querer + infinitivo

El valor principal de la perífrasis «querer + infinitivo» es volitivo: «Diana quiere salir a pasear» (*Diana wants to go for a walk*), y muchos autores no consideran que el verbo *querer* esté suficientemente gramaticalizado para constituir verdadera perífrasis en la mayor parte de los casos (p. ej. Fernández de Castro 1999: 229 y García Fernández 2012:10). Gómez Torrego tampoco la considera perífrasis cuando presenta dicho valor semántico y lleva un sujeto de persona. Sí cuando el sujeto es de cosa o cero, como en «Un suspiro quería escapar de su boca» (*A sigh was about to escape from his mouth*) o «Quiere llover» (*It looks as if it is going to rain*).

En estos casos creemos que, a pesar de las dudas sobre la naturaleza perifrástica de estas estructuras verbales, los valores enumerados por Gómez Torrego como válidos (1988: 102) son suficientemente interesantes por sí mismos para que valga la pena investigar la manera en la que son utilizados en originales y traducciones.

Nos interesan, concretamente, los valores potencial e incoativo que adopta en estos casos y, además, los conativos éticos y de recriminación: «Quiero verte trabajar ya» (*Start working right now*); «¡Te quieres callar de una vez!» (*Will you be quiet?*). Consideraremos locución la expresión «querer decir».

Existen numerosas opciones para trasladar estos valores al inglés, como distintas expresiones de futuro e inminencia o los modales *could*, *would*, *will*, por ejemplo, pero el uso de esta perífrasis no es completamente evidente al traducir estas opciones al español, por lo que hemos podido comprobar en diccionarios bilingües.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: *AR*/*ER*/*IR*

But excluding: decir*

Context words: QUERI*/QUIER*/QUIÉR*/QUER*/QUISO/QUISI*

Context Search Horizons: 1L to 0R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Querer + INF	22	6,8	1	0,2	+34,85	< 0,0001	SÍ

En este caso, el auxiliar fue incluido como parte del contexto, por lo que el horizonte de búsqueda se estableció en una palabra a la izquierda de la palabra clave.

Esto se hizo para poder establecer un parámetro adicional en la búsqueda de la palabra clave. Uno de los verbos más habituales que sigue a *querer*, sobre todo en traducciones del inglés, debido a la expresión sumamente habitual «sujeto + *to mean*», es *decir*. Debido a que se la considera una locución, creímos necesario excluirla de la búsqueda de infinitivos. Es difícil buscar infinitivos en un corpus no etiquetado; la expresión debe incluir un asterisco al principio y al final (para tener en cuenta los clíticos), por lo que aun así hubo que revisar a fondo todas las concordancias. Otra expresión que vimos exageradamente representada en el corpus de traducciones es «quiso saber», como acotación en los diálogos, quizás por ser más frecuente el «s/he wanted to know» en inglés.

De los muchos casos en los que aparecía esta perífrasis, conservamos sólo los enumerados en la caracterización, realizada más arriba, de esta perífrasis (potencial, incoativo, ético y de recriminación): en total, veintidós en español original y uno solo válido en las traducciones.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

`[word="(¿|i)?qu(i?)(e|é)r.*| (¿|i)?quis(o|i)e).*" %c] [word=".*(a|á|e|é|i|í)r.*" & lem!="decir"];`

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Querer + INF	26	7,1	15	3,1	+6,97	< 0,01	SÍ

Como en la fase anterior, hemos tenido problemas para la selección de casos válidos debido a la gran cantidad de concordancias obtenidas¹³⁷. Esta perífrasis se utiliza más en TRAD: 692 concordancias frente a 489 en ES, aunque la diferencia no es significativa (LL -1,08).

Además, algunas formas de *decir*, aunque en principio tendrían que haberse omitido según la expresión regular que hemos utilizado, se incluyen en las

¹³⁷ Ante la dificultad de distinguir los casos que Gómez Torrego parece aceptar como perifrásticos de los casos de personificación hallados en el corpus, hemos optado por incluir todos los valores que difieren del uso más habitual, el volitivo con sujeto de persona; es decir, los casos de personificación, de personificación claramente incoativa o de disposición, de cortesía e imperativos. De todos modos, el resultado final no se ve alterado de ninguna forma, pues en todos los supuestos se registra una diferencia significativa a favor del subcorpus de originales.

concordancias, debido con toda certeza a errores de lematización del programa TreeTagger.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Vázquez	7	8	Cornwell	4	3,6
Martín	6	10,7	Ellroy	4	3,6
Giménez	4	4,7	Grafton	4	4,3
Madrid	4	9,6	Leonard	2	2,6
Núñez	4	9,4	Highsmith	1	1
Martínez	1	1,8			
Total	26	7,1		15	3,1

Tabla 40 - Distribución de «querer + INF» por texto

Encontramos bastante diferencia con respecto a la versión antigua del subcorpus de traducciones, en el que habíamos observado un único caso, en la traducción de Núñez. La mayor parte de los casos encontrados en TRAD (trece de quince) son de tipo conativo, de recriminación o llamada a hacer o dejar de hacer algo, o petición cortés (la cortesía más marcada en el caso de utilizar el futuro o el condicional), mediante una interrogación —que podría ser también exclamación, pero no hemos hallado ningún caso de ello—:

<Grafton>: - <¿Quieres callarte> ?

-->par_en: "Would you knock it off ?

<Cornwell>: <¿Querrás hacerme> este favor para que yo pueda ocuparme de las cosas ?

-->par_en: Will you do that for me so I can take care of things ?

En ES también encontramos once casos de este tipo, pero se añaden seis casos de una modalidad diferente de apelación que no aparece en TRAD: la que no se vale de una interrogación, sino de un enunciado en primera persona que adquiere función imperativa en el contexto, un valor que Gómez Torrego denomina ético (1988:101); siempre con el verbo *ver* en los ejemplos extraídos de nuestro subcorpus.

<Vázquez>: En cualquier caso arréglate como puedas , pero en una semana <quiero ver> tu nombre borrado de todos los documentos que todavía nos unen .

En TRAD hallamos dos casos de utilización de «querer + infinitivo» con sujeto de cosa personificado, pero vemos que este uso es mucho más frecuente en español original (nueve en ES).

<Núñez>: En otras circunstancias yo te lo hubiera propuesto - y me abrazó con una energía que parecía <querer fundir> todo el miedo que yo confesaba tener .

<Núñez>: Al fin apareció la noticia . Estaba como <queriéndose ocultar> entre otras en la página de sucesos .

Estos últimos casos se asemejan mucho a ciertos ejemplos que proporciona Gómez Torrego (1988: 100, 199, 205), a los que él otorga valor estilístico descriptivo y clasifica como incoativos o con el significado de «estar a punto de + infinitivo»: «Los ojitos parecían querer mover los párpados»; «Este verano parece que quiere portarse mejor que el pasado»; «Parece como si el sol quisiera ocultarse entre las nubes».

Sin embargo, a nosotros nos parece que en todos podría considerarse que *querer* conserva su sentido volitivo original, salvo en tres ejemplos pertenecientes al subcorpus de originales, en los que parece acercarse más al de «intentar» que al de inminencia al que se refiere Gómez Torrego (y también Fernández de Castro 1999: 230).

<Madrid>: Raquelita sonreía en mi puerta con lo que <quería ser> una mueca sexy .

<Vázquez>: Los ojos del escritor <quisieron sellar> los labios de su mujer , pero ya era tarde .

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Will	7	46,7
Would	4	26,7
Want	3	20
Could	1	6,7
TOTAL	15	100

Tabla 41 – Equivalentes de «querer + INF» en EN

No ha habido sorpresas con respecto a los elementos activadores de las perífrasis con *querer* hallados en nuestro corpus. En los dos casos de utilización de «querer + infinitivo» con sujeto de cosa personificado, el inglés utilizaba *want*, por lo que el recurso retórico está presente en inglés y podría haber sido copiado tal cual, sin tener en cuenta los usos estilísticos del español. De todas formas, hemos comprobado que en español original también es un recurso utilizado y, además, en mayor medida que en traducciones.

<Highsmith>: Heloise estaba sentada junto a la ventanilla , porque Tom había insistido en ello . Y allí estaban : las dos afiladas puntas del puerto de Tánger , curvadas hacia dentro , y dirigidas hacia el estrecho como si <quisieran capturar> algo .

-->par_en: Heloise had the window seat , Tom had insisted on that , and there it was-the dramatic two prongs , curved inward , of the port of Tangier , reaching out into the Strait as if **wanting** to capture something .

En los trece ejemplos de tipo conativo, hallamos un 46,7 % de casos en los que se traduce a partir del modal *will* (*quiere(s)* en cinco concordancias y *querrás* en dos), la mayor parte dentro de una petición o ruego cortés, pero también, con *quiere(s)*, en exclamaciones cargadas de impaciencia o recriminación.

<Cornwell>: <¿Querrás ir> con Lucy a buscar el vino , dentro de un rato ? preguntó mi madre desde la ventana .

-->par_en: **Will** you and Lucy go out after a while and get the wine ? " my mother said from the window .

<Ellroy>: - <¿Quieres hacer> el jodido favor de cantar de una vez ?

-->par_en: Pete said , " **Will** you _please_ fucking give ? "

Would es el segundo detonante por frecuencia de aparición: en tres casos se traduce por *quiere(s)* y en uno por el condicional *querría*. En este último caso hablamos de una pregunta con connotaciones de urgencia y petición impaciente. En uno de los casos en los que se utiliza el presente, la petición es cortés y en los otros dos, el sentido es imperativo y de recriminación.

<Grafton>: - Sí , pero <¿querría volver> a hablar conmigo si no contesta ?

-->par_en: " Yes , but if he does n't answer , **would** you come back on the line for me ?

<Cornwell>: <¿Quiere rezar> conmigo , doctora Scarpetta ? - Extendió la mano - . Por favor .

-->par_en: **Would** you pray with me , Dr. Scarpetta ? " He held out his hand . " Please . "

<Grafton> - <¿Quieres dejar> de hacer el ganso ? - dije - .

-->par_en: " **Would** you quit being such a butt ? " I said .

Could también se utiliza para una orden o petición que implica insistencia o impaciencia:

<Grafton>: <¿Quieres hacer> lo que te he dicho ?

-->par_en: " **Could** you get busy with that tape like I said ? "

Por último, hay un único caso en el que el verbo *want* también se utiliza para formular una petición:

<Ellroy>: Jadeos fingidos : - <¿Quiere echarme> una mano con esto ?

-->par_en: Fake grunts : " You want to give me a hand with this ? "

Vemos que estos modales, más allá de los valores más o menos formales o corteses que les asignen las gramáticas a la hora de modalizar una orden o petición, pueden utilizarse para formar enunciados tanto amables como recriminatorios, y que, al traducir este tipo de frases, el elemento más utilizado en español es el verbo *querer* en presente, independientemente de si en inglés se ha utilizado *would* o *will*.

Sea como sea, esta perífrasis cuenta con elementos activadores en inglés, sobre todo en uno de sus valores, el de recriminación y petición cortés. Su menor utilización en traducciones se explica por la menor frecuencia de sujetos de cosa personificados mediante el verbo *querer* en dicha lengua, por un lado, y, por otro, la ausencia del valor ético expresado a través de la perífrasis en infinitivo con el verbo *ver*. En este último caso, en el subcorpus de traducciones encontramos frases con *querer* seguido de un objeto o de una oración de relativo con otros verbos, pero no de la perífrasis:

<Ellroy>: Brusque : " I want you to stay on the Kafesjian burglary .

-->par_trad: Brusco : - Quiero que siga con el robo Kafesjian .

4.4.3.2. Ponerse a + infinitivo

Se trata de una perífrasis incoativa o ingresiva, como en el caso de «empezar/comenzar + infinitivo», pero, según Gómez Torrego (1988: 110), aporta un matiz de subjetividad en el sujeto que expresa mayor voluntariedad, decisión o esfuerzo, como en «Se puso a estudiar», o de comienzo más repentino o enérgico en el caso de verbos impersonales, como en «Se puso a diluviar».

Junto con «echar/romper a + infinitivo», funciona como una opción más específica o estilística para traducir el valor semántico incoativo de «*to start/begin to + infinitivo*» o «*to start/begin + gerundio*». Aventurábamos en el apartado correspondiente a la perífrasis «empezar a + infinitivo» que, al ser esta una opción de

traducción más evidente, el resultado sería que estas otras perífrasis aparecerían con menor frecuencia en traducciones que en originales españoles, lo que indicaría que nuestro idioma cuenta con una amplio abanico de medios para expresar este significado, y no todos se activan en el proceso de traducir *start* o *begin*. De hecho, esto fue así en el caso de «comenzar a + infinitivo».

«Ponerse a + infinitivo» puede traducirse, por supuesto, como «*start/begin* + -ing», «*start/begin* + infinitivo» y también por medio de expresiones más específicas como *get (to)*, *get down to*, *set to*, *set about*, *set off*, *break into*, *burst into*, *burst out* + -ing, etcétera. No incluimos en el cómputo expresiones como «ponerse a parir», que consideramos locución.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words:

PONE*/PUEST*/PONIENDO/PONIENDO*/PON/PONTE/PONÉ*/PONG*/PÓNG*/PONÍA*/PUS*/PÚS*/PONDR*

Context words: A

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Ponerse a + INF	42	12,9	33	6,3	+9,48	< 0,01	SÍ

Ya hemos explicado anteriormente el motivo por el cual creemos que esta perífrasis confirma nuestra hipótesis. El traductor precisa ir un paso más allá para no obedecer el impulso más natural de utilizar «empezar a + infinitivo», perífrasis sobrerrepresentada en las traducciones, y escoger «ponerse a + infinitivo». Sea cual sea la razón, el subcorpus de originales en español presenta 6,6 apariciones más por cien mil de palabras que el de traducciones.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?p(o|ó)n(e|é|i|t|g|i).*(¿|i)?pondr.*(¿|i)?p(u|ú)(e?)s.*" %c] []{0,5} "a"

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Ponerse a + INF	43	11,7	115	23,5	-16,69	< 0,0001	NO

Una vez más, nos llevamos una sorpresa al comprobar que se ha invertido completamente la tendencia al alterar la composición del subcorpus de traducciones y la perífrasis ha pasado de estar infrarrepresentada —y con motivos que considerábamos más que justificados— a sobrerrepresentada en traducciones del inglés al español, con un nivel de certeza muy alto.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	13	15,3	Grafton	86	92,6
Madrid	11	26,5	Leonard	9	11,6
Martínez	9	16,6	Cornwell	8	7,3
Vázquez	5	5,7	Ellroy	7	6,4
Núñez	4	9,4	Highsmith	5	5,1
Martín	1	1,8			
Total	43	11,7		115	23,5

Tabla 42 - Distribución de «ponerse a + INF» por texto

Si nos centramos primero en ES, sorprende Madrid, que presenta la frecuencia normalizada más alta en el subcorpus ES a pesar de ser el último en número de palabras. Es evidente que tiene cierta preferencia por evitar el excesivo uso de *empezar*, pues en el caso de esta última perífrasis sí que está al final de la lista, mientras que en el de *comenzar* ocupa, como aquí, un puesto superior al esperado (el segundo). El caso opuesto es Núñez, con una abundancia llamativa de *empezar* pero solo cuatro casos de *ponerse* y dos de *comenzar*.

En TRAD, asimismo, sorprende la traducción de Moya, hecho que ya mencionábamos en el caso de la contraposición entre las perífrasis «empezar + infinitivo» y «comenzar + infinitivo», dado que pasaba de ocupar el último puesto en utilización del auxiliar *empezar* al primero en *comenzar*, lo que denotaba un afán por alejarse de la perífrasis incoativa más habitual. En el caso de la perífrasis que nos ocupa ahora, la diferencia con el resto de los autores es sobrecogedora, puesto que este traductor es responsable del 75% de las concordancias válidas.

Esto hace que se desequilibren las cifras por completo y «ponerse a + infinitivo» pase de estar infrarrepresentada a sobrerrepresentada en traducciones. De hecho, la distribución de la versión antigua del subcorpus TRAD era bastante equilibrada (9, 7, 7, 6 y 4 casos) y más parecida a la que obtendríamos si elimináramos el texto de Grafton

de la tabla actual (9, 8, 7 y 5). Una vez más, vemos cómo la idiosincrasia de un autor o traductor puede influir en los resultados en una investigación del lenguaje basada en corpus, convirtiendo en aparente normalidad lo que es particularidad de un texto, por lo que es importante en esta metodología tener en cuenta las aportaciones de los diferentes elementos que conforman el corpus.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Begin	25	21,7
∅	19	16,5
Marcador temporal	19	16,5
Start	16	13,9
Reformulación	14	12,2
Present participle clause	11	9,6
Get + suj + -ing/SP/inf	3	2,6
Verbo + pronombre reflexivo + -ing/SP	3	2,6
Phrasal verb o verbo + prep (go off/to)	3	2,6
Verbo <i>feel</i> seguido de -ing	2	1,7
TOTAL	115	100

Tabla 43 – Equivalentes de «ponerse a + INF» en EN

Como el resto de perífrasis incoativas, también es activada por los verbos *start* y *begin*, pero en bastante menor medida que «empezar a + infinitivo» (35 % frente al más de 60 % de esta última perífrasis).

<Grafton>: Me senté ante el pequeño escritorio y me <puse a> abrir cajones de manera sistemática .

-->par_en: I sat down at his small desk and **began** to open drawers systematically .

<Grafton>: Te <pusiste a> hablarle y pensé que todo era invención tuya , hasta que empuñaste la escopeta .

-->par_en: You **started** talking to him , I thought it was all bullshit until you pulled out that shotgun .

Las adiciones y reformulaciones, por su parte, conforman un porcentaje mayor (29 % contra el 13,5 % de «empezar a + infinitivo»).

<Grafton>: <Ponerse a> marcar números al azar era absurdo .

-->par_en: There was n't any point in **trying to guess the right one** .

<Grafton>: Meto el hocico en el agujero y me <pongo a> escarbar hasta que saco lo que hay dentro .

-->par_en: I 'm like a little terrier when it comes to the truth . I have to stick my nose down the hole **and dig until I find** out what 's in there .

En traducciones se usa la perífrasis más de lo que cabía esperar, pero los patrones que hemos encontrado son bastante diversos y afectan a porcentajes pequeños de las concordancias, menos del 10 %, excepto las que incluyen algún tipo de marcador temporal, como adverbios (*now, already, then* o *again*) y subordinadas de tiempo precedidas por las preposiciones *when, while, as, since* y *after*. Estas conforman casi un 17 %. El marcador temporal indica que la acción empieza en determinado momento aunque no se utilice una estructura incoativa explícita:

<Ellroy>: ¿Y busca un tipo loco ? : pruebe con Tommy K. Suele rondar por el Bido Lito's con sus amigos y se levanta y se <pone a> tocar ese maldito saxo cada vez que le dejan , que es a menudo porque , ¿quién se atreve a decir que no a un tío loco , aunque sea un tipo canijo como Tommy?

-->par_en: He hangs out with the suedes at Bido Lito 's and gets up and plays this godawful tenor sax whenever they let him , which is frequently , 'cause who wants to refuse a crazy man , even a little skinny twerp like Tommy ?

<Cornwell>: Se había <puesto a> llover con fuerza , y la presencia de coches que viajaban a velocidades imprudentes con los faros encendidos hacía que la circulación fuera peligrosa .

-->par_en: It was raining hard now , and streets were treacherous with cars moving at imprudent rates of speed with their headlights on .

<Cornwell>: Le di a Lucy las llaves del coche , le expliqué cómo llegar allí y , cuando se fue , me <puse a> leer ante el fuego y me quedé dormida.

-->par_en: I gave Lucy directions and my car keys , and after she left , I read in front of the fire and fell asleep .

En cuanto al resto de categorías, podemos mencionar que, como en «empezar + infinitivo», los traductores vuelven a valerse de una perífrasis incoativa para enlazar la subordinada con participio presente a la oración principal, dotando así de mayor cohesión al texto o dándole una estructura más natural en español pero, al mismo tiempo, explicitando ciertas relaciones semánticas implícitas en el original.

<Grafton>: Apreté el paso y el bolso se <puso a> golpearme la cadera mientras recorría al trote el corto trecho que había hasta la entrada .

-->par_en: I doubled my pace , my shoulder bag banging against one hip as I jogged down the short walk to the entrance .

<Grafton>: Sin duda le gustaba quejarse , y se <ponía a> recitar sus males para justificar su forma de tratar a los demás .

-->par_en: He probably liked to bitch , reciting his grievances in order to justify his ill treatment of other people .

Los demás patrones que hemos encontrado son los siguientes:

- Verbo *get* + someone/something + verbo/sintagma preposicional:

<Grafton>: Los tres guardamos silencio durante unos segundos , en espera de que la imaginación se <pusiera a> trabajar .

-->par_en: The three of us were silent for a moment , trying to **get our imaginations to work** .

- Verbo + pronombre reflexivo + participio presente o sintagma preposicional:

<Grafton>: Me <puse a> buscar el número de Bucky mientras el hombre salía de la tienda .

-->par_en: **I busied myself** looking up Bucky 's number while the guy came out of the gift shop behind me

- Verbo *feel* + objeto + participio presente:

<Grafton>: El estómago se me <puso a> murmurar en previsión de lo que iba a suceder .

-->par_en: **I could feel my stomach churning with anticipation** .

- Verbo seguido de preposición o *phrasal verb*:

<Grafton>: En cuanto abrí la puerta , se <puso a> sonar la alarma contra incendios .

-->par_en: When I opened the door , the fire alarm **went off** .

4.4.3.3. Echar(se) a + infinitivo

«Echar a + infinitivo» o «echarse a + infinitivo» es también una perífrasis incoativa pero aporta un matiz de brusquedad e inicio puntual y repentino. Se encuentra muy restringida en cuanto a los infinitivos con los que puede usarse. Con la forma no pronominal, principalmente *andar*, *correr*, *volar* y *nadar*; con la forma pronominal, fundamentalmente *reír*, *llorar*, *temblar*, *correr*, *volar* (aunque en teoría puedan darse otras combinaciones como las que menciona García Fernández (2012: 72): *bailar*, *caminar*, *gritar*, etcétera. Consideraremos locuciones las expresiones «echar a perder» y «echar a rodar (p. ej. una moneda)».

Recordemos una vez más que habíamos hipotetizado que existiría una infrarrepresentación de expresiones incoativas alternativas o menos habituales que

«empezar a + infinitivo» como «comenzar a + infinitivo» y «ponerse a + infinitivo» —en este último caso, por cierto, hemos visto que nuestra intuición ha sido refutada en la segunda versión del corpus—, pero no de las variantes, como «echar(se) a + infinitivo» o «romper a + infinitivo», que cuentan con equivalentes estilísticos obvios en inglés que pueden encontrarse en cualquier diccionario bilingüe («*burst/break out + -ing*», «*break/burst into a + sustantivo*», «*set/run/fly off*», etcétera.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: ECH*/ÉCH*

Context words: A

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Echar(se) a + INF	31	9,5	77	14,8	-4,47	< 0,05	NO

Como habíamos previsto, esta perífrasis se encuentra sobrerrepresentada en el subcorpus de traducciones, probablemente debido a los equivalentes estilísticos con los que cuenta en inglés, además de tratarse de expresiones muy típicas del lenguaje literario que pueden encontrarse fácilmente en diccionarios bilingües, por lo que son inmediatamente accesibles al traductor.

En cuanto a la cuestión de la reducida variedad de infinitivos que admite, el uso difiere entre subcorpus: en los originales en español se utiliza, de 31 apariciones, *reír* en 22 ocasiones (71 % del total), *llorar* en 7 (22,6 %) y *correr* y *temblar* en 1 (3,2 % cada una). En los textos traducidos, sin embargo, de 77 apariciones, aparte de una amplia mayoría de combinaciones con *reír* (42 casos, 54,5 %) y 5 ejemplos de *llorar* (6,5 %), 2 de *volar* (3 %), 2 de *temblar* (3 %) y 1 de *gritar* (1 %), principalmente encontramos los verbos *andar* (14 casos, 18 %) y *correr* (11 casos, 14 %), que apenas aparecen en el subcorpus de español original (de hecho, *andar* no aparece ni una vez combinado con «echar(se) a» en ES). Esto puede deberse simplemente a que el subcorpus de traducciones es más grande, pero sería interesante averiguar si hay expresiones específicas del inglés que desencadenan el uso de estas dos últimas combinaciones que no parecen ser tan frecuentes en lenguaje no traducido. Sea como sea, estos datos

confirman la tendencia de los textos traducidos a mostrar patrones colocacionales disntintos de los no traducidos (Mauranen 2000).

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?(e|é)ch.*" %c] []{0,5} "a";

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Echar(se) a + INF	37	10,1	85	17,4	-8,17	< 0,01	NO

Se confirma la sobrerrepresentación de esta perífrasis en textos traducidos, esta vez con mayor nivel de significación estadística. Encontramos las siguientes combinaciones con auxiliares: en español original, 23 casos de *reír* (62 % del total), 8 de *llorar* (22 %), 3 de *correr* (8 %), 2 de *andar* (5 %) y una de *temblar* (3 %). Vemos entonces que *correr* aumenta su presencia en originales y *andar* aparece ahora gracias al texto de Núñez. Por otro lado, en traducciones disminuye el número de apariciones de *reír* a 34 (40 %), pero aumentan las de *andar* y *correr* a 29 (34 %) y 15 (18 %) respectivamente. La variedad de verbos disminuye con respecto a la versión anterior del corpus, ya que solo quedan por enumerar 6 casos de *llorar* (7 %) y uno de *volar* (1 %). Sin embargo, la frecuencia de *correr* (8 % en ES frente a 18 % en TRAD) y *andar* (5 % en ES frente a 34% en TRAD) sigue siendo excesivamente elevada en traducciones si la comparamos con su utilización en originales.

Podríamos referirnos aquí a la hipótesis de la simplificación en textos traducidos (Laviosa 1997, 1998). Ambos subcorpus cuentan ahora con cinco verbos distintos cada uno, aunque no sean los mismos. Pero hay que tener en cuenta que los casos de TRAD duplican a los de ES, por lo que cabría esperar que la variedad de verbos utilizada fuera mayor (de hecho, tanto en ES como en la versión anterior de TRAD encontramos otros verbos, lo que indica que otras combinaciones son posibles). Lo que observamos, por el contrario, es una menor variedad de combinaciones que se repiten más.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Vázquez	13	14,8	Highsmith	23	23,4
Giménez	12	14,1	Leonard	23	29,6
Núñez	6	14,1	Grafton	18	19,4
Martín	3	5,3	Ellroy	11	10
Martínez	2	3,7	Cornwell	10	9,1
Madrid	1	2,4			
Total	37	10,1		85	17,4

Tabla 44 - Distribución de «echar(se) a + INF» por texto

A partir de la distribución por textos vemos que en ES se distinguen dos grupos: en uno de ellos, los tres escritores presentan prácticamente la misma frecuencia normalizada. Sin embargo, todos presentan frecuencias inferiores a las de las traducciones, al menos las de los tres primeros textos. Esta diferencia puede tener que ver con que estamos ante una perífrasis que funciona como una solución de traducción habitual, ya sea porque existen estructuras en inglés que la activan o porque, por tradición de uso, ya ha pasado a formar parte de los recursos estilísticos de lo que se considera el lenguaje de una traducción hoy en día.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
∅	40	47,1
Phrasal verb	20	23,5
Start	9	10,6
Verbo + preposición	7	8,2
Reformulación	3	3,5
Break/burst into + sust	3	3,5
Begin	3	3,5
TOTAL	85	100

Tabla 45 - Equivalentes de «echar(se) a + INF» en EN

Las apariciones de *start* y *begin* son escasas en esta ocasión (14,1%), comparadas con las de las otras perífrasis incoativas. Tampoco observamos demasiados ejemplos de los equivalentes que podrían considerarse más directos según la caracterización que hemos hecho de la perífrasis: hallamos solo un «*burst into (tears)*» y dos «*break into (a run/trot)*».

<Grafton>: Pensaba que sabías dónde estaba - susurró , <echándose a> llorar otra vez .

-->par_en: I thought you knew where it was , " she whispered , **beginning** to weep again .

<Cornwell>: Le dio un beso en la boca y se <echaron a> reír los dos .

-->par_en: He kissed her on the mouth and both of them **started** laughing .

Con respecto a nuestra idea de que esta perífrasis ha pasado ya a formar parte de los recursos estilísticos habituales del español traducido, podemos decir que con algunos verbos esto parece especialmente cierto, puesto que, por ejemplo, *laugh* aparece casi invariablemente solo pero el traductor no puede evitar añadir el «echarse a» al verbo *reír*. La expresión parece ya casi fosilizada. Este tipo de ejemplos conforman el 47,1 % de adiciones.

<Leonard>: Las demás mujeres se <echaron a> reír , algunas soltando gritos de admiración para agradar a la de la peluca , o sólo para hacer resonar sus voces entre aquellas paredes de cemento , hasta que la señora Kay ordenó : - Ya vale - y se callaron .

-->par_en: The other women laughed , some with screams of appreciation , to please the woman in the wig or to let go and hear the sound of their own voices , loud inside the cement-block walls , until Miss Kay said , " Zip it , " and they shut up .

<Ellroy>: Me <eché a> reír .

-->par_en: I laughed .

En cuanto a *run*, *walk* y sinónimos (p. ej. *stroll*, *move* o *start*), parecen ir acompañados casi siempre de una preposición o partícula adverbial (*off*, *away*, *up*) cuando se traducen por «echar a + infinitivo» (31,7 %).

<Leonard>: - Deberías - contestó ella , <echando a> andar hacia la arena mojada .

-->par_en: She said , " You should , " and walked out to the hard-packed wet sand .

<Leonard>: En cuanto estuviera de nuevo fuera y a la vista, tendría que parecer ansiosa , desesperada , y <echar a> correr para buscar a Nicolet , a cualquiera , y contarle lo que le había ocurrido .

-->par_en: Once she was out on the floor in plain sight she would have to appear anxious , helpless , and run off looking for Nicolet , someone , to tell what happened .

<Grafton>: Llamé a la puerta , pero como pasaron varios minutos sin que nadie respondiese , <eché a> andar por el sendero , hacia la parte trasera .

-->par_en: I knocked on the door , but after several minutes went by and nobody answered , I headed down the driveway toward the back .

Al parecer, esta reducida variedad de verbos combinados con preposición o *phrasal verbs*, dependiendo del caso —o incluso sin preposición, teniendo en cuenta que las adiciones representan casi un 50%—, activan en la mente del traductor la perífrasis de la que hablamos, que puede llegar a sonar mejor estilísticamente o más idiomática, cuando en realidad en español no parece utilizarse con tanta asiduidad.

4.4.3.4. Romper a + infinitivo

Esta quinta perífrasis incoativa posee mayor fuerza expresiva que las anteriores, puesto que sugiere una contención previa por parte del sujeto de la acción que comienza a realizar. Aporta matices estilísticos de brusquedad y violencia y su uso también está restringido a unos cuantos infinitivos: los más habituales son *reír*, *llorar*, *hablar*, *andar*, *gritar*, *chillar* o *cantar*, aunque García Fernández (2012: 72) recoge varios más. Al igual que «*echar a + infinitivo*», según los diccionarios y obras contrastivas, cuenta con detonantes obvios en inglés, como *break into*, *burst into*, *burst out*, *break down* y adverbios como *suddenly*.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: ROMP*/ROT*

Context words: A

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Romper a + INF	0	0	10	1,9	-9,69	< 0,01 ¹³⁸	NO

Sorprendentemente, «romper a + infinitivo» ni siquiera aparece en el subcorpus de originales españoles. Sí lo encontramos en el de traducciones en diez ocasiones, siempre con los verbos *llorar* (en siete ocasiones, y una con el verbo *sollozar*) y *reír* (dos ocasiones). Sería necesario un corpus de dimensiones mayores para averiguar si se utiliza en español original y con qué frecuencia.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|)r(o|ó)mp.*|(¿|)roto" %c] [0,5} "a";

¹³⁸ Test exacto de Fisher: 0,009054 (p < 0,01)

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Romper a + INF	0	0	2	0,4	-2,24	n. s. ¹³⁹	NO

En español original sigue sin haber ejemplos de esta perífrasis. Casi todos los casos que faltan en TRAD (siete: seis «echar a reír» y un «echar a llorar») correspondían a la traducción de la novela *L. A. Confidential*, de Ellroy, lo que indica una preferencia por esta expresión en el traductor argentino Carlos Gardini, que lo empuja a utilizarla por encima de la frecuencia en español que hemos encontrado en nuestro subcorpus de originales o simplemente una mayor frecuencia de las posibles estructuras equivalentes en el texto original (los otros dos ejemplos eran del texto de Hunter, *llorar* y *sollozar*).

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	0	0	Grafton	1	1,1
Madrid	0	0	Highsmith	1	1
Martín	0	0	Cornwell	0	0
Martínez	0	0	Ellroy	0	0
Núñez	0	0	Leonard	0	0
Vázquez	0	0			
Total	0	0		2	0,4

Tabla 46 - Distribución de «romper a + INF» por texto

En esta nueva versión del corpus solo encontramos dos casos: uno en la novela de Highsmith (con *reír*) y otro en la de Grafton (con *llorar*). Las cifras, que en principio indican una sobrerrepresentación de esta perífrasis en traducciones, no son estadísticamente significativas, por lo que no puede tenerse en cuenta esta diferencia en el cómputo final.

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Burst out	1	50
Suddenly	1	50
TOTAL	2	100

Tabla 47 - Equivalentes de «romper a + INF» en EN

Evidentemente, no podemos extraer datos significativos de tan escasos ejemplos de utilización, pero tanto los verbos principales utilizados —*llorar* y *reír*— como los elementos activadores de los dos casos —el adverbio de modo *suddenly* y el phrasal

¹³⁹ Test exacto de Fisher: 0,510077 (n. s.)

verb *burst out*, ambos utilizados para dotar a la acción de un carácter brusco y repentino— se ajustan a lo que mencionábamos en la caracterización de la perífrasis:

<Grafton>: A veces , cuando bebe , <rompe a> llorar como un niño .

-->par_en: Sometimes when he drinks , he **bursts out** crying like a baby .

<Highsmith>: De pronto , Tom <rompió a> reír a carcajadas , casi se bamboleó en la sala , por donde estaba paseando despacio .

-->par_en: Tom **suddenly** doubled over with laughter , nearly staggered in the living room , where he had been slowly walking about .

4.4.3.5. Estar a punto de + infinitivo

«Estar a punto de + infinitivo» es una construcción cuyo auxiliar tiene la particularidad de ser una locución, por lo que se incluye sólo en el repertorio de perífrasis de Fernández de Castro y Gómez Torrego, quienes no la consideran completamente perifrástica. García Fernández (2012: 46) opina asimismo que su grado de gramaticalización es bajo.

Posee valor aspectual de inminencia y, con el auxiliar en pasado, describe una acción que no llegó a realizarse pero estuvo muy cerca de hacerlo: «Estuvimos a punto de abandonar» (*We nearly gave up*). Entre los equivalentes de esta perífrasis se incluyen adverbios de grado como *almost* o *nearly* y expresiones como «*be about to*», «*come within (...) of + -ing*» o «*be on the point/verge/brink of*», que cuentan con traducciones al español formalmente más similares: «casi», «por poco» o «estar al borde/a las puertas de».

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: A PUNTO

Context words: DE

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Estar a punto de +INF	39	12	44	8,4	+2,53	n. s.	NO

A pesar de que Gómez Torrego (1999: 3375) afirma que el auxiliar de esta perífrasis es una locución («estar a punto de»), hemos encontrado combinaciones con otros verbos como *hallar* o *parecer* en lugar de *estar* (aunque sólo hemos contado las formadas con *estar*). Esto, y el hecho de que haya frases sin verbo, sólo con la secuencia «a punto de», hace pensar que es ésta la que otorga el significado de inminencia, no la expresión en su conjunto (opinión expresada también por Fernández de Castro 1999: 228). Además de todo lo mencionado, la frecuente intercalación de elementos entre auxiliar y auxiliado hace dudar de su carácter perifrástico. El cómputo de apariciones indica que la diferencia de utilización de esta estructura entre originales y traducciones no es significativa.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

"(¿|i)?a" %c "punto" [{"0,5} "de";

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Estar a punto de +INF	42	11,4	49	10	+0,39	n. s.	NO

La diferencia entre subcorpus, aunque parece favorecer en número de apariciones al subcorpus de originales españoles, sigue careciendo de significación estadística. En todos los textos de ambos subcorpus encontramos casos; lo único que llama la atención es la escasez de casos en Giménez Bartlett, la segunda más extensa de ES, y la abundancia de casos en la traducción del texto de Leonard, cuya novela es la menos extensa de TRAD.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martínez	11	20,2	Leonard	18	23,1
Martín	9	16	Highsmith	9	9,1
Vázquez	9	10,2	Cornwell	8	7,3
Giménez	6	7,1	Ellroy	7	6,4
Núñez	4	9,4	Grafton	7	7,5
Madrid	3	7,2			
Total	42	11,4		49	10

Tabla 48 - Distribución de «estar a punto de + INF» por texto

Hemos elegido tener en cuenta solo los casos en los que la perífrasis aparece en su forma completa. En ambos subcorpus encontramos casos de utilización de verbos de

significado similar a *estar*, como *hallarse* (1 en ES) o «parecer» (1 en ES y 4 en TRAD).

<Giménez>: Nos hallábamos <a punto de> lograr una cota bastante absoluta : la profunda y total enemistad del mundo entero .

<Vázquez>: El mejor novelista gay de las dos Castillas parecía <a punto de> llorar y Sagalés de reír .

También encontramos un ejemplo de orden invertido de los componentes en el siguiente ejemplo:

<Núñez>: <A punto estuve de> cantar de plano para ahorrarle molestias , detalle de rebanada de pelotas incluido .

En un número reducido de concordancias encontramos elementos intercalados entre el auxiliar y el resto de la perífrasis (2 en ES y 1 en TRAD).

<Giménez>: Su cara pequeña y poco atractiva estaba desvaída , <a punto de> borrarse por completo .

Este último supuesto podría considerarse parte del que hallamos en mayor medida en ambos subcorpus, que es la elisión del auxiliar (4 en ES y 11 en TRAD):

<Vázquez>: Fue una entrevista tormentosa , así me lo dijo más o menos en clave cuando le llamé desde el hotel , <a punto de> salir para este premio y quedamos en hablar de tú a tú , en un aparte .

No hemos querido contar estos casos porque, al faltar dicho auxiliar, la expresión «a punto de» se convierte a veces casi en adjetivo:

<Grafton>: Gilbert cerró los ojos con la paciencia <a punto de> agotársele .

-->par_en: Gilbert closed his eyes , his patience wearing thin .

Sin embargo, aun habiendo tenido en cuenta estos casos, la diferencia entre subcorpus habría continuado siendo estadísticamente no significativa (50 casos en ES, 66 en TRAD; LL +0,0).

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
About	26	53,1
Almost	8	16,3
Reformulación	7	14,3
To be ready + inf	3	6,1
Be/come close to + -ing	2	4,1
Near(ly)	2	4,1
On the verge of + -ing	1	2
TOTAL	49	100

Tabla 49 - Equivalentes de «estar a punto de + INF» en EN

En las concordancias paralelas de esta perífrasis encontramos una amplia variedad de equivalentes: se encuentran representados por primera vez casi todos los elementos que nombrábamos en la introducción e incluso encontramos más opciones, como «*to be ready* + infinitivo».

<Ellroy>: - Creo que estabas <a punto de> repartir un poco de leña sindical .

-->par_en: " I think you **almost** did some union thumping . "

<Highsmith>:Madame Annette estaba <a punto de> irse a sus compras matinales .

-->par_en: Mme . Annette was **ready** to depart for morning shopping

<Leonard>: Max estuvo <a punto de> responderle : « Sí , y ¿sabe qué me ha dicho ? » .

-->par_en: Max **came close** to saying , yeah , and you know what she said to me ?

<Highsmith>: Estaba <a punto de> desmayarse .

-->par_en: He was **near** fainting .

<Grafton> Laura dio un gemido e hice amago de sujetarla como si estuviera <a punto de> desmayarse .

-->par_en: Laura moaned and leaned forward as if she were **on the verge of** fainting .

No obstante, el más frecuente con diferencia es *about*, que abarca el 53,1 % de concordancias en su forma más habitual «*to be about* + infinitivo» pero también en otras expresiones sin verbo:

<Cornwell>: Mi madre estaba <a punto de> llorar .

-->par_en: My mother was **about** to cry .

<Grafton>: Estaba ya <a punto de> irme cuando vi movimiento en el porche del bungalow .

-->par_en: Just about the time I was ready to pack it in , I saw some movement on the cottage porch

No hemos tenido que incluir la categoría de apariciones sin elemento activador, lo que no suele ser habitual. Es evidente que a esta perífrasis no le faltan detonantes en inglés, y esto hace probablemente que no se encuentren diferencias significativas entre originales y traducciones con respecto a su utilización.

4.4.3.6. Dejar de + infinitivo - Parar de + infinitivo

Estas dos perífrasis de valor aspectual terminativo expresan la interrupción de una acción o proceso: «¡Deja de comer tanto!» (*Stop/quit eating so much!*). Cuando la frase es negativa, el valor pasa a ser reiterativo, ya que la acción continúa realizándose: «No he dejado de correr desde que salí de casa» (*I haven't stopped running since I left home*). «Dejar de + infinitivo», en enunciados negativos, enfatiza una afirmación: «No deja de sorprenderme que se comporte así» (*I still find it surprising that he behaves that way*). Otras opciones de traducción para enunciados negativos incluyen «to (not) forget + infinitivo», «to make/be sure + infinitivo», «keep + -ing» o «without fail», que pueden traducirse de manera más literal al español por medio de expresiones como «no te olvides de», «asegúrate de» o «sin falta». *Parar* admite menos variedad de infinitivos y no adquiere el valor enfático de *dejar* en negaciones pero se utiliza más en negaciones de valor reiterativo.

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: DEJ*/DÉJ*
 PARO/PARE*/PARÉ*/PARA*/PARÁ*
 But excluding: PAREC*/PARED*
 Context words: DE
 Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Dejar de + INF	81	25	80	15,4	+9,38	< 0,01	SÍ
Parar de + INF	11	3,4	6	1,1	+4,80	< 0,05	SÍ

Tanto «dejar de + infinitivo» como su expresión sinónima «parar de + infinitivo» se encuentran infrarrepresentadas en traducciones al español. En ambos subcorpus, esta segunda perífrasis es mucho menos frecuente y aparece en enunciados negativos en prácticamente todos los casos.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?d(e|é)j.*" %c] []{0,5} "de";

[word="(¿|i)?par(o|ó|a|á|e|é)(n|s)?|(¿|i)?par(a|á)(r|n|d|m).*(¿|i)?par(a|á|e|é)(i|s|b|m).*" %c] []{0,5} "de"

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Dejar de + INF	95	25,9	101	20,7	+2,46	n. s.	NO
Parar de + INF	12	3,3	9	1,8	+1,72	n. s.	NO

Quizás la diferencia en el número de palabras entre subcorpus que existía en la fase I realizaba o sobrevaloraba la importancia de la diferencia en las frecuencias de aparición de las perífrasis: en aquella versión la diferencia de frecuencias no era muy grande y, sin embargo, alcanzaba para que el test de significación estadística confirmara la infrarrepresentación de ambas perífrasis en traducciones. En esta segunda versión del corpus, no obstante, las diferencias no resultan lo suficientemente importantes para superar el umbral de significación en ninguno de los dos casos.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	33	38,9	Cornwell	28	25,5
Madrid	17	41	Leonard	26	33,4
Núñez	13	30,6	Grafton	22	23,7
Martín	12	21,3	Ellroy	16	14,6
Martínez	11	20,2	Highsmith	9	9,2
Vázquez	9	10,2			
Total	95	25,9		101	20,7

Tabla 50 - Distribución de «dejar de + INF» por texto

Las obras con menor cantidad de palabras son las que presentan, en el caso de «dejar + infinitivo», las frecuencias de uso más elevadas (Madrid en ES y Leonard en TRAD). Por otra parte, podemos afirmar que el valor de reiteración (el uso de la perífrasis en negativo) es bastante más frecuente en TRAD (40 casos) que en ES (15 casos).

<Martín>: Desde que ese tipo empezó a hacer de las suyas , la Policía no <deja de> acosarnos , persiguiendo macarras y camellos .

El uso enfático es similar en ambos subcorpus (4 y 5 casos respectivamente):

<Madrid>: ¿Sabe como llaman a esto ? Financiación . No <deja de> tener su gracia , ¿verdad ?

En cuanto a «parar de + infinitivo», parece presentar unas preferencias semánticas más específicas: confirmamos el uso casi exclusivo de la perífrasis en enunciados negativos y el predominio de verbos de expresión (*hablar, llorar, decir, repetir o quejarse*), por lo que ambas perífrasis no son del todo intercambiables en sus valores.

<Núñez>: A primera hora de la mañana , Beatriz tuvo una llamada de teléfono y no <paró de> llorar el resto del día .

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Madrid	3	7,2	Cornwell	4	3,6
Martínez	3	5,5	Highsmith	3	3
Giménez	2	2,3	Grafton	2	2,1
Vázquez	2	2,3	Ellroy	0	0
Núñez	1	2,3	Leonard	0	0
Martín	1	1,8			
Total	12	3,3		9	1,8

Tabla 51 - Distribución de «parar de + INF» por texto

Mientras que todos los autores de ES cuentan con al menos una concordancia de «parar de +infinitivo», las nueve de TRAD se distribuyen entre solo tres de los cinco textos. Madrid vuelve a presentar la frecuencia más alta de ES, mientras que Leonard no utiliza esta perífrasis en ningún caso.

- CORPUS PARALELO

Vemos en la tabla de equivalentes que la variedad de estructuras que pueden activar el uso de «dejar + infinitivo» es amplísima.

Elemento activador EN	Frecuencia	Porcentaje
Quit	18	17,8
Stop	17	16,8
Reformulación	13	13,9
Still	9	8,9
Present participle clause	7	6,9
Marcador temporal	7	6,9
Phrasal verb	6	5,9
Enunciado negativo	5	4,9
Keep	5	4
As + suj + -ed	4	4
Cease	3	3
∅	2	2
Fail to	2	2
Continue	2	2
Can't help but	1	1
TOTAL	101	100

Tabla 52 - Equivalentes de «dejar de + INF» en EN

Los verbos *stop* y *quit*, incluidos en diversas estructuras con infinitivos, sintagmas nominales, gerundios o participios presentes, son los más frecuentes, y junto con su sinónimo *cease* conjugan un 37,6 % de las concordancias.

<Grafton>: <Deja de> llorar .

-->par_en: Quit with the crying .

<Cornwell>: Afirmó también que esperaba la muerte con impaciencia y que prefería que yo <dejara de> solicitar un habeas corpus federal .

-->par_en: He said that he was looking forward to dying and preferred that I cease pursuing federal habeas corpus relief .

<Highsmith>: Tom había pronunciado el nombre de Bernard ante Murchison (según recordaba) , cuando intentaba persuadirlo de que <dejara de> investigar sobre las falsificaciones .

-->par_en: Tom had uttered Bernard's name to Murchison (as Tom recalled) when trying to persuade Murchison to stop his forgery inquiry .

De los cuatro casos enfáticos de TRAD, en tres se utiliza *still* (el otro es una adición). Los seis casos restantes de utilización de este adverbio son parte de enunciados reiterativos:

<Leonard>: Aún lo andaba buscando , pues ya llevaba cinco días con el Rolex falso que no era feo y marcaba bien la hora , pero no <dejaba de> costar sólo unos cientos de pavos .

-->par_en: Still trying to find him , five days now with the fake Rolex that wasn't bad-looking , kept the right time , but still wasn't worth a thousand bucks .

<Leonard>: Él la agarró sin <dejar de> moverse .

-->par_en: He took hold of her , still moving

Dentro de los 40 casos reiterativos encontramos, además, las siete concordancias que incluyen un marcador temporal que indica la repetición de la acción o su prolongación en el tiempo y las siete formadas con un participio presente yuxtapuesto a la frase principal, que expresa acción simultánea.

<Cornwell>: Suponía que nunca <dejaría de> intrigarme el que poesía y crueldad pudieran residir en el mismo corazón .

-->par_en: I supposed it would always intrigue me that poetry and cruelty could reside in the same heart .

<Leonard>: Louis hizo un gesto con la cabeza y Ordell lo siguió a la sala de espera sin <dejar de> hablar :
- Tío , te he estado llamando sin parar.. .

-->par_en: Louis motioned with a nod of his head and Ordell followed him into the meeting room , Ordell saying , " Man , I been calling and calling you.. . "

<Leonard>: El defensor se volvió hacia ella sin <dejar de> mirar un expediente y le preguntó cómo pensaba declararse .

-->par_en: He turned to her looking at a case file and asked how she wanted to plead .

También dentro de los casos reiterativos encontramos aquellas concordancias que incluyen los verbos *keep* y *continue*, las que expresan acción simultánea mediante subordinadas introducidas por la conjunción *as*, y la que utiliza la expresión «*can't help but + -ing*»

<Ellroy>: Está metido en no sé qué extraña actuación y no <deja de> hablar de su « interpretación » y de su « llamada a escena » .

-->par_en: He 's got some kind of strange acting bug , and he keeps talking about his ' cue ' and his 'curtain call . ' "

<Cornwell>: Su respuesta fue abrir el sobre que contenía las fotografías de Robyn Naismith y sacar un grueso fajo de copias en color que no <dejarían de> impresionarme por más veces que las viera .

-->par_en: His answer was to open the envelope containing Robyn Naismith 's photographs and slide out a thick stack of color prints that would continue to shock me no matter how many times I looked at them .

<Highsmith>: Tom echó a andar , se enjugó el sudor de la frente con la parte superior de la chilaba verde claro , y luego la arremangó y se la quitó por la cabeza sin <dejar de> andar .

-->par_en: Tom started walking , wiped sweat from his forehead with the upper part of the pale green djellaba , then hoisted it up at the sides , and pulled it over his head as he walked .

<Cornwell>: Francamente , no puedo <dejar de> preguntarme si el ataque contra Heath no pretendía ser simbólico , por alguna razón .

-->par_en: Frankly , I ca n't help but wonder if the attack on Heath was n't , for some reason , intended to be symbolic . "

Otros equivalentes presentes en los originales, aunque cada uno conforma un porcentaje muy pequeño de las concordancias, son diferentes *phrasal verbs*, simples enunciados negativos, y «fail to».

<Grafton>: Mantuve y sostuve que me habían obligado a punta de pistola , y que me habían dado una paliza y tirado a la cuneta cuando había <dejado de> serles útil .

-->par_en: I maintained stoutly that I 'd been forced along at gunpoint , clobbered and abandoned when my usefulness ran out .

<Cornwell>: Es cierto que llamé varias veces a casa de Susan cuando <dejó de> acudir al trabajo .

-->par_en: I did call Susan 's house when she did n't show up at work .

<Grafton>: Puede que haya <dejado de> mencionar un par de cosas .

-->par_en: " I might have failed to mention a couple things . "

Finalmente, un 15,9 % corresponde a reformulaciones y concordancias sin elementos activadores identificables, aunque de estas últimas apenas hay casos. «Dejar de + infinitivo» se utiliza para solucionar multitud de problemas de traducción, pero prácticamente siempre de manera motivada por estructuras de valor similar a la perífrasis y el contexto.

Elemento activador EN	Frecuencia	Porcentaje
Expresión de tiempo o cantidad	3	33,3
Stop	2	22,2
Reformulación	2	22,2
∅	2	22,1
TOTAL	9	100

Tabla 53 - Equivalentes de «parar de + INF» en EN

Los pocos casos que hemos encontrado en el subcorpus de traducciones no nos permiten establecer patrones claros en la utilización de la perífrasis «parar de + infinitivo». Vemos que el valor terminativo es anecdótico, pues solo hay un caso en el subcorpus y se traduce por *stop*.

<Cornwell>: El sábado , cuando me levanté a las cuatro de la mañana y fui a despertar a Lucy , había <parado de nevar> .

-->par_en: Saturday , the snow had stopped when I got up at four A.M. and went into Lucy 's bedroom to wake her .

El resto son enunciados negativos de valor reiterativo: por un lado, volvemos a encontrar una utilización de *stop* con este valor, y, por otro, vemos que tres enunciados se valen de expresiones de duración temporal o cantidad para transmitir la sensación de reiteración de la acción durante un período de tiempo determinado, que el traductor reexpresa por medio de la perífrasis.

<Cornwell>: Estaba alerta y pensaba con lucidez , pero no <paraban de temblarme> las manos - ..

-->par_en: I was alert and thinking clearly but my hands would not **stop** shaking .

<Highsmith>: Las escaleras que llevaban arriba quedaban a un lado de la recargada chimenea . Tom subió con Janice , que no <paraba de hablar> .

-->par_en: The stairs up were on one side of the gaudy fireplace , and he climbed them with Janice , who was talking **all the while** .

<Grafton>: Y no <paraba de repetir> « ¿Qué ? » .

-->par_en: He said , " What ? " **quite a lot** . The others tended to ignore him .

Los últimos cuatro casos corresponden a dos adiciones y dos reformulaciones, que conforman un 44,3 % de las concordancias:

<Grafton>: El tipo no <para de decir> : « Oye , tú , ¿por qué me haces esto ? Te he salvado la vida

-->par_en: **He 's all** , 'Hey , why 'd you do that when I saved your life ? ' And the snake goes

<Cornwell>: Y menos aún si pensamos que iba vestido de blanco y que no <paraba de quejarse> porque le habían llenado la casa de polvo .

-->par_en: Especially since he **was dressed in white and bitching** about the dusting powder everywhere .

4.4.3.7. Soler + infinitivo

«Soler + infinitivo» es una perífrasis frecuentativa que expresa acciones habituales o frecuentes. *Soler* se utiliza actualmente sólo como verbo auxiliar, por lo que la naturaleza perifrástica de esta construcción es indudable.

En frases de valor frecuentativo en tiempo presente (p. ej. «Suelo comprarme la ropa en tiendas de segunda mano»), el inglés no posee una estructura paralela para traducir este valor y requiere la utilización de partículas, marcadores temporales o expresiones adverbiales como *frequently*, *usually*, *often*, *tend to* o *be in the habit of*.

Esto significa que la perífrasis podría utilizarse menos frecuentemente en traducciones, ya que el español también dispone de expresiones temporales, adverbios o construcciones similares a las del inglés que expresan el mismo matiz, como «generalmente», «habitualmente», «por lo común», «normalmente», «tender a», etcétera.

Sin embargo, para enunciados en tiempo pasado, existen los modales *used to* y *would*, de significado prácticamente idéntico, que podrían activar fácilmente el uso de «soler + infinitivo» en traducciones: «Solía salir a pasear todos los domingos» (*She used to go out for a walk every Sunday*).

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: SUEL*/SOLI*/SOLÍ*/SOLAMOS/SOLÁIS/SOLEMOS/SOLÉIS

Context words: *AR*/*ER*/*IR*

Context Search Horizons: 0L to 4R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	p	Verificación de la hipótesis
Soler + INF	76	23,4	88	16,9	+4,28	< 0,05	SÍ

En principio podríamos pensar que esta perífrasis estaría bien representada en el subcorpus de traducciones, dado que en inglés existe la construcción *used to*, que muchas veces es traducido al español por *soler* en pretérito imperfecto (a veces, sin necesidad). Lo que encontramos en el corpus, sin embargo, es una infrarrepresentación de esta perífrasis, pues, como hemos explicado, el grado de similitud entre ambas lenguas en este sentido cambia según el tiempo verbal (pasado o presente) en que se use la perífrasis. Mientras que en inglés existen *used to* y *would* como traducciones estandarizadas para traducir «soler + infinitivo» en tiempo pasado, para enunciados de valor frecuentativo en tiempo presente, este idioma requiere el uso de determinadas partículas o expresiones adverbiales (como *frequently*) o marcadores temporales (como *everyday*), lo que puede hacer que el traductor utilice una opción más literal de las disponibles en nuestra lengua para expresar acciones habituales («frecuentemente» o «todos los días») y deje de lado la perífrasis que nos ocupa.

Sea como sea, existe una infrarrepresentación de esta perífrasis en el subcorpus de traducciones con respecto al de originales. Además, es notoria la diferencia de

tiempos verbales entre ambos subcorpus. En nuestro corpus solo hay ejemplos de *soler* en presente y pretérito imperfecto. En el de originales españoles predomina el auxiliar en presente (e.g. *suele*) en el 74 % de las concordancias (56 apariciones de 76), mientras que en el de traducciones, el auxiliar aparece en pretérito imperfecto (e.g. *solía*) también en el 74 % de los casos (65 apariciones de 88), debido probablemente a lo que apuntábamos antes.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?soler|(¿|i)?sol(e|a)mos|(¿|i)?suel.*|(¿|i)?sol(i|f).*|(¿|i)?sol(á|é)is" %c];

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Soler + INF	77	21	102	20,9	+0,00	n. s.	NO

Una vez más, comprobamos cómo se desvanecen las diferencias significativas entre subcorpus al modificar la selección de textos. La perífrasis «soler + infinitivo» no se encuentra infrarrepresentada en esta versión del subcorpus de traducciones, puesto que el test estadístico nos indica que no hay diferencia en la frecuencia de utilización de la misma por parte de autores españoles y traductores del inglés. Los dos traductores argentinos era quienes menos utilizaban la perífrasis (solo un caso en el texto de Ellroy y cinco casos en el de Mosley), por lo que, al cambiar sus textos por otros, la frecuencia de aparición aumentó considerablemente.

En esta ocasión, además, la diferencia entre el uso de pasado y presente en las concordancias de «soler + infinitivo» no es tan acusada como en la primera versión del corpus. Vemos que el porcentaje de concordancias en pasado de TRAD ha disminuido de un 74 % a un 47%. Aun así, sigue siendo bastante mayor que el 26 % registrado en ES (74 % en presente), lo que, a nuestro entender, es un hecho muy significativo que nos indica que, o bien *used to* es una expresión más frecuente en inglés de lo que «soler + infinitivo» en pasado lo es en español, o bien la utilización de la perífrasis en presente se reduce en traducciones por la influencia de los adverbios temporales del inglés, que empujan al traductor a utilizar soluciones más cercanas al idioma de partida.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	24	28,3	Highsmith	32	32,5
Vázquez	17	19,4	Cornwell	26	23,7
Madrid	16	38,6	Leonard	22	28,3
Martín	13	23,1	Grafton	16	17,2
Martínez	5	9,2	Ellroy	6	5,5
Núñez	2	4,7			
Total	77	21		102	20,9

Tabla 54 - Distribución de «soler + INF» por texto

En presente existe una utilización paralela de dos recursos (perífrasis más expresión temporal) en un 21 % más de casos en TRAD que en ES, como si el traductor intuyera que la utilización de «soler + infinitivo» es adecuada en este caso pero no quisiera dejar de lado el adverbio que aparece en el texto inglés. De todas formas, una revisión en profundidad de estas concordancias nos permite darnos cuenta de que esto sucede solo en el caso de dos traductores, principalmente en el de Mustieles, por lo que es probable que estemos ante un hábito personal y no un comportamiento generalizado.

<Cornwell>: - Normalmente no <suelo> buscar cebras .
 -->par_en: " I do n't usually look for zebras . "

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Usually / as usual	20	19,6
∅	12	11,8
Used to	12	11,8
Often	12	11,8
Marcador temporal	12	11,8
Enunciados generalizadores	10	9,8
Generally	8	7,8
Reformulación	4	3,9
Would	3	2,9
Tend to /make a habit of	3	2,9
Be inclined / apt to	2	2
Routinely	2	2
Wont	1	1
Typically	1	1
TOTAL	102	100

Tabla 55 - Equivalentes de «soler + INF» en EN

Como vemos, los equivalentes de esta perífrasis, a pesar de no ser formalmente similares, son muy numerosos y de variada naturaleza. *Usually* y *often* son las expresiones que activan la utilización de «soler + infinitivo» de manera más habitual en el presente, y *used to* en el pasado.

<Grafton>: Un montón de toallas en el brazo <suele> ser una buena coartada para entrar en una habitación .

-->par_en: An armload of towels usually provides good cover for getting into a room .

<Highsmith>: Bueno , y qué negocio no lo era , <solía> pensar Tom .

-->par_en: Well , whose weren't , Tom often thought .

<Leonard>: - <Solía> volar a Freeport un par de veces al mes - explicó Nicolet - .

-->par_en: " He used to fly over to Freeport a couple of times a month , " Nicolet said .

De más de 90 apariciones de *use(d) to* en el subcorpus EN de originales en inglés, sin embargo, solo 13 se traducen por «soler + infinitivo» (14,4 %). El porcentaje es mayor en el caso de los adverbios *usually* (20 de 37: 54 %) y *often* (12 de 40: 30 %).

No obstante, encontramos que en un 16,7 % de los casos, el traductor ha elegido usar «soler + infinitivo» *motu proprio*, mediante una reformulación de la expresión del original o sin un detonante identificable. En prácticamente todos los casos de adición, se utiliza el pasado o el presente de manera que se entiende por contexto el carácter habitual de la acción pero no hay un elemento textual que lo indique.

<Highsmith>: Se imaginó que Pritchard debía de acabar demasiado cansado al anochecer como para salir al bar a tomarse una última cerveza , o lo que <soliera> beber .

-->par_en: Tom supposed that Pritchard might well be too tired in the evening to come out to the bar-tabac for a late beer , or whatever he drank .

<Highsmith>: - <Suelen> ponerse conferencias ?

-->par_en: " You talk with each other-across the Atlantic ? "

<Leonard>: - Si la caga , y ya se sabe que <suele> hacerlo - añadió Melanie - , se quedarán con mi dinero , se quedarán con mi droga y se quedarán contigo .

-->par_en: " He fucks up , which he's been known to do , " Melanie said , " and they get the cash , they get my dope , and they've got you hanging .

El siguiente elemento activador en orden de frecuencia de aparición son los distintos marcadores temporales que aparecen en la frase para indicar la repetición de la acción o su prolongación en el tiempo (un 11,8 %). Se trata de adverbios como *still*, *ever* o *sometimes* y subordinadas de tiempo precedidas por las preposiciones *when* y *while*, pero también incluimos aquí el adjetivo *most* cuando se utiliza en expresiones de cantidad que también se traducen por la perífrasis (2 casos):

<Grafton>: Tenía el estómago contraído por el curioso retortijón de inquietud y emoción que <suelo> sentir cuando entro en piso ajeno ilegalmente .

-->par_en: I could feel my stomach squeeze in the same curious twist of anxiety and excitement I experience **when I 've entered** someone 's premises illegally .

<Highsmith>: Me palese muy bien - dijo Tom , adoptando su acento chino , que <solía> divertir a Heloise .

-->par_en: " Velly good , " said Tom , putting on his Chinese accent , which **sometimes** amused Heloise .

<Cornwell>: El plumón del eider común es blanco inmaculado ; **normalmente** el plumón <suele> tener un tono sucio .

-->par_en: The common eider 's down is snow-white . **Most** down is dingy . "

Como expresiones exclusivas de enunciados en tiempo presente encontramos casi todos los enunciados generalizadores (es decir, sentencias u opiniones que los hablantes consideran verdades establecidas, conocidas y compartidas por todos), la expresión «*to make a habit of*» y los adverbios *generally*, *routinely* y *typically*. No así el sinónimo de estos tres últimos *usually*, ya que, a pesar de ser más frecuente en este tiempo verbal, 6 de las 20 apariciones se encuentran en pasado.

<Leonard>: - Cuando uno está separado , no <suele> hacer eso .

-->par_en: " When you're separated , " Max said , " you **don't get to do that** .

<Cornwell>: Aunque por lo **general** las nuevas tecnologías no <suelen> aportar nada en los casos que llevan mucho tiempo cerrados , hay excepciones .

-->par_en: Though new technology **generally** is irrelevant in cases that have long been closed , there are exceptions .

<Leonard> El cartero <suele> dejarlos abajo , junto al ascensor , pero éste lo subió .

-->par_en: " The mailman **usually** leaves them downstairs by the elevator , but he brought this one up .

<Leonard>: Las demás <solían> estar ansiosas por llegar .

-->par_en: **Usually** they were anxious to get out of here .

Únicamente en enunciados en tiempo pasado, encontramos las expresiones *to be inclined to* y *to be apt to* y *to be one's wont*. No sucede lo mismo con *tend to*, que se utiliza en ambos tiempos.

<Highsmith>: El día era cálido y soleado , y Tom abrió una rendija en ambas ventanillas , esperando que el taxista no se quejara por le courant d'air , como <solían> hacer los franceses con la más ligera de las brisas.

-->**par_en**: The day was warm and sunny , and Tom opened the taxi windows a slit on both sides , hoping the driver would not complain of a courant d'air , which the French **were apt to** do at the mildest of breezes .

<Highsmith>: Tom se volvió y subió las escaleras hacia su habitación , se lavó las manos deprisa en su cuarto de baño , como <solía> hacer tras un episodio desagradable .

-->**par_en**: Tom turned and went up the stairs to his own room , washed his hands quickly in his bathroom as **was his wont** after a disagreeable episode such as the one he'd just been through .

<Cornwell>: Era tarde ; en esta época del año **siempre** <suele> ocurrir .

-->**par_en**: It was late , **tends to** be that way this time of year .

<Grafton>: <Solía> madrugar , pero a las cinco de la mañana era realmente un atropello .

-->**par_en**: He **tended to** rise early , but five a.m. was really pushing it .

Used to, como es evidente, aparece exclusivamente en pasado, mientras que *would* no se utiliza solo para indicar una acción habitual en el pasado sino también en presente condicional como una manera de hipotetizar o declarar que una situación es o sería natural o lógica:

<Grafton>: La piel le olía a productos farmacéuticos , a esa colonia que <solemos> ponernos en la sala de cuidados intensivos , compuesta a partes iguales de alcohol alcanforado , esparadrapo y yodo

-->**par_en**: His skin smelled medicinal , the sort of aftershave you **find** sport in an emergency room , composed of equal parts rubbing alcohol , adhesive tape , and suturing material .

4.4.3.8. Llevar + gerundio

Esta perífrasis de gerundio posee valor durativo progresivo —pues enmarca el desarrollo de una acción en un período de tiempo hasta un punto en el presente o pasado y prevé la continuación de la acción a partir de ese momento— y, en ocasiones, frecuentativo. Exige la utilización de complementos temporales casi siempre explícitos, como en «Lleva comiendo pizza tres días» (*She has been eating pizza for three days in a row*). Yllera (1999: 3419-20) duda de su carácter perifrástico, dada su escasa gramaticalización; el complemento temporal tiende a intercalarse entre el verbo en forma personal y el gerundio, lo que indica poca cohesión en la construcción. Sin embargo, no es posible separar los verbos en dos oraciones («lleva tres días» y «come pizza») ni responde a una pregunta con «cómo» («¿Cómo lleva tres días?»), por lo que el gerundio no tiene valor adverbial.

Es una estructura muy característica del español (Fente Gómez 1971:137); el inglés no posee fórmulas similares, pero los tiempos perfectos en forma continua combinados con expresiones de tiempo funcionan como un detonante evidente para su utilización. El español puede valerse también de la preposición «desde» («Trabajo en esa fábrica desde hace diez años») o el verbo «hacer» (Hace diez años que trabajo en esa fábrica») (García Fernández 2012: 68).

- CORPUS COMPARABLE – FASE I

Términos de búsqueda en WordSmith:

Search words: LLEV*/LLÉV*

Context words: *ANDO*/*IENDO*/*ÁNDO*/*IÉNDO*/*YENDO*/*YÉNDO*

Context Search Horizons: 0L to 6R

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Llevar + GER	9	2,8	17	3,3	-0,16	n. s.	NO

Debido a la presencia exigida de complementos temporales, que suelen encontrarse entre el auxiliar y la forma no personal, consideramos necesario ampliar el horizonte de búsqueda a seis palabras a la derecha de la palabra clave para asegurarnos de encontrar todos los casos de esta perífrasis en el corpus. A pesar de ser una construcción muy característica del español, la traducción del tiempo perfecto en forma continua combinado con un complemento temporal activa sin duda su utilización en las traducciones. Por este motivo, la diferencia entre textos originales y traducidos es mínima, y no llega a ser significativa estadísticamente. La mayor parte de los casos se dan, de nuevo, en la novela de Hunter, por lo que podría aventurarse que su traductor dispone de una mayor variedad de recursos estilísticos que los demás.

- CORPUS COMPARABLE – FASE II

Términos de búsqueda en Corpus Workbench:

[word="(¿|i)?ll(e|é)v.*" %c] []{0,6} [word=".*(a|á)ndo.*|.*(y|i)(e|é)ndo.*"];

Perífrasis	Frec. ES	Frec. ES normalizada	Frec. TRAD	Frec. TRAD normalizada	LL	P	Verificación de la hipótesis
Llevar + GER	10	2,7	13	2,7	+0,00	n. s.	NO

En esta segunda fase aumentan los casos en ES y disminuyen en TRAD, por lo que las frecuencias normalizadas se equiparan completamente y la diferencia entre ambos subcorpus sigue sin ser significativa.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Madrid	4	9,6	Leonard	5	6,4
Giménez	2	2,3	Ellroy	4	3,6
Martín	2	3,5	Cornwell	2	1,8
Martínez	1	1,8	Highsmith	2	2
Núñez	1	2,3	Grafton	0	0
Vázquez	0	0			
Total	10	2,7		13	2,7

Tabla 56 - Distribución de «llevar + GER» por texto

No se utiliza demasiado en originales, y es curioso que sea Madrid, una vez más, quien encabece la clasificación de frecuencias y, más curioso aún, el hecho de que en TRAD suceda lo mismo con Leonard. Siete de los diez ejemplos hallados en ES intercalan el complemento circunstancial de tiempo entre auxiliar y gerundio, lo que confirma que esta es la pauta y no la excepción en esta perífrasis, por lo que contamos todos los casos a excepción de los dos siguientes, en los que los complementos son demasiados y parecen incidir más sobre el auxiliar que sobre el conjunto perifrástico.

<Giménez>: <Lleva un año con nosotros siempre demostrando> buena conducta .

<Martínez>: Hay alguna disculpa para mi conducta : <llevaba seis meses en paro , sobreviviendo> malamente de préstamos piadosos , ridículos ahorros y pequeñas colaboraciones sobre ornitología , metalurgia avanzada o la última ola de pirateo informático (...)

De contar estos casos, las cifras seguirían sin ser significativas (12 en ES, 13 en TRAD; LL +0,26), por lo que el resultado no se vería alterado. En cuanto a TRAD, se cumple esta pauta en mayor medida, puesto que solo uno de los trece casos válidos no presenta el complemento o sujeto intercalado.

<Núñez>: <Llevo casi un mes viviendo> con ellas y no saben nada de mi vida .

<Leonard>: <Llevo casi veinte años sirviendo> a la gente .

-->par_en: I've been waiting on people for almost twenty years .

- CORPUS PARALELO

Elemento activador EN	Frecuencia	Porcentaje
Tiempo perfecto continuo + expresión de tiempo	8	61,5
Otros tiempos verbales + expresión de tiempo	4	30,8
Tiempo continuo solo	1	7,7
TOTAL	13	100

Tabla 57 - Equivalentes de «llevar + GER» en EN

Más del 60 % de las concordancias corresponde al patrón sugerido por las gramáticas y estudios contrastivos: tiempos perfectos continuos combinados con marcador temporal.

<Cornwell>: Hay mucha gente que <lleva años esperando> ver cómo se desangra .

-->par_en: There 's a lot of people who 've been waiting for years to see you bleed . "

<Cornwell>: Jenny y Ronnie <llevaban muchos años escribiéndose> , y a ella se le hacía imposible aceptar la idea de que lo mataran .

-->par_en: Jenny and Ronnie had been writing to each other for many years and she just could n't deal with the thought of him being put to death . "

Sin embargo, aunque es verdad que encontramos expresiones de tiempo prácticamente en todas las concordancias, vemos que otros tiempos verbales también pueden activar la utilización de esta perífrasis mientras vayan acompañados de una expresión de tiempo adecuada. Incluimos en esta categoría el participio presente del tercer ejemplo.

<Ellroy>: Están condenadas al fracaso , porque las medidas tomadas por el departamento de Policía de Los Ángeles <llevan muchos años demostrando> su eficacia .

-->par_en: They will fail , because the enforcement measures employed by the Los Angeles Police Department have proven successful for many years .

<Ellroy>: Narcóticos y la familia Kafesjian <llevan veintitantos años relacionándose> de un modo autónomo .

-->par_en: Narco and the Kafesjian family go back autonomously twenty-odd years .

<Leonard>: Durante los dos años que <llevaba viviendo> solo , una vez había estado a punto de decirle a una mujer que la quería .

-->par_en: Once , during the past two years living alone , he had almost told a woman he loved her .

Solo hay un caso en el que el detonante es únicamente una oración en *present perfect continuous*, sin presencia de marcador temporal explícito, pero el español añade esa expresión de tiempo ausente en el original, pues de lo contrario, como hemos visto, la perífrasis no podría utilizarse y habría que recurrir a otra solución de traducción.

<Leonard>: <Llevo tiempo esperando> que aparezcas .

-->par_en: I've been waiting for you to stop by . "

En el caso de esta perífrasis, la estructura activadora está tan clara que no encontramos ejemplos de adiciones o reformulaciones. Se podría aventurar que si la estructura paralela en inglés fuera menos habitual, la perífrasis se utilizaría en menor medida. Sin embargo, sabemos que esta combinación puede traducirse simplemente valiéndose de la conjugación verbal española, por ejemplo, sin necesidad de recurrir a perífrasis verbales. Por eso, en última instancia, la utilización de «llevar + gerundio» depende del hábito, del conocimiento o de la preferencia puntual del traductor. No obstante, también puede pensarse que este caso es parecido al de «venir + gerundio»: la influencia del factor 2 de la Gravitational Pull Hypothesis desempeña un papel importante, pues las formas continuas son muy prominentes en inglés, y la utilización de esta perífrasis podría responder a la necesidad del traductor de variar las estructuras utilizadas para traducirlas.

4.4.3.9. Recapitulación de resultados: grupo III

- CORPUS COMPARABLE – FASE I

Subtotal Grupo III	311	95,9	356	68,4	+18,67	< 0,0001	Sí
-------------------------------	------------	------	------------	------	---------------	----------	-----------

En el grupo de las perífrasis a medio camino, las del grupo III, confirmamos algunas intuiciones, como la de la sobrerrepresentación de las perífrasis incoativas de valor estilístico «echar a + infinitivo» y «romper a + infinitivo». Estas dos, junto con «llevar + gerundio» —en cuyo caso, el tiempo perfecto en forma continua combinado con un complemento temporal en inglés activa sin duda su utilización en las traducciones al español— y «estar a punto de + infinitivo», no confirmaron la hipótesis

de la infrarrepresentación. Por lo tanto, a pesar de que los resultados para el conjunto del grupo indican que hay una diferencia estadísticamente significativa de estas perífrasis en el subcorpus de traducciones, lo cierto es que el test de log-likelihood da positivo en solo cinco de nueve casos.

- CORPUS COMPARABLE – FASE II

Subtotal Grupo III	342	93,2	491	100,5	-1,16	n. s.	NO
-------------------------------	------------	------	------------	-------	--------------	--------------	-----------

En este tercer grupo las cosas han cambiado mucho. De contar con cinco perífrasis infrarrepresentadas («dejar de + infinitivo», «ponerse a + infinitivo», «querer + infinitivo», «parar de + infinitivo» y «soler + infinitivo»), dos sobrerrepresentadas («echar(se) a + infinitivo» y «romper a + infinitivo») y dos casos no significativos, pasamos a tener una única infrarrepresentación (se mantiene «querer + infinitivo») y seis resultados no significativos. En el caso de «dejar de + infinitivo» y «soler + infinitivo», la diferencia no llega a ser significativa, mientras que «ponerse a + infinitivo» pasa a ser un claro caso de sobrerrepresentación en traducciones. «Parar de + infinitivo» deja de estar infrarrepresentada. Las perífrasis «echar(se) a + infinitivo» y «romper a + infinitivo» continúan estando mejor representadas en el subcorpus de traducciones, pero en el caso de esta última, la diferencia deja de ser estadísticamente significativa. «Estar a punto de + infinitivo» y «llevar + gerundio» no alteran su estado.

El resultado final es que pasamos de una rotunda infrarrepresentación del grupo en general a una pequeña sobrerrepresentación no significativa estadísticamente. Es el grupo en el que el cambio ha sido más llamativo y creemos que un factor determinante podría ser que el número de novelas es demasiado reducido como para que las características individuales de cada traductor no afecten al conjunto del corpus. Queda claro, por ejemplo, en la enorme diferencia de casos hallados en «ponerse a + infinitivo» —que pasa de sumar 33 concordancias a 115, y casi todas, fruto de la pluma de un único traductor— o en «querer + infinitivo» —que pasa de un único caso en TRAD, a 15—. De lo que podemos extraer que, quizás, la muestra lingüística tendría que ser bastante más amplia para poder sacar conclusiones válidas, algo que, como hemos comentado en el capítulo de metodología, no siempre es defendido por los trabajos sobre estudios de traducción basados en corpus.

Otro motivo del cambio podría ser que al acercar el número de palabras reales de ambos subcorpus, la estadística deja de compensar, en cierto modo, la diferencia de tamaño, y podemos ver que en realidad no hay tanta diferencia entre lenguaje original y traducido. Podría resultar curioso que esto quede tan patente en las perífrasis de este tercer subcorpus, que incluye expresiones acerca de las cuales no estábamos seguros de cuál sería su comportamiento. Sin embargo, no lo es tanto si pensamos que, por el mismo motivo, podría considerarse que forman parte de un repertorio más estándar de la lengua: estaríamos hablando de expresiones que ni se utilizan mucho por tener claros elementos activadores en la lengua origen (de este caso excluiríamos a «echarse a + infinitivo» y «romper + infinitivo», las cuales hemos comprobado que podríamos haber incluido en el grupo I) ni dejan de utilizarse por no contar con ellos; sino que su uso está supeditado a otros factores; depende de los traductores y de su bagaje lingüístico la decisión de si incluirlas o no. Podríamos pensar que, cuando no hay una influencia fuerte proveniente de las estructuras de la lengua original —el segundo factor de la Gravitational Pull Hypothesis— y los elementos en cuestión son recursos habituales en la lengua de llegada y, por tanto, poseen un grado de prominencia cognitivo elevado en la mente del traductor (el primer factor de dicha hipótesis), tienden a reflejar una frecuencia similar en ambos subcorpus. Por supuesto, las diferencias en cuanto a prototipicidad entre las perífrasis presentadas en este apartado podrían ser grandes, por lo que no podemos generalizar sin realizar estudios de mayor alcance.

- CORPUS PARALELO

En cuatro casos hemos encontrado una variedad amplísima de estructuras en lengua original que se traducen por la perífrasis estudiada: en la incoativa «ponerse a + infinitivo», la inminente «estar a punto de + infinitivo», la terminativa «dejar de + infinitivo» y la frecuentativa «soler + infinitivo», todas ellas utilizadas de manera similar en traducciones y originales, según los tests estadísticos, excepto «ponerse a + infinitivo», que resultó estar sobrerrepresentada en traducciones.

En tres casos, «querer + infinitivo», «romper a + infinitivo» y «llevar + gerundio», no había tanta diversidad de patrones pero, sin embargo, resultaba muy claro cuáles eran las estructuras activadoras. El primer caso es el único de este grupo en que la perífrasis resultó infrarrepresentada en traducciones, porque, a pesar de la claridad de los patrones activadores para el resto de valores, estaba igual de claro que faltaban

estructuras en inglés que equivaliesen a dos de los valores utilizados en español. Sin embargo, al analizar la perífrasis «romper a + infinitivo», el resultado fue el contrario. A pesar de que los casos son tan pocos que es arriesgado sacar conclusiones a partir de ellos, existen dos (o más, seguramente) elementos activadores claros que parecen haber sido establecidos como equivalentes estandarizados de esta perífrasis, la cual se utiliza de manera casi automática para traducirlos, aunque en español original no tengamos constancia de su utilización.

Esto queda más claro en el caso de «echarse a + infinitivo», una de las dos perífrasis sobrerrepresentadas en traducciones aunque sin demasiada variedad de estructuras activadoras ni patrones claros, pues el 50 % de sus concordancias corresponden a adiciones. Sin embargo, si nos fijamos en dichas adiciones, casi todas son frases que incluyen las acciones de empezar a reír, correr o llorar, y activan en la traducción la perífrasis de manera automática, porque, como hemos dicho, es el recurso estándar para la expresión de estas acciones en lenguaje traducido. «Parar de + infinitivo» tampoco cuenta con patrones claros, y un gran porcentaje de concordancias (aunque no hay que olvidar que los casos son muy pocos) está conformado por adiciones y reformulaciones. A pesar de que su sinónimo «dejar de + infinitivo» se utiliza de manera mucho más frecuente, esta perífrasis no se encuentra infrarrepresentada porque es la que mejor coloca en un determinado contexto semántico, que es el de enunciados negativos con verbos de expresión.

A partir de lo expuesto, vemos que se repite la idea de que, si una perífrasis cuenta con equivalentes claros en inglés, es difícil que se encuentre infrarrepresentada. Lo mismo ocurre cuando, a pesar de no haber patrones claros, su utilización es muy flexible y puede servir como solución de traducción para muy diversas estructuras —que es lo que parece suceder con varias de las perífrasis de este grupo—, o para un valor muy específico, como en «querer + infinitivo» o «parar + infinitivo», lo que a veces puede dar lugar a una utilización mayor de lo que en principio parece normal en originales, como es el caso de «romper a + infinitivo» o «echarse a + infinitivo».

La última de las perífrasis con estructuras activadoras claras es «llevar + gerundio». Hemos visto que siempre hay un patrón concreto en la concordancia paralela, pero también que este patrón podría haberse traducido de otras maneras, lo que puede hacernos pensar que es una estructura común en español que el traductor tiende a utilizar de manera similar a un autor que escribe en español sin estar expuesto a la influencia de un texto original en otra lengua. No obstante, hemos visto también que en

algunas de las perífrasis de gerundio, el factor 2 de la Gravitational Pull Hypothesis desempeña un papel importante en la mayor utilización de dichas perífrasis en traducciones. A pesar de no contar con un elemento activador directo, la prominencia de los tiempos continuos o simplemente perfectos en inglés combinados con una expresión temporal podría estar influyendo en la decisión del traductor de no valerse de la solución de traducción más evidente («estar + gerundio») y, al diversificar las soluciones de traducción, la frecuencia de utilización de las perífrasis verbales analizadas se incrementa.

4.5. Conclusiones

En este capítulo hemos seleccionado un grupo de 27 perífrasis verbales de infinitivo y gerundio, las hemos dividido en tres grupos basándonos en nuestras expectativas de verlas o no representadas en la misma medida en traducciones al español y en originales españoles, y hemos realizado búsquedas en nuestro corpus comparable para extraer la frecuencia de aparición de cada una de dichas perífrasis en originales y en traducciones. Posteriormente, hemos realizado las mismas búsquedas en el corpus paralelo de originales en inglés, para analizar las estructuras equivalentes en lengua origen que podrían haber afectado de alguna manera la utilización de las perífrasis en español traducido.

Nuestro objetivo era averiguar si el grupo de perífrasis que consideramos típico del español, en el sentido de que sus valores y significados no se manifiestan de la misma manera o a través de estructuras similares en inglés, presentan una frecuencia de aparición menor en traducciones de dicha lengua a español que la que puede encontrarse en textos escritos originalmente en español.

El primer análisis que realizamos lo hicimos sobre un corpus comparable de novela negra y policiaca en español compilado sin tener en cuenta ciertos parámetros referentes a origen de los autores y traductores y diferencia en número de palabras. En este primer análisis, los resultados fueron contundentes a favor de nuestra hipótesis, la infrarrepresentación de elementos típicos en traducciones. Sin embargo, para la realización de esta tesis, como explicamos en el capítulo correspondiente, creímos necesario aumentar el número de palabras del subcorpus de originales para intentar

igualarlo al de traducciones, modificar la composición del subcorpus de traducciones para incluir una única variedad lingüística tanto en español (la de España) como en inglés (la estadounidense), y añadir el componente de originales en inglés del que partían las traducciones utilizadas.

En esta nueva versión del corpus, hemos vuelto a confirmar la infrarrepresentación de las perífrasis como conjunto, aunque con un margen de error estadístico un poco mayor al de la versión anterior. Sin embargo, podríamos afirmar que con mayor seguridad, puesto que esta infrarrepresentación se produce a pesar de haber igualado el número de palabras de los subcorpus.

En este caso, la infrarrepresentación se constata solo en 11 de las 27 perífrasis, en contraposición con las 17 de la versión anterior del corpus. No obstante, la infrarrepresentación general se entiende si tenemos en cuenta, por un lado, que el número de perífrasis sobrerrepresentadas sigue siendo más reducido (aunque hay 2 más que antes) y, por otro lado, que aumenta el número de casos no significativos (de 7 a 11), pero que 7 de ellos corresponden a perífrasis que, a pesar de estar mejor representadas en originales, no llegan a pasar el test de significación estadística. Esto significa que encontramos efectivamente una menor utilización de estas perífrasis en el subcorpus de traducciones pero la diferencia con el subcorpus de originales españoles no resulta lo suficientemente importante como para tenerla en cuenta en términos de rigor estadístico.

Esta reducción en el número de perífrasis que confirman la hipótesis de la infrarrepresentación constata la importancia de contar con subcorpus de similar tamaño, pues al acercar los números de palabras, lo que más ha aumentado es el número de casos no significativos; es decir, se han reducido las diferencias estadísticamente significativas entre subcorpus. Esto quiere decir que no existe tanta diferencia como pensábamos entre lenguaje original y traducido; la estadística probablemente realizaba las diferencias para compensar el desequilibrio en número de palabras. Sucede lo mismo si miramos los resultados por grupos. En la primera versión del corpus había una infrarrepresentación significativa en los tres grupos. En la segunda versión, solo queda la infrarrepresentación significativa del grupo II, pues pasamos a tener dos sobrerrepresentaciones no significativas, en el grupo I y el II.

Pasamos a resumir los resultados grupo por grupo, dado que nuestras expectativas con respecto a los resultados de cada uno de ellos diferían. El grupo I

estaba formado por perífrasis con equivalentes, en principio, evidentes en inglés, que esperábamos que estuvieran bien representadas en traducciones por ese motivo. Sin embargo, cinco de las perífrasis de este grupo resultaron infrarrepresentadas, por no tener equivalentes claros según el análisis de nuestro corpus paralelo («acabar + gerundio» y, en el caso de «terminar por + infinitivo», debido también a la preferencia por su estructura sinónima en traducciones), o por contar con una estructura alternativa de significado similar pero de uso más generalizado («comenzar a + infinitivo», «continuar + gerundio»). En general, no resultaron infrarrepresentadas aquellas perífrasis con una variedad amplia de estructuras equivalentes en inglés capaces de activar su utilización en traducciones, como «empezar a + infinitivo» y «volver a + infinitivo» —que son también recursos muy habituales en español original—, las que no se utilizan prácticamente en español original y, por ello, cualquier aparición en TRAD se magnifica («terminar + gerundio») o las que, a pesar de no ser muy frecuentes, según nuestro subcorpus, por algún motivo se prefieren por sobre sus estructuras sinónimas (como «acabar por + infinitivo»). Nos sorprendió la infrarrepresentación de «seguir + gerundio», puesto que pensábamos que existirían varios equivalentes inmediatos en inglés que activarían su utilización en traducciones. Sin embargo, observamos que la utilización de esta perífrasis se producía también en casos sin un elemento activador claro en el texto original. Estas dos circunstancias, sin embargo, no bastaron para que la perífrasis alcanzara la frecuencia esperada según el modelo del corpus de originales en español.

En el segundo grupo, el de las perífrasis que esperábamos que presentaran una clara infrarrepresentación en traducciones por no contar con estructuras formalmente equivalentes en inglés que pudieran funcionar como estímulos para su utilización al traducir, solo cinco de nueve perífrasis cumplieron estas expectativas. Resultaron infrarrepresentadas aquellas que no tenían equivalente claramente determinado en la lengua de partida («andar + gerundio»), a pesar de servir algunas de traducción a estructuras muy diversas («haber que + infinitivo», «llegar a + infinitivo» e «ir + gerundio»), así como «acertar a infinitivo», cuyos valores, no demasiado frecuentes en español, son expresados por otras estructuras en traducciones. La perífrasis que contaba con equivalentes claros en inglés no resultó infrarrepresentada («haber de + infinitivo»), mientras que algunas, a pesar de no contar con dichos equivalentes claros, tampoco mostraron diferencias con el lenguaje de los originales puesto que en español original no son en absoluto frecuentes («alcanzar a + infinitivo» y «venir a + infinitivo»). «Venir

+ gerundio» tampoco es frecuente en español original según nuestros datos, pero la prominencia de la estructura equivalente más habitual en la lengua de partida hace que esta perífrasis se utilice en traducciones con una frecuencia similar e incluso algo más elevada que en ES, aunque esta pequeña diferencia no sea estadísticamente significativa.

El tercer grupo es el que menos infrarrepresentación presenta, lo que resulta llamativo, dado que, en principio, esperábamos que esto sucediera con el grupo I. Quizás sea, justamente, porque se trata de perífrasis sin claros equivalentes pero que son de uso bastante habitual en español, por lo que se utilizan para traducir estructuras diversas de manera espontánea aunque el texto original no predisponga a los traductores a ello. Cuatro contaban con estructuras equivalentes en el original de muy diversa naturaleza («ponerse a + infinitivo», «dejar de + infinitivo», «soler + infinitivo» y «estar a punto de + infinitivo») y dos, con elementos activadores claros («romper a + infinitivo» y «llevar + gerundio»). «Echar a + infinitivo» y «parar de + infinitivo» son equivalentes estandarizados para determinados valores, y la única perífrasis infrarrepresentada lo está porque, a pesar de contar con equivalentes fácilmente accesibles, algunos de los valores que hemos seleccionado no se encuentran con una frecuencia similar en inglés («querer + infinitivo»).

A modo de resumen, en la tabla 58 encontramos los distintos factores que, basándonos en los resultados obtenidos, parecen influir en la mayor o menor utilización de las perífrasis en textos traducidos. Como vemos, muchos de los factores se repiten, y alguno puede conducir tanto a una infrarrepresentación en traducciones como a frecuencias similares en ambos subcorpus o incluso a una sobrerrepresentación. Esto es así porque, por ejemplo, cuando encontramos una amplia variedad de estructuras equivalentes en inglés, puede significar tanto que la perífrasis sea un recurso sumamente habitual en español que se utilice para traducir multitud de estructuras y resolver diversos problemas de traducción, como es el caso de, por ejemplo, «empezar a + infinitivo», «soler + infinitivo» o «volver + infinitivo» (que, por tanto, contarán con gran número de concordancias), como que no haya un estímulo claro en inglés que la active al traducir y su utilización dependa de si al traductor se le ocurre utilizarla en determinados casos o bajo la influencia de algún estímulo, por lo que no encontraremos patrones fijos, y la perífrasis tendrá relativamente pocas concordancias (p. ej. «ir + gerundio»).

INFRARREPRESENTACIÓN		NO INFRARREPRESENTACIÓN	
FACTOR	EJEMPLO	FACTOR	EJEMPLO
Falta de patrones equivalentes claros	Terminar por + INF; Acabar + GER, Andar + GER	Patrones equivalentes claros	Haber de + INF, Romper a + INF, Llevar + GER
Amplia variedad de estructuras equivalentes + Falta de patrones equivalentes claros	Llegar a + INF, Haber que + INF, Ir + GER	Amplia variedad de estructuras equivalentes	Ponerse a + INF, Estar a punto de + INF, Dejar de + INF, Soler + INF
Existencia en español de una estructura sinónima de uso más habitual	Comenzar a + INF, Continuar + GER	Amplia variedad de estructuras equivalentes + Recurso habitual en español	Empezar a + INF, Volver a + INF
Estructura poco habitual en español original + Falta de patrones equivalentes claros	Acertar a + INF	Estructura poco habitual en español original + Falta de patrones equivalentes claros + Preferencia del traductor	Venir + GER, Venir a + INF, Alcanzar a + INF, Terminar + GER,
Ausencia en inglés del valor expresado por la perífrasis	Querer + infinitivo	Poca variedad de estructuras equivalentes + Patrones equivalentes claros	Echar a + INF, Parar de + INF
Escasez de estructuras activadoras + Existencia en español de estructuras sinónimas más similares a las inglesas	Seguir + GER	Preferencia del traductor en detrimento de otra perífrasis equivalente	Acabar por + INF

Tabla 58 - Factores que influyen en la utilización de las distintas perífrasis en TRAD

En cuanto a la combinación de factores «Estructura poco habitual en español original + Falta de patrones equivalentes claros», que se repite en ambas columnas, tenemos tanto el caso extremo de «acertar a + infinitivo», que aparece en español pero no cuenta con ninguna concordancia en el corpus paralelo, como las tres («venir + gerundio», «venir a + infinitivo» y «alcanzar a + infinitivo») que presentan una frecuencia de utilización muy reducida en español y, por tanto, si un traductor tiene cierta preferencia por utilizarlas, la diferencia entre subcorpus se incrementa de inmediato. Una vez más, consideramos importante contar con subcorpus lo suficientemente grandes y variados para reducir la posibilidad de no hallar ejemplos de algunas perífrasis y para que determinadas cuestiones de idiosincrasia de autores o traductores no afecten excesivamente a los resultados.

Si seguimos a Hareide (2013: 35-40), no obstante, podemos enlazar los factores que hemos observado en este trabajo con los que Halverson menciona como causantes de una posible sobrerrepresentación o infrarrepresentación, y que hemos transcrito en la introducción teórica. Cualquier afirmación sobre prototipicidad en la lengua de partida o de llegada solo puede ser tentativa, ya que debería confirmarse a través de estudios

específicos sobre cada elemento. Nos basaremos fundamentalmente en los datos de frecuencias arrojados por nuestro análisis y en la información que hemos podido obtener a través de la lectura de los distintos trabajos consultados sobre los elementos analizados en este capítulo, para intentar analizar qué factores pueden haber sido los más influyentes a la hora de provocar la infrarrepresentación de las perífrasis analizadas.

A partir de los datos obtenidos, podríamos decir que dichas perífrasis cumplen una o varias de las siguientes condiciones:

- ✓ no son elementos prototípicos de la lengua meta o no se utilizan con frecuencia, ya que hay otros equivalentes de significado y funciones similares (factor 1), o
- ✓ no cuentan con elementos equivalentes lo suficientemente prominentes en la red conceptual de la lengua de partida que puedan ejercer su influencia en la selección del elemento (factor 2), o
- ✓ no existen enlaces directos o claros entre la perífrasis y los conceptos relacionados de la lengua de partida (factor 3).

Según lo que se desprende de nuestro análisis, podríamos agrupar las perífrasis infrarrepresentadas según si cumplen o no —limitándonos a los resultados de nuestro análisis— los supuestos de la definición de Hareide de *unique item*:

- has no direct counterpart in the source language: no en el caso de «comenzar a + infinitivo», «seguir + gerundio», «continuar + gerundio», «terminar por + infinitivo» y «acabar + gerundio» (que pertenecen al grupo con equivalentes claros en inglés); sí en el de «haber que + infinitivo», «acertar a + infinitivo», «llegar a + infinitivo», «andar + gerundio», «ir + gerundio» y «querer + infinitivo» (6 de 11).
- has no salient or prototypical source language structure that may impact its choice in the target language: no en el caso de «comenzar a + infinitivo», «haber que + infinitivo», «andar + gerundio» (creemos que determinadas estructuras —como las incoativas, las de obligación o las reiterativas— sí son prominentes en inglés, aunque luego puedan no traducirse por nuestras perífrasis); sí en el de «seguir + gerundio», «continuar + gerundio», «terminar

por + infinitivo», «acabar + gerundio», «acertar a + infinitivo», «querer + infinitivo», «llegar a + infinitivo», «ir + gerundio» (hemos visto que las estructuras inglesas que pueden activarlas son muy variadas y no hay una que destaque especialmente) (8 de 11).

- It is characterized by being infrequent and non-prototypical: dado que es difícil delimitar el grupo de elementos dentro del cual deberíamos establecer la prototipicidad de una determinada perífrasis, pues los valores que expresan son muy diversos, nos basaremos entonces en la frecuencia de uso: no en el caso de «seguir + gerundio», «haber que + infinitivo», «llegar a + infinitivo» e «ir + gerundio»¹⁴⁰, puesto que en ambas fases del subcorpus de originales ES han superado la media de frecuencia de uso de perífrasis —que en nuestro corpus es 43,6 en la fase I y 56,8 en la fase II—, o bien entendemos que son la forma más habitual de expresar el valor que poseen (o una de las formas más habituales en el caso de «haber que + infinitivo»); sí en el de «comenzar a + infinitivo»¹⁴¹, «continuar + gerundio», «andar + gerundio», «terminar por + infinitivo», «acabar + gerundio», «acertar a + infinitivo» y «querer + infinitivo» (7 de 11).

Como vemos, los tres supuestos no se cumplen en todos los casos (salvo para «acertar a + infinitivo» y «querer + infinitivo»). Recordemos que en el grupo I, el de las perífrasis con equivalentes claros, también ha habido bastantes casos de infrarrepresentación. Por tanto, uno, dos o los tres factores juntos están ejerciendo, en el caso de estas perífrasis, una fuerza conjunta hacia la infrarrepresentación. Sucede a la inversa en el caso de las perífrasis sobrerrepresentadas. Por ejemplo, en «haber de + infinitivo» solo hemos registrado un factor, el de la existencia de patrones equivalentes claros, que coincidiría con la tercera condición de la definición proporcionada por Hareide. La fuerza ejercida por este único factor es, al parecer, suficiente para contrarrestar los otros dos. En algunos casos, por otra parte, como hemos visto en la cuarta fila de la segunda columna de la tabla 80, la voluntad del traductor o la traductora (que puede tener causas muy diversas, tanto externas como de su propio bagaje lingüístico o cultural) contrarresta todos los factores y la perífrasis no resulta infrarrepresentada a pesar de todo.

¹⁴⁰ Según la RAE (2009: 2186), se trata de la segunda perífrasis de gerundio más frecuente después de «estar + gerundio».

¹⁴¹ Según la RAE (2009: 2173) posee índices de frecuencia más bajos pero iguales propiedades que «empezar a + infinitivo».

Es evidente que no es solo el texto original o la lengua de partida lo que ejerce influencia sobre las cifras registradas, sino diversos motivos que pueden ser inherentes al proceso de traducción o, simplemente, al traductor. Como hemos observado una y otra vez a lo largo de este análisis, en última instancia, las decisiones dependen de las actitudes, el bagaje lingüístico y cultural y las preferencias de la persona que traduce.

	EMPEZAR	COMENZAR	PONERSE A	ECHARSE A	ROMPER A
ES	Núñez 71 Giménez 64 Vázquez 33 Martín 12 Madrid 12 Martínez 8	Martínez 37 Madrid 12 Vázquez 2 Núñez 2 Giménez 1 Martín 1	Giménez 13 Madrid 11 Martínez 9 Vázquez 5 Martín 4 Núñez 1	Vázquez 13 Giménez 12 Núñez 6 Martín 3 Martínez 2 Madrid 1	-
TRAD	Cornwell 143 Leonard 60 Highsmith 48 Ellroy 35 Grafton 26	Grafton 26 Cornwell 1	Grafton 86 Leonard 9 Cornwell 8 Ellroy 7 Highsmith 5	Highsmith 23 Leonard 23 Grafton 18 Ellroy 11 Cornwell 10	Grafton 1 Highsmith 1
INFRARREPRESENTACIÓN	N. S.	Sí	NO	NO	N. S.

Tabla 59 - Comparación de las frecuencias de aparición de las cinco perífrasis incoativas

Detengámonos un momento en las perífrasis incoativas que hemos estudiado en los diferentes grupos y analicémoslas en conjunto, para observar un fenómeno interesante. En la tabla 59, vemos que los escritores españoles utilizan «empezar a + infinitivo» en mayor proporción que las demás. Sin embargo, dos de seis se decantan por «comenzar a + infinitivo», y los segundos y terceros puestos son ocupados alternativamente por «echarse a + infinitivo» y «ponerse a + infinitivo». «Comenzar a + infinitivo» predomina entre los últimos puestos a excepción de los dos autores mencionados.

Sin embargo, en TRAD vemos un patrón bastante más marcado: todos los traductores menos el de Grafton se decantan por «empezar a + infinitivo» (n. s.) en primer lugar y «echarse» en segundo (sobrerrepresentada). «Ponerse a + infinitivo» (sobrerrepresentada) aparece casi siempre en los últimos puestos (excepto en el caso de Grafton) y «comenzar a + infinitivo» (infrarrepresentada) apenas aparece en una ocasión aparte de la de Grafton. Por supuesto, como ya hemos aclarado, «romper a + infinitivo» (n. s.) aparece en el último puesto en dos ocasiones a pesar de no figurar en el repertorio de perífrasis de los autores españoles de nuestro corpus.

¿Por qué nos fijamos en este detalle? Dado que estábamos hablando de factores que podrían estar influyendo en la utilización de nuestras perífrasis, nos pareció

interesante apuntar que este uso más estereotipado de las perífrasis incoativas (puesto que existe más uniformidad en la selección de las que aparecen en traducciones que en las de los originales) podría significar que nos encontramos ante un caso de *levelling out*, el candidato a universal de traducción —también llamado «convergencia» (Baker 1996, Laviosa 2002)—, que sugeriría que puede existir mayor homogeneidad con respecto a ciertos factores como densidad léxica o longitud media de oraciones entre las traducciones de un corpus comparable determinado que entre los textos originales de dicho corpus escritos en el mismo idioma. También hemos apuntado ya, por un lado, al universal de explicitación (Blum-Kulka 1986, Baker 1996) en el caso de ciertas perífrasis, como «empezar a + infinitivo», «volver a + infinitivo» o «ponerse a + infinitivo», cuyo uso parte, en ocasiones, de adiciones al texto original motivadas por el impulso de hacer más explícitas ciertas relaciones sintácticas o semánticas del texto original, y, por otro, al universal de simplificación (Baker 1996, Laviosa 1997) en el caso de la variedad de verbos con los que se combina la perífrasis «echar a + infinitivo» en originales y traducciones. En estos casos, estaríamos hablando de un factor que depende del proceso de traducción en sí, y no de la influencia de factores contrastivos como estructuras concretas de la lengua de partida en la selección de una u otra perífrasis o la utilización en mayor o menor medida en traducciones. Probablemente, nuestros resultados sean una combinación de factores de ambos tipos.

Nos quedamos con el hecho de que las perífrasis verbales analizadas siguen arrojando datos que confirman que, en conjunto, estos elementos se utilizan en menor medida en textos traducidos. Nuestra clasificación en grupos, sin embargo, no ha resultado demasiado fiable a la hora de predecir resultados. En cada uno de los grupos ha habido algunas que han cumplido con nuestras expectativas y otras que no, por los diversos factores que hemos ido mencionando. En cuanto a una posible división según los valores que adoptan las distintas perífrasis (incoativo, perfectivo, etcétera), tampoco vemos posible realizar una distinción clara o una clasificación posible de acuerdo a los resultados obtenidos. Reorganicemos la tabla del principio —«Tabla 5 – Comparación de frecuencias entre subcorpus ES y TRAD y significación estadística de los resultados (FASE II)»— según las diferentes categorías que han surgido de los resultados del test estadístico:

- Grupo I: si la perífrasis aparece más frecuentemente en textos originales y, además, esta mayor frecuencia es estadísticamente significativa; es decir, si está infrarrepresentada en traducciones (en verde).
- Grupo II: si la perífrasis aparece más frecuentemente en textos originales pero esta mayor frecuencia no es estadísticamente significativa (en amarillo).
- Grupo III: si la perífrasis aparece más frecuentemente en textos traducidos pero esta mayor frecuencia no es estadísticamente significativa (en fucsia).
- Grupo IV: si la perífrasis aparece más frecuentemente en textos traducidos y, además, esta mayor frecuencia es estadísticamente significativa; es decir, si está sobrerrepresentada en traducciones (en rojo).

Perífrasis	Total ES	Frecuencia ES normalizada	Total TRAD	Frecuencia TRAD normalizada	LL	P	Verificación de la hipótesis
Grupo I							
Acabar + GER	38	10,3	22	4,5	+10,11	< 0,01	SÍ
Andar + GER	35	9,5	26	5,3	+5,15	< 0,05	SÍ
Haber que + INF	159	43,3	75	15,3	+59,61	< 0,0001	SÍ
Terminar por + INF	7	1,9	0	0	+11,85	< 0,001 ¹⁴²	SÍ
Ir + GER	148	40,3	80	16,4	+44,67	< 0,0001	SÍ
Comenzar a + INF	55	15	27	5,5	+19,42	< 0,0001	SÍ
Querer + INF	26	7,1	15	3,1	+6,97	< 0,01	SÍ
Continuar + GER	24	6,5	14	2,9	+6,30	< 0,05	SÍ
Acertar a + INF	7	1,9	0	0	+11,85	< 0,001 ¹⁴³	SÍ
Llegar a + INF	66	18	48	9,8	+10,32	< 0,01	SÍ
Seguir + GER	170	46,3	183	37,4	+3,95	< 0,05	SÍ
Subtotal Grupo I	735	200,2	490	100,3	+144,31	< 0,0001	SÍ
Grupo II							
Dejar de + INF	95	25,9	101	20,7	+2,46	n. s.	NO
Parar de + INF	12	3,3	9	1,8	+1,72	n. s.	NO
Terminar + GER	9	2,4	4	0,8	+3,67	n. s.	NO
Estar a punto de + INF	42	11,4	49	10	+0,39	n. s.	NO
Acabar por + INF	7	1,9	5	1	+ 1,15	n. s.	NO
Soler + INF	77	21	102	20,9	+0,00	n. s.	NO
Llevar + GER	10	2,7	13	2,7	+0,00	n. s.	NO
Subtotal Grupo II	252	68,6	283	57,9	+3,83	n. s.	NO
Grupo III							
Romper a + INF	0	0	2	0,4	-2,24	n. s. ¹⁴⁴	NO
Empezar a + INF	200	54,5	312	63,9	-3,11	n. s.	NO
Venir + GER	1	0,3	7	1,4	-3,51	n. s. ¹⁴⁵	NO
Venir a + INF	1	0,3	6	1,2	-2,68	n. s. ¹⁴⁶	NO
Subtotal Grupo III	202	55	327	66,9	-4,86	< 0,05	NO
Grupo IV							
Volver a + INF	199	54,2	378	77,4	-17,04	< 0,0001	NO
Echar(se) a + INF	37	10,1	85	17,4	-8,17	< 0,01	NO
Ponerse a + INF	43	11,7	115	23,5	-16,69	< 0,0001	NO
Haber de + INF	63	17,2	131	26,8	-8,86	< 0,01	NO
Alcanzar a + INF	1	0,3	10	2	-6,20	< 0,05 ¹⁴⁷	NO
Subtotal Grupo IV	343	93,4	719	147,2	-50,19	< 0,0001	NO
Total	1532	417,3	1819	372,3	+10,79	< 0,01	SÍ

Tabla 60 - Reorganización de los grupos de perífrasis según los resultados arrojados por el test de LL

¹⁴² Test exacto de Fisher: 0,002675 (p < 0,01)

¹⁴³ Test exacto de Fisher: 0,002675 (p < 0,01)

¹⁴⁴ Test exacto de Fisher: 0,510077 (n. s.)

¹⁴⁵ Test exacto de Fisher: 0,149489 (n. s.)

¹⁴⁶ Test exacto de Fisher: 0,25095 (n. s.)

¹⁴⁷ Test exacto de Fisher: 0,02973 (p < 0,05)

Resulta curioso ver de manera gráfica como, en los dos grupos intermedios, muchos casos no significativos llegan a serlo (grupo III) o casi (grupo II —dado que el umbral de significación está establecido en 3,84) al ser tomados en conjunto. De ahí nuestras dudas iniciales acerca de la fiabilidad de sumar los resultados de distintas perífrasis. Pero la tabla también nos permite comprobar visualmente por qué se da esta infrarrepresentación general: son más las perífrasis en los grupos con infrarrepresentación significativa en traducciones que con sobrerrepresentación, y también son más las perífrasis con diferencias no significativas que, aun así, aparecen con frecuencias mayores en originales. Si solo tuviéramos en cuenta los casos significativos los resultados serían los siguientes, más significativos si cabe:

Perífrasis	Total ES	Frecuencia ES normalizada	Total TRAD	Frecuencia TRAD normalizada	LL	P	Verificación de la hipótesis
Grupo I							
Acabar + GER	38	10,3	22	4,5	+10,11	< 0,01	SÍ
Andar + GER	35	9,5	26	5,3	+5,15	< 0,05	SÍ
Haber que + INF	159	43,3	75	15,3	+59,61	< 0,0001	SÍ
Terminar por + INF	7	1,9	0	0	+11,85	< 0,001 ¹⁴⁸	SÍ
Ir + GER	148	40,3	80	16,4	+44,67	< 0,0001	SÍ
Comenzar a + INF	55	15	27	5,5	+19,42	< 0,0001	SÍ
Querer + INF	26	7,1	15	3,1	+6,97	< 0,01	SÍ
Continuar + GER	24	6,5	14	2,9	+6,30	< 0,05	SÍ
Acertar a + INF	7	1,9	0	0	+11,85	< 0,001 ¹⁴⁹	SÍ
Llegar a + INF	66	18	48	9,8	+10,32	< 0,01	SÍ
Seguir + GER	170	46,3	183	37,4	+3,95	< 0,05	SÍ
Subtotal Grupo I	735	200,2	490	100,3	+144,31	< 0,0001	SÍ
Grupo IV							
Volver a + INF	199	54,2	378	77,4	-17,04	< 0,0001	NO
Echar(se) a + INF	37	10,1	85	17,4	-8,17	< 0,01	NO
Ponerse a + INF	43	11,7	115	23,5	-16,69	< 0,0001	NO
Haber de + INF	63	17,2	131	26,8	-8,86	< 0,01	NO
Alcanzar a + INF	1	0,3	10	2	-6,20	< 0,05 ¹⁵⁰	NO
Subtotal Grupo IV	343	93,4	719	147,2	-50,19	< 0,0001	NO
Total	1078	293,7	1209	247,5	+16,62	< 0,0001	SÍ

Tabla 61 - Comparación de frecuencias de los grupos I y IV

Chesterman (2007: 11-12) propone una manera alternativa de hallar un *unique item* sin tener que recurrir a una definición a priori de niveles formales de equivalencia. En primer lugar, utilizar estudios contrastivos para determinar qué elementos muestran una frecuencia de aparición diferente en textos traducidos y originales; en nuestro caso, las perífrasis de este trabajo. A continuación, dividir aquellas que presenten diferencias de uso en dos grupos: las que aparecen más frecuentemente en textos traducidos (las 9 de los grupos III y IV) y las que lo hagan en textos originales (las 18 de los grupos I y

¹⁴⁸ Test exacto de Fisher: 0,002675 (p < 0,01)

¹⁴⁹ Test exacto de Fisher: 0,002675 (p < 0,01)

¹⁵⁰ Test exacto de Fisher: 0,02973 (p < 0,05)

II). En último lugar, ordenar los elementos según el grado de diferencia de frecuencia entre subcorpus y concentrarse en los que muestren las diferencias más grandes sobre un umbral establecido de manera arbitraria. Once perífrasis no presentan diferencias estadísticamente significativas entre subcorpus (las 6 del grupo II y las 5 del III). Si establecemos el umbral en una diferencia de, por ejemplo, más de 10 puntos en el test de log-likelihood (percentil 99,9; significativo al nivel de 0,1 %; $p < 0,001$; valor crítico = 10,83)¹⁵¹, nos quedamos únicamente con 5 perífrasis claramente infrarrepresentadas en las traducciones y 2 claramente sobrerrepresentadas. Con respecto a estas últimas, una pertenece al grupo con equivalentes claros («volver a + infinitivo») y otra al grupo III original, es decir, el de perífrasis en situación intermedia («ponerse a + infinitivo»). En cuanto a las infrarrepresentadas, 3 son del grupo II original, es decir, el que incluía aquellas perífrasis sin equivalentes evidentes en la lengua de partida («ir + gerundio», «haber que + infinitivo» y «acertar a + infinitivo») y 2 son del grupo I original, es decir, con equivalentes claros en la lengua de partida («comenzar a + infinitivo» y «terminar por + infinitivo»). Si tenemos en cuenta el test de Fisher, no podríamos contar a «acertar a + infinitivo» ni a «terminar por + infinitivo».

¿Serían estos los mejores candidatos a elemento único? Habría que realizar más estudios con distintos corpus para averiguarlo, puesto que, por ejemplo, la inclusión de «ponerse a + infinitivo» en esta selección de elementos significativos y, específicamente, en la categoría de la sobrerrepresentación, es un hecho puntual y particular de nuestro corpus que se debe a su utilización excesiva por parte de un único traductor. La perífrasis «haber que + infinitivo», no obstante, podría considerarse un ejemplo incuestionable de la influencia que tiene la falta de estructuras paralelas en inglés en las decisiones del traductor.

¹⁵¹ Valores consultados en Rayson 1993-2010 [fecha de consulta: 10/03/2010].

5. Demostrativos

5.1. Demostrativos: definición y caracterización

La Real Academia Española (RAE), en su *Nueva gramática de la lengua española* (2009: 49) afirma que «(...) el español es peculiar en el hecho de manifestar tres grados de proximidad en el paradigma de los demostrativos, frente a otras muchas lenguas que los limitan a dos». Presenta como ejemplos de este fenómeno el francés (*celui-ci ~ celui-là*), el italiano (*questo ~ quello*) o el rumano (*acest ~ acel*), pero, sin ir más lejos, el catalán (Klee y Lynch 2009: 57) también cuenta con una distinción dicotómica de sus demostrativos (*aquest* y *aquell*) y locativos (*aquí* y *allà*)¹⁵². Este hecho justifica nuestro interés en uno de los grados de dicho paradigma tripartito como candidato a elemento único del español para este estudio.

El inglés, por su parte, también cuenta con un sistema de demostrativos de solo dos grados, como puede comprobarse en cualquier gramática de la lengua inglesa (Thomson y Martinet 1987: 23; Sinclair et ál. 1996: 35, 48; Mackenzie y Martínez Caro 2012: 189). En este capítulo se pretende comprobar si la infrarrepresentación de elementos únicos en textos traducidos se produce en la combinación inglés-español con respecto a alguno de los elementos de la estructura tripartita del sistema de demostrativos —*este, ese* y *aquel* (tanto determinantes como pronombres) y sus correspondientes formas neutras y flexionadas—, por un lado, y de los adverbios locativos *aquí, ahí* y *allí* y sus variantes, por otro.

Dado que la división formal de nuestra lengua de contraste, el inglés, es bipartita (*this* y *that*; *here* y *there*) y que, en principio, no hay motivo para pensar que el primer grado de lejanía (*este* y *aquí*) vaya a estar infrarrepresentado en traducciones, puesto que la equivalencia con los términos ingleses *this* y *here* es *a priori* bastante clara, nuestra hipótesis es que alguno de los dos elementos que en español pueden

¹⁵² Según Macías Villalobos (2006: 150), el catalán poseía al principio un paradigma demostrativo ternario constituido por las formas *aquest/aqueix/aquell*, cuyo elemento intermedio, *aqueix*, ha ido cayendo en desuso en casi todo el dominio lingüístico catalán. Este fenómeno no se ha dado en valenciano (Ainaud, Espunya y Pujol 2003: 181) que conserva los tres grados (*este/eixe/aquell* y *ací/ahí/allí/allà*) debido a su mayor contacto con el español (Klee & Lynch 2009: 60 citando a Blas Arroyo 2000) o el catalán antiguo (Montiel Ponsoda 1999: 77 [nota 2]). Hablamos del uso espontáneo, pues la normativa (Montiel Ponsoda 1999: 58) sigue describiendo tres formas.

corresponder a cada una de las formas que en teoría expresan mayor lejanía deíctica en inglés, es decir, *that* y *there* (*ese* y *aquel* o *ahí* y *allí*, respectivamente), podría verse favorecido en las traducciones, en detrimento del otro. Estudios anteriores sobre demostrativos que han descrito el sistema español (Macías Villalobos 2006; Eguren Gutiérrez 1999: 940; RAE 2009: 1280-81; de Kock y Gómez Molina 1992: 19, 20) o lo han contrastado con otras lenguas (Goethals 2013: 528; Labrador 2011: 83) dejan constancia de que *ese* es una forma que aparece con mayor frecuencia que *aquel* en textos tanto escritos como orales y que, en ciertos ámbitos, la utilización de este último se encuentra cada vez más restringida. Por lo tanto, nuestra hipótesis de trabajo es que el tercer grado de lejanía del español podría encontrarse infrarrepresentado en las traducciones debido a que el segundo grado del inglés tiene que repartirse entre ambos y, a partir de los datos con los que contamos, creemos que es poco probable que esta repartición se lleve a cabo de manera proporcional entre las dos formas. No disponemos de datos similares concernientes a los adverbios demostrativos, por lo que, en su caso, la hipótesis no puede ser tan específica. Sin embargo, en la norma peninsular encontramos al menos dos formas de elevada frecuencia de uso para la deixis de lejanía (*ahí* y *allí*), por lo que una de ellas podría encontrarse infrarrepresentada en detrimento de la otra.

5.1.1. *Determinantes y pronombres demostrativos*

Los demostrativos son un grupo de palabras que se utiliza para hacer referencia a uno o varios elementos en un contexto espacial, temporal o textual determinado. Pueden tener función deíctica, cuando señalan a un elemento del contexto extraverbal, fórica, cuando hacen referencia al contexto textual-discursivo, o ambas a la vez, dado que los límites entre lo que se suele considerar deixis o anáfora no siempre están claros, como explican Botley y McEnery (2001: 7). Por ejemplo, si en un fragmento de discurso directo nos encontramos con la frase siguiente: «Ya le había dicho a ese tipo que se largara», puede ser que el *ese* funcione como elemento anafórico, pues el hablante podría haberlo mencionado antes en el discurso, pero también es posible que esté señalando a «ese tipo» mientras habla, con lo que el demostrativo ejercería una función deíctica.

Dejando dichos casos dudosos de lado por el momento, en principio, podemos afirmar que sus funciones (Gómez Torrego 2002; Macías Villalobos 2006; RAE 2009) son principalmente las de señalar la ubicación espacial de un elemento en relación con el emisor del enunciado (deixis espacial), señalar una ubicación de naturaleza temporal, es decir, hacer referencia a un punto concreto en el tiempo, cercano o alejado del momento en el que el hablante emite el enunciado (deixis temporal) y, también, hacer referencia a un elemento concreto del texto que ya se ha mencionado (antecedente) o que se mencionará a continuación (consecuente): son los usos anafóricos o catafóricos.

Macías Villalobos (2006: 176) añade a estas tres funciones principales los usos estilísticos o expresivos —como los valores de aproximación afectiva («¡Qué chico este!»), los despreciativos («el bicho ese») o la deixis evocadora, en la que se alude a algo remoto o imaginario, sobre todo mediante el demostrativo *aquel*—; los casos de anadeixis (2006: 72) —es decir, cuando es difícil delimitar si el señalamiento es espacio-temporal o fórico—; y, dentro de la deixis espacial, incluye también la textual, cuyo referente no es extralingüístico sino que forma parte del discurso. La distingue de la anáfora explicando (2006: 56) que el antecedente de la expresión deíctica remite a su vez a un objeto real, mientras que en la deixis textual la referencia se hace únicamente a una entidad intratextual, en un uso metalingüístico (p. ej.: «¿Has dicho *déjà vu*? Nunca he sabido pronunciar esta palabra»). Por otro lado, a pesar de mantener la distinción tradicional entre deixis y anáfora para los fines prácticos de su trabajo, considera que los usos anafóricos y catafóricos son simplemente otra posible manifestación de la deixis, que es el fenómeno general de señalamiento de la lengua (2006: 72). En este sentido, estamos de acuerdo con Goethals (2013: 519) cuando afirma:

Demonstratives are typical deictic expressions because they locate a referent vis-à-vis a deictic center, and thus also evoke, or establish, this deictic center. Since the deictic center can be evoked both in exophoric and anaphoric relations, terminologically speaking, deictic and anaphoric are categories that are not mutually exclusive.

Existe un amplio debate al respecto de esta cuestión entre los distintos autores. Por ejemplo, Eguren Gutiérrez (1999: 936) defiende lo siguiente:

La anáfora, por su parte, no es un tipo de deixis. (...) Las anáforas o las catáforas «tienen el mismo referente» que otro término o expresión que ha aparecido antes, o que aparecerá después, en el discurso. La anáfora consiste, por tanto, en una relación de ‘correferencia’.

Sin embargo, dado que la RAE también parece estar de acuerdo con la opinión de Macías Villalobos y Goethals, puesto que señala en su *Nueva gramática de la lengua española* (2009: 1272) que las referencias fóricas son manifestaciones más abstractas del fenómeno de la deixis, así lo consideraremos en este trabajo.

Por último, los demostrativos pueden encontrarse dentro de construcciones lexicalizadas y semilexicalizadas (RAE 2009: 1332-35). En ocasiones no poseen valor deíctico ni anafórico sino sentido idiomático, pero a veces conservan un valor anafórico incluso a pesar de estar lexicalizados en locuciones. Por ejemplo, mientras que *este* y *esto* se utilizan como muletillas para rellenar pausas en la conversación, *eso* en la expresión «eso es todo» y *esas* en «ni por esas» mantienen su valor anafórico.

Las tres formas básicas de los demostrativos, tradicionalmente relacionadas con las tres personas gramaticales y el grado de proximidad o lejanía con respecto al emisor, son las siguientes:

- ESTE: según las gramáticas consultadas, se utiliza para referirse a una entidad cercana al emisor, ya sea físicamente o mentalmente, y, con respecto al tiempo, al presente o al pasado y al futuro cercanos. En usos anafóricos, a algo o a alguien que se acaba de mencionar o al elemento más cercano cuando en el discurso se menciona más de un antecedente.
- ESE: se utiliza para referirse a un objeto más alejado del emisor o en la cercanía inmediata del receptor (o a una distancia intermedia entre emisor y receptor). También al futuro y al pasado.
- AQUEL: se utiliza para referirse a elementos que están lejos (física o mentalmente) del hablante o distantes tanto del emisor como del receptor, para hablar del pasado lejano o para señalar el elemento más alejado cuando en el discurso se mencionan dos antecedentes.

A pesar de lo que parecen ser unas clasificaciones claramente estructuradas en la teoría, es común detectar en la práctica la alteración de estas pautas. El hablante se expresa a menudo de acuerdo con su propia escala deíctica, y puede utilizar *aquel* para referirse a elementos cercanos de los que intenta desvincularse de algún modo o *este* para señalar algo que está más alejado pero considera cercano en términos afectivos. Los motivos de este fenómeno pueden ser emocionales, pueden tener que ver con cuestiones de solidaridad o empatía o depender del contexto o las circunstancias de la interacción, de manera que los demostrativos pueden adoptar un papel modalizador que no siga a rajatabla las reglas gramaticales (Eguren Gutiérrez 1999: 935, 941; Borrero Barrera y Cala Carvajal 2002 y RAE 2009: 1282).

Algunos autores, como de Kock y Gómez Molina (1992: 67-68) —que en su trabajo se refieren solo a los pronombres—, consideran que el sistema deíctico, tradicionalmente dividido en tres grados, en realidad ha evolucionado para configurarse solo en dos. Así ha ocurrido, de hecho, como explica Macías Villalobos (2006: 150)¹⁵³, con la mayor parte de las lenguas romances menos el portugués. Según esta postura, *ese* habría pasado a ser un elemento no marcado capaz de adoptar valores de cercanía o lejanía o neutralización (RAE 2009: 1280; Macías Villalobos 2006: 107).

Macías Villalobos (2006: 106) defiende que esto sucede porque el nuestro es un sistema demostrativo sumamente flexible en el que la referencia principal es el hablante, a partir de cuyos criterios subjetivos el espacio puede organizarse en un sistema bipartito (lo que cae dentro de la esfera del hablante y lo que queda fuera de ella) o tripartito, como hemos explicado al principio. Aun así, este autor (2006: 155) considera que, en la práctica, el tiempo y los valores expresivos suelen adoptar una organización bipolar (el presente del hablante frente al no-presente y la proximidad afectiva al que habla frente al alejamiento afectivo).

Con respecto a la clasificación sintáctica de estos elementos, existe un debate aún abierto y con diversas ramificaciones. Según las gramáticas tradicionales y trabajos más antiguos, los demostrativos pueden clasificarse como adjetivos o pronombres. Se clasifican como adjetivos si van acompañando a un sustantivo; si lo sustituyen, son pronombres. Asimismo, las formas neutras (*eso*, *esto* y *aquello*), que se utilizan para referirse a hechos, objetos desconocidos, conceptos abstractos o proposiciones enteras,

¹⁵³ Según Macías Villalobos, el español ha conservado la estructura ternaria heredada del latín, mientras que en el resto de lenguas romances, el sistema deíctico se volvió binario. Así, en italiano, la distinción entre *questo/codesto/quello* pasó a ser entre *questo* y *quello* y el francés distingue entre *ce-ci* y *ce-là* e *ici* y *là*. Algo parecido sucede con el ruso y el alemán (Montiel Ponsoda 1999: 61-63).

pueden ser únicamente pronombres, puesto que no existen como adjetivos demostrativos. Entre otros autores que utilizan las denominaciones de *adjetivo* o *adjetivo determinativo* encontramos a de Kock y Gómez Molina (1992), Sánchez Benedito (1990), Marcos Marín (1975), Gili Gaya (1981) o Seco (1989). En inglés, encontramos el término *demonstrative adjective*, por ejemplo, en Thomson y Martinet (1987), en Stockwell, Bowen y Martin (1965) o en la *Collins Cobuild English Grammar* (Sinclair et ál. 1996), que recoge ambas denominaciones, la de adjetivo y la de determinante. Algunos textos, en general más modernos, los consideran un tipo específico de «determinante», como por ejemplo los de Gómez Torrego (1999, 2002) o Marcos Marín, Satorre Grau y Viejo Sánchez (1998). En inglés hablan de *demonstrative determiners* Close (1985), Chalker (1984) y Mackenzie y Martínez Caro (2012).

Por otro lado, hay autores que no diferencian entre formas adjetivas o determinantes y pronominales. Por ejemplo, Alarcos (1995: 89) considera que no es necesario distinguir entre estos dos tipos de palabra, pues todos los adjetivos pueden funcionar como sustantivos mediante la sustantivación. Por su parte, Macías Villalobos (2006: 113) no está de acuerdo con el estudio en dos paradigmas separados, dado que considera que el paradigma de las formas demostrativas es único, mientras que lo que cambia son las funciones. No obstante, en el análisis de su corpus distingue siempre entre ocurrencias de formas pronominales y adjetivas para realizar un estudio lo más completo posible. La RAE (2009: 1278) reconoce que el tema no está resuelto, pero se decanta por considerarlos *determinantes* porque, a pesar de que, como los adjetivos, concuerden en género y número con el sustantivo, los adjetivos no convierten al sustantivo en una expresión referencial, como sí lo hacen, según esta gramática, el artículo y los demostrativos. Además, escoge la opción simplificadora, es decir, la no duplicación de las clases de demostrativos en determinantes y pronombres, pues considera que los demostrativos pronominales no neutros deben interpretarse como determinantes que inciden sobre elementos nulos o tácitos. Por tanto, los únicos pronombres demostrativos desde este punto de vista serían *esto*, *eso* y *aquello*.

En nuestro caso, sí distinguiremos entre ambas funciones sintácticas, opción más tradicional pero considerada igualmente válida por la RAE, porque consideramos que esta distinción puede aportar un análisis más detallado de los datos y un parámetro más de comparación a la hora de extraer conclusiones, si bien es verdad que en ciertos momentos hemos tenido dificultades para la clasificación de algunos ejemplos, como

explicaremos más adelante. La denominación que utilizaremos, por consiguiente, será la de *pronombres demostrativos y determinantes demostrativos*.

Estos pronombres y determinantes demostrativos adoptan diversas formas según el género y el número del objeto al que hacen alusión. Así, pueden clasificarse como se muestra en la tabla 62.

Forma	ESTE		ESE		AQUEL	
Clase sintáctica Género y número	DET	PRO	DET	PRO	DET	PRO
Masc. singular	Este	Este	Ese	Ese	Aquel	Aquel
Fem. singular	Esta	Esta	Esa	Esa	Aquella	Aquella
Masc. plural	Estos	Estos	Esos	Esos	Aquellos	Aquellos
Fem. plural	Estas	Estas	Esas	Esas	Aquellas	Aquellas
Neutro	-	Esto	-	Eso	-	Aquello

Tabla 62 - Clasificación de los determinantes y pronombres demostrativos en español

En inglés, las formas equivalentes serían las que vemos en la tabla 63.

Forma	THIS		THAT	
Clase sintáctica Número	DET	PRO	DET	PRO
Singular	This	This	That	That
Plural	These	These	Those	Those

Tabla 63 - Clasificación de los determinantes y pronombres demostrativos en inglés

Como vemos, frente a las quince formas diferentes del español, el inglés cuenta solo con cuatro, contraponiendo *this* (*este*) a *that* (*ese* o *aquel*), es decir, utilizando solo dos grados de proximidad, según el objeto al que se hace referencia se encuentre cercano a la persona que habla o no. Para transmitir el matiz de *aquel*, algunos autores afirman que se hace uso de paráfrasis como «that one (way) over there». Dado que nuestro corpus no posee un cuarto componente de traducciones de español a inglés, no podremos comprobar si esto es así, pero sí podremos ver cuáles son los elementos ingleses que los traductores eligen expresar mediante este demostrativo.

5.1.2. *Adverbios demostrativos locativos*

A pesar de que Macías Villalobos (2006: 131) no los considera demostrativos, pues, a diferencia de estos, son invariables en cuanto a género y número (como todos los adverbios) y su paradigma no está cerrado, la RAE denomina a las formas que estudiaremos bajo este epígrafe *adverbios demostrativos* —de tipo locativo, dado que también los hay de tiempo (p. ej. *ayer*), de grado (p. ej. *tanto*) y de manera (p. ej. *así*)— y afirma que coinciden con los pronombres y determinantes demostrativos en su naturaleza deíctica (Eguren Gutiérrez 1999: 931; RAE 2009: 1311).

Eguren Gutiérrez (1999: 958) diferencia entre dos subsistemas, uno ternario (*aquí, ahí y allí*) y uno binario (*acá y allá*), y resume las razones que llevan a algunos gramáticos a hacer esta distinción en tres puntos: por un lado, la serie ternaria indica una localización estática frente a la binaria, que expresa movimiento o dirección; por otro lado, la primera identifica lugares concretos mientras que la segunda indica una localización más vaga; por consiguiente, y en tercer lugar, la serie binaria admite modificadores de grado como *más* o *tan*. Sin embargo, cada región hispanohablante posee usos lingüísticos particulares relacionados con estos adverbios, por lo que no es posible diferenciarlos claramente a través de los rasgos semánticos mencionados. Por ejemplo, en algunas zonas de países hispanohablantes como Argentina, *aquí* y *allí* apenas se utilizan, pues suenan mucho más formales o afectados que los más habituales *acá* y *allá*.

En este trabajo consideraremos que el español presenta una división tripartita en la que *aquí, ahí y allí* y sus variantes se contraponen a la división formal dicotómica del inglés (*here* y *there*).

- ◆ AQUÍ (equivalente de *here*), o *acá* en algunos usos o regiones hispanohablantes, se utiliza para señalar lo que se encuentra en la cercanía inmediata (física o emocional) del emisor.
- ◆ AHÍ (equivalente aproximado de *there*), se utiliza si hablamos de elementos más distantes del emisor o, según algunos autores, más cercanos al interlocutor.
- ◆ ALLÍ (equivalente aproximado de *there*), se utiliza cuando se habla de elementos aún más distantes del emisor o, según algunos autores, alejados tanto del emisor como del receptor. En estos casos también se utiliza *allá*, aunque,

como hemos mencionado, algunos autores atribuyen a esta forma —como a *acá*— un valor añadido de movimiento o direccionalidad (Bello 1978: 140; RAE 2009: 1315).

Como sucede en el caso de los pronombres y determinantes, los hablantes pueden alterar los valores enunciados en la teoría para expresar un punto de vista subjetivo. Hablamos, por ejemplo, de lo que la RAE (2009: 1313) denomina ubicación en un «espacio analógico», que permite a un hablante señalar un punto en un mapa y poder utilizar tanto *aquí* como *allí* para referirse a él, según si lo considera un punto ubicado en su mismo espacio o en el territorio real representado en dicho mapa.

Además del hecho de que en ciertas partes de Latinoamérica predomina la utilización de *acá* en lugar de *aquí* y de *allá* en lugar de *allí*, también podemos mencionar que el uso del demostrativo *aquel* está cada vez menos extendido. En su lugar, se tiende a utilizar el sistema binario *este/ese* y el demostrativo *aquel* queda reservado para los usos literarios o la deixis evocadora (Eguren Gutiérrez 1999: 940; RAE 2009: 1280-81). Dado que uno de los criterios de compilación de nuestro corpus es la procedencia geográfica de los escritores y traductores, lo esperable es no encontrar representados estos fenómenos que, en principio, no se consideran característicos de la variedad peninsular.

5.2. Características del corpus y métodos de análisis

5.2.1. Corpus utilizado para el análisis

Para este estudio hemos utilizado tanto el componente comparable como el componente paralelo de nuestro corpus, pero únicamente en su segunda versión. Por lo tanto, no podremos realizar comparaciones con la antigua configuración del corpus, como hicimos en el capítulo del análisis de las perífrasis.

Al tratarse de unidades gramaticales y no léxicas, y con poca variación morfológica, pues en total no hay más de quince formas diferentes de demostrativos en español, nos pareció interesante aprovechar esta oportunidad para comprobar la

fiabilidad del etiquetado del programa TreeTagger de manera más sistemática. Veremos a continuación que los errores de etiquetado son numerosos, lo que no nos permitió en este trabajo tampoco omitir la revisión manual de los resultados.

Como hemos comentado en el capítulo de metodología, consideramos que corregir manualmente el etiquetado y volver a realizar las búsquedas era duplicar el trabajo y el tiempo requerido para realizarlo, puesto que las mismas búsquedas que nos permiten detectar los errores son las que arrojan los resultados necesarios para nuestro análisis. Por este motivo, nuestra decisión en su momento fue no hacer esta doble tarea, pero sí dejar constancia de los errores que hemos ido hallando, algo que podría resultar útil en trabajos futuros, tanto en el caso de que nosotros decidiéramos corregir el etiquetado para volver a utilizar este mismo corpus, como en el caso de que otros investigadores decidieran etiquetar sus corpus con este programa.

5.2.2. Fiabilidad y utilidad del etiquetado

La búsqueda de concordancias se llevó a cabo en un primer momento tanto en el corpus original sin etiquetar como en el etiquetado, para asegurarnos de que la asignación de etiquetas se había realizado correctamente y de que los programas de análisis léxico recogían todas las apariciones de los elementos que buscábamos en ambos casos, permitiéndonos, de esta forma, poder utilizar uno u otro corpus según nuestras necesidades.

El etiquetador TreeTagger incluye la etiqueta «pronombre demostrativo» («DM»), que, según el conjunto de etiquetas que se proporciona junto al programa, busca tanto determinantes como pronombres, puesto que recoge las formas con y sin acento diacrítico. La búsqueda [pos="DM"] sobre el etiquetado del subcorpus ES arroja 3368 concordancias: 1070 formas de *este* (entre las cuales hallamos las siguientes formas acentuadas: 46 *ésta*, 5 *ésta*s, 70 *éste* y 6 *ésto*s), 1595 de *ese* (ninguna forma acentuada) y 703 de *aquel* (4 *aquél*).

Si realizamos las búsquedas sin apoyarnos en el etiquetado, es decir mediante expresiones regulares, los resultados nos indican que hay un total de 3670 formas demostrativas y, por tanto, una diferencia de 302 ocurrencias entre un cómputo y otro (17 en *este*, 178 en *ese* y 107 en *aquel*). La tabla 64 muestra las expresiones regulares

que hemos usado para la segunda búsqueda, la cantidad de demostrativos hallada en cada caso y las formas acentuadas extraídas.

"(¿ i)?(e é)st(e o a)(s)?" %c	1087	(46 <i>ésta</i> , 5 <i>ésta</i> s, 73 <i>éste</i> y 6 <i>ésto</i> s)
"(¿ i)?(e é)s(o a)(s)?" %c (e é)se" %c ¹⁵⁴	1773	(42 <i>é</i> sa, 7 <i>é</i> sas, 74 <i>é</i> se y 11 <i>é</i> sos)
"(¿ i)?aqu(e é)l(l)?(o a)?(s)?" %c	810	(5 <i>aquél</i> y 14 <i>aquélla</i>)

Tabla 64 - Resultados de la búsqueda de demostrativos en ES mediante expresiones regulares

En el caso de *ese* (v. tabla 64, segunda fila) la práctica totalidad de la diferencia se explica por el etiquetado incorrecto de todas las formas acentuadas. En 131 casos podemos ver que el programa no reconoce las formas acentuadas como demostrativos y las etiqueta alternativamente como adjetivos («ADJ») y nombres propios («NP») —sobre todo a *é*sa—, nombres comunes («NC») —sobre todo a *é*se y *é*sos— y hasta verbos («VLfin»).

Query: COMP_ESOR_BAR; "Ésa";

#

1. 23114: in a/PREP hostias/NC ./FS <Ésa/ADJ> es/Vsfin la/ART ley/NC .
2. 35533: RT sanción/NC ./FS -/DASH <Ésa/NP> es/Vsfin otra/QU histori
3. 87814: in muerto/ADJ ./FS -/DASH <Ésa/NP> es/Vsfin una/ART gran/AD
4. 92776: ntificó/VLfin ?/FS -/DASH <Ésa/NP> es/Vsfin una/ART buena/A
5. 104715: PREP progresar/VLinfin ./FS <Ésa/ADJ> fue/Vsfin la/ART señal/N
6. 105422: in ./CM Cifuentes/NC ./FS <Ésa/ADJ> es/Vsfin la/ART verdad/N
7. 148727: asé/VLfin dentro/ADV ./FS <Ésa/NP> era/VEfin la/ART sede/NC
8. 159060: /PPX separemos/VLfin ./FS <Ésa/ADJ> es/Vsfin mi/PPO vida/NC
9. 188563: P repetición/NC ?/FS)/RP <Ésa/NC> es/Vsfin otra/QU ./FS Yo
10. 231426: DV conseguirla/VLinfin ./FS <Ésa/ADJ> fue/Vsfin mi/PPO segunda
11. 251336: cia/NC inusitada/ADJ ./FS <Ésa/ADJ> es/Vsfin una/ART caracte
12. 344633: pitalista/ADJ ./FS -/DASH <Ésa/NP> es/Vsfin la/ART coartada

#

Query: COMP_ESOR_BAR; "ésa";

#

1. 2374: ouflés/NC ./CM era/VEfin <ésa/ADJ> una/ART práctica/NC que/
2. 11175: e/ADV ha/VHfin sido/Vsadj <ésa/ADJ> mi/PPO filosoffa/NC ./FS
3. 25114: REP poder/VMinfin ./FS Y/CC <ésa/ADJ> solía/VLfin ser/VSinfin la
4. 37206: cocina/NC árabe/ADJ ./CM <ésa/ADJ> es/Vsfin otra/QU razón/N
5. 57914: aparecer/VLinfin ./FS Y/CC <ésa/ADJ> fue/Vsfin la/ART razón/N
6. 63723: /DASH Sí/PPX ./CM de/PREP <ésa/ADJ> que/CQUE siente/VLfin cu
7. 77670: amiseta../NC ./FS ¿era/NC <ésa/ADJ> la/ART compañera/NC de/P
8. 91395: ente/ADJ también/ADV ./CM <ésa/ADJ> era/VEfin una/ART mujer/
9. 93718: /FS -/DASH ¿Y/NC cómo/ADV <ésa/ADJ> chica/ADJ ha/VHfin podid
10. 97447: ual/ADJ ./CM y/CC de/PREP <ésa/NC> ya/ADV tenía/VLfin todos
11. 98337: in nada/QU fácil/ADJ ./CM <ésa/ADJ> es/Vsfin la/ART verdad/N
12. 106371: gues/VLfin la/ART caja/NC <ésa/ADJ> ./FS Oye/VLfin ./CM ¿has
13. 109852: Ladj carteras/NC como/ADV <ésa/ADJ> a/PREP casi/ADV todos/QU
14. 125277: uelita/NP ./CM como/CSUBX <ésa/ADJ> de/PREP que/CQUE e/ART
15. 145917: CM a/PREP la/ART cueva/NC <ésa/ADJ> de/PREP Isaac/NP de/PREP
16. 181183: ué/INT misión/NC es/Vsfin <ésa/ADJ> ./CM pero/CCAD la/ART pr
17. 205950: mpoco/ADV ./FS Chúpate/NC <ésa/ADJ> ./CM marrana/NC ./FS Sí/
18. 208891: Linf trabajo/NC ./FS Y/CC <ésa/ADJ> será/Vsfin la/ART histor
19. 222329: emostrearle/VLinfin que/CQUE <ésa/VLfin> no/NEG era/VEfin mi/PPO
20. 243951: Carmen/NP ./FS Pero/CCAD <ésa/ADJ> no/NEG estaba/VEfin disp
21. 290140: e/CQUE parezca/VLfin ./CM <ésa/ADJ> fue/Vsfin la/ART imagen/
22. 315034: suspirando/NC -/DASH ./CM <ésa/ADJ> es/Vsfin la/ART pregunta
23. 320266: n/PREP fin/NC ;/SEMICOLON <ésa/ADJ> fue/Vsfin la/ART imbecil
24. 324365: CM vaya/VLfin ;/SEMICOLON <ésa/ADJ> es/Vsfin la/ART teoría/N
25. 329795: P proposición/NC es/Vsfin <ésa/ADJ> ?/FS -/DASH No/NEG te/PP
26. 331142: cia/NC ./FS Bien/ADV ./CM <ésa/ADJ> es/Vsfin la/ART historia
27. 331566: M Celia/NP ./FS Pero/CCAD <ésa/ADJ> es/Vsfin tu/PPX película

¹⁵⁴ Separamos esta expresión en dos partes para evitar la extracción de concordancias del sustantivo *eses*.

28. 392970: n /COLON Lorenzo/NP /CM <ésa/ADJ> ha/VHfin de/PREP ser/VSi
 29. 403114: pesetas/NC /FS Que/CQUE <ésa/VLfin> era/VEfin su/PPO tarifa/

#-----
 # Query: COMP_ESOR_BAR; "ésas";
 #-----

1. 29735: cral/ADJ /FS ¿Eran/VLfin <ésas/ADJ> las/ART etapas/NC de/PRE
2. 76424: pornográficas/NC de/PREP <ésas/NC> raras/ADJ /FS -/DASH ¿D
3. 141853: /FS -/DASH ¿Ves/NC a/PREP <ésas/NC> ?/FS Pues/CSUBF trabajan
4. 248382: /CARD mujeres/NC como/ADV <ésas/ADJ> /FS Venci/VLfin los/ART
5. 314278: S -/DASH ¿Con/NC que/CQUE <ésas/NC> tenemos/VLfin ?/FS Pues/
6. 315155: Una/ART placa/NC de/PREP <ésas/NC> es/Vsfin fácil/ADJ de/CS
7. 348318: iterario/ADJ /FS En/PREP <ésas/NC> circunstancias/NC la/ART

#-----
 # Query: COMP_ESOR_BAR; "Ése";
 #-----

1. 14139: o/NC de/PREP puta/NC /FS <Ése/NC> era/VEfin hasta/PREP el/
2. 39529: /NC de/PREP plata/NC /FS <Ése/NC> era/VEfin el/ART hilo/NC
3. 46820: Lfin una/ART copa/NC /FS <Ése/NC> hubiera/VHfin sido/VSadj
4. 62344: s/CARD millones/CARD /FS <Ése/NC> sería/Vsfin el/ART últim
5. 70090: M más/ADV astuto/ADJ /FS <Ése/NC> sería/Vsfin el/ART autén
6. 78221: eñor/NC Masderius/NP ?/FS <Ése/NC> era/VEfin un/ART caso/NC
7. 104074: í/PPX /CM hombre/NC /FS <Ése/NC> tan/ADV chuleta/NC él/PP
8. 104317: /PREP tu/PPX vida/NC /FS <Ése/NC> era/VEfin el/ART tema/NC
9. 123145: e/NC sinvergüenza/NC ?/FS <Ése/NC> viene/VLfin a/PREP la/AR
10. 136632: u/PPX historia/VLfin /FS <Ése/NC> es/Vsfin el/ART trato/NC
11. 143363: adj de/PREP ir/VLinf /FS <Ése/NC> era/VEfin el/ART único/A
12. 146946: /ART urbanización/NC /FS <Ése/NC> ha/VHfin sido/VSadj siem
13. 177979: /CM y/CC dirá/VLfin «/SYM <Ése/NC> es/Vsfin »/SYM /CM que/
14. 187899: CARD nada/QU más/ADV /FS <Ése/NC> es/Vsfin el/ART punto/NC
15. 209290: REP la/ART Navaja/NC ?/FS <Ése/NC> sí/ADV que/CQUE sería/VS
16. 216655: er/Vsinf escritor/NC /FS <Ése/NC> hace/VLfin el/ART bisnes
17. 246183: dicho/QU Bonafoux/NC ?/FS <Ése/NC> era/VEfin un/ART periodi
18. 252988: NC de/PREP ellos/PPX /FS <Ése/NC> era/VEfin mi/PPO error/N
19. 350379: NP ?/FS -/DASH No/PE /FS <Ése/NC> es/Vsfin un/ART coche/NC
20. 353080: CM el/ART Conesal/NC /FS <Ése/NC> sabe/VLfin hasta/PREP ec
21. 382231: EP carne/NC viva/ADJ /FS <Ése/NC> era/VEfin muy/ADV amigo/
22. 382677: ormaciones/NC /FS -/DASH <Ése/NC> no/NEG sabe/VLfin ni/CCN
23. 427488: merece/VLfin /FS -/DASH <Ése/NC> sí/ADV que/CQUE lo/PPC t

#-----
 # Query: COMP_ESOR_BAR; "ése";
 #-----

1. 629: hecho/VLadj siendo/Vsger <ése/NC> mi/PPO segundo/ORD matri
2. 4968: DV del/PDEL brazo/NC /CM <ése/NC> es/Vsfin el/ART guión/NC
3. 7544: abe/NC quién/INT es/Vsfin <ése/NC> caballero/NC que/CQUE ac
4. 49616: un/ART polizonte/NC /CM <ése/NC> hubiera/VHfin sido/VSadj
5. 56132: arla/VLinf /CM pero/CCAD <ése/NC> es/Vsfin un/ART pensamie
6. 63885: /ART boca/NC /CM en/PREP <ése/NC> mismo/ADJ momento/NC /C
7. 102759: co/VLfin al/PAL viejo/ADJ <ése/NC> /CM pero/CCAD no/NEG sé
8. 106951: stá/VEfin el/ART chalé/NC <ése/NC> ?/FS -/DASH Creo/VLfin q
9. 114185: DASH /FS ¿No/NC es/Vsfin <ése/NC> Tomi/NP ?/FS Me/PPX puse
10. 117129: Linf dónde/ADV está/VEfin <ése/NC> /CM eh/NP ?/FS Yo/PPX n
11. 117205: Lfin dónde/ADV está/VEfin <ése/NC> de/PREP la/ART cazadora/
12. 120887: do/VLadj del/PDEL memo/NC <ése/NC> de/PREP Ventura/NC -/DAS
13. 123245: é/INT asunto/NC era/VEfin <ése/NC> tan/ADV grave/ADJ y/CC m
14. 133146: M ya/ADV sabes/VLfin /CM <ése/NC> de/PREP la/ART calle/NC
15. 136697: Mfin whisky/NC como/CSUBX <ése/NC> /CM una/ART mujer/NC co
16. 142573: /FS El/ART Jesucristo/NP <ése/NC> te/PPX dio/VLfin la/ART
17. 143753: C ?/FS -/DASH Sí/PPX /CM <ése/NC> tan/ADV raro/ADJ de/PREP
18. 157626: RT último/ADJ día/NC /CM <ése/NC> fue/Vsfin su/PPO error/N
19. 158907: /FS Ah/NP /CM no/PE /CM <ése/NC> se/SE jode/VLfin /FS El
20. 160662: PREP coches/NC como/CSUBX <ése/NC> en/PREP la/ART ciudad/NC
21. 160732: PREP coches/NC como/CSUBX <ése/NC> en/PREP la/ART ciudad/NC
22. 169283: CC supongo/VLfin que/CQUE <ése/NC> era/VEfin un/ART incenti
23. 183428: /CM dicen/VLfin que/CQUE <ése/NC> es/Vsfin el/ART límite/N
24. 204251: C tal/QU vez/NC sea/Vsfin <ése/NC> el/ART único/ADJ modo/NC
25. 205868: del/PDEL frenillo/NC /CM <ése/NC> y/CC no/NEG otro/QU /CM
26. 210307: ué/NC hombre/NC era/VEfin <ése/NC> ?/FS La/ART mirada/NC de
27. 212745: era/VEfin ?/FS /CM y/CC <ése/NC> es/Vsfin el/ART punto/NC
28. 215707: «/SYM ¿quién/ADV es/Vsfin <ése/NC> ?/FS ¿quién/ADV es/Vsfin
29. 215711: ?/FS ¿quién/ADV es/Vsfin <ése/NC> ?/FS »/SYM /FS Tú/PPX i
30. 215857: -/DASH ¿quién/NC es/Vsfin <ése/NC> ?/FS -/DASH Un/ART escri
31. 216640: -/DASH ¿quién/NC es/Vsfin <ése/NC> ?/FS -/DASH Un/ART escri
32. 239155: Hfin dicho/VLadj que/CQUE <ése/NC> era/VEfin mi/PPO territo
33. 282363: /FS Quizá/ADV sea/Vsfin <ése/NC> el/ART motivo/NC por/PRE
34. 298078: S -/DASH ¿Qué/NC mamón/NC <ése/NC> Verbenas/NP /FS .../DOT
35. 300841: rabajabas/VLfin para/PREP <ése/NC> -/DASH giró/VLfin la/ART
36. 306407: /PREP particular/ADJ /CM <ése/NC> /FS El/ART del/PDEL cem
37. 320530: a/ADV abajo/ADV /FS Y/CC <ése/NC> fue/Vsfin su/PPO saludo/
38. 323492: dj /FS -/DASH Pues/CSUBF <ése/NC> es/Vsfin el/ART camino/N
39. 330553: re/PREP nosotros/PPX /CM <ése/NC> fue/Vsfin su/PPO saludo/
40. 342663: ente/ADV hervido/ADJ /CM <ése/NC> sería/Vsfin el/ART defec
41. 345782: fin aguantar/VLinf a/PREP <ése/NC> /CM ni/CCNEG a/PREP su/
42. 348163: le/PPC pasa/VLfin a/PREP <ése/NC> ?/FS -/DASH Es/Vsfin el/
43. 361697: /FS -/DASH ¿Escritor/NC <ése/NC> ?/FS -/DASH Como/CSUBX h
44. 365073: ero/CCAD no/NEG fue/Vsfin <ése/NC> el/ART talante/NC adopta

45. 374003: do/VLadj de/PREP que/CQUE <ése/NC> era/VEfin el/ART plato/N
 46. 402333: /PREP casa/NC /CM ¿es/NC <ése/NC> su/PPO criterio/NC ?/FS
 47. 416454: Manzanque/NC ?/FS ¿Y/NC <ése/NC> quién/INT es/VSfin ?/FS
 48. 416681: EP un/ART bacalao/NC /CM <ése/NC> es/VSfin el/ART título/N
 49. 422642: desierto/ADJ ?/FS ¿Es/NC <ése/NC> el/ART ganador/NC ?/FS L

#-----
 # Query: COMP_ESOR_BAR; "Ésos";
 #-----
 1. 344288: ra/CSUBI matar/VLinf /FS <Ésos/NC> consideran/VLfin a/PREP
 2. 415523: in para/PREP nada/QU /FS <Ésos/NC> son/VSfin unos/ART muert

#-----
 # Query: COMP_ESOR_BAR; "ésos";
 #-----
 1. 90070: CC eficaz/ADJ /CM a/PREP <ésos/ADJ> me/PPX los/PPC conozco/V
 2. 106524: /NC adosado/VLadj de/PREP <ésos/NC> /CM con/PREP jardincito
 3. 137358: ebes/VLfin ser/NC de/PREP <ésos/NC> /FS -/DASH No/NEG sé/VL
 4. 174696: Sfin sicópata/NC de/PREP <ésos/NC> /CM o/CC no/NEG ?/FS Yo
 5. 222230: sted/PPX uno/CARD de/PREP <ésos/NC> /FS Negué/VLfin con/PRE
 6. 281996: P púas/VLfin /CM de/PREP <ésos/NC> que/CQUE se/SE utilizan/
 7. 287218: de/PREP grandeza/NC /CM <ésos/NC> se/SE perdieron/VLfin co
 8. 328392: e/PPX dices/VLfin de/PREP <ésos/NC> ?/FS Eran/VLfin de/PREP
 9. 409678: P Vargas/NP Llosa/NP y/CC <ésos/NC> son/VSfin nombres/NC con

Por otra parte, los demostrativos que van precedidos por signos de interrogación o exclamación (40 casos), estén o no acentuados, tampoco se reconocen como tales.

Query: COMP_ESOR_BAR; "(¿!)(e)és(e)l(a)(s)?" "%c|"(¿!)(e)ése" %c;
 #-----
 1. 8795: desnudos../NC /FS -/DASH <¿Eso/NC> es/VSfin impropio/ADJ de
 2. 11945: adj las/ART tetas/NC ?/FS <¿Eso/NC> quiere/VLfin decir/VLinf
 3. 37343: imitación/NC /FS -/DASH <¿Eso/NC> es/VSfin ridículo/ADJ !/
 4. 41815: estigación/NC /FS -/DASH <¿Eso/NC> es/VSfin lo/ART que/CQUE
 5. 43973 fin mucho/ADV /FS -/DASH <¿Eso/NC> le/PPC ha/VHfin comentad
 6. 50541: l/ART dato/NC /FS -/DASH <¿Eso/NC> lo/PPC hace/VLfin espec
 7. 60712: ligaciones/NC /FS -/DASH <¿Eso/NC> también/ADV os/VLadj lo/
 8. 75553: UBF no/NEG /CM pero/CCAD <¿eso/NC> le/PPC parece/VLfin dese
 9. 75613: omo/CSUBX testigo/NC ?/FS <¿Esa/VLfin> sí/ADV que/CQUE es/VSfin
 10. 77151: rendido/VLadj /FS -/DASH <¿Eso/NC> es/VSfin lo/ART que/CQUE
 11. 87101: NC segura/ADJ ?/FS -/DASH <¿Esa/NC> periodista/NC tiene/VLfi
 12. 90151: /CM Petra/NC ?/FS -/DASH <¿Eso/NC> le/PPC ha/VHfin parecido
 13. 99553: ntactan/VLfin /FS -/DASH <¿Eso/NC> es/VSfin magnífico/ADJ !
 14. 100019: existencia/NC /FS -/DASH <¿Eso/NC> es/VSfin lo/ART que/CQUE
 15. 104054: ra/QU mano/NC -/DASH /FS <¿Ese/NC> ?/FS -/DASH Sí/PPX y/CC
 16. 116136: a/ART tele/NC ?/FS -/DASH <¿Eso/NC> qué/INT quiere/VLfin dec
 17. 117452: /CM señor/NC ?/FS -/DASH <¿Eso/NC> espero/VLfin !/FS -/DASH
 18. 123142: llegará/VLfin ?/FS -/DASH <¿Ese/NC> sinvergüenza/NC ?/FS Ése
 19. 135051: compañero/NC /FS -/DASH <¿Eso/NC> va/VLfin diciendo/VLger
 20. 144356: RT mujeres/NC /FS -/DASH <¿Eso/NC> era/VEfin lo/ART que/CQUE
 21. 168752 CC el/ART Cuesco/NP «/SYM <¿Ese/NC> ?/FS Imposible/ADJ /FS
 22. 175034: REP otra/QU forma/NC !/FS <¿Ese/NC> hombre/NC es/VSfin un/AR
 23. 175072: /VEfin él/PPX -/DASH /FS <¿Ese/NC> hombre/NC es/VSfin un/AR
 24. 190963: epetidos/VLadj /LP «/SYM <¿Ese/VLfin> hombre/NC de/PREP la/ART
 25. 206078: /PREP puta/NC ?/FS -/DASH <¿Eso/NC> no/NEG te/PPX lo/PPC con
 26. 206136: a/ART mano/NC -/DASH /FS <¿Eso/NC> sí/ADV que/CQUE no/NEG !
 27. 302151: SH dijo/VLfin /FS -/DASH <¿Eso/NC> era/VEfin todo/QU ?/FS L
 28. 313302: /PREP ónix/NC ?/FS -/DASH <¿Ese/NC> grandote/ADJ cuajado/VLa
 29. 324109: Bonnal/NP ?/FS .../DOTS <¿Ese/VLfin> tipo/NC elegante/ADJ de/
 30. 328428: /ART trato/NC /FS -/DASH <¿Eso/NC> te/PPX hizo/VLfin creer/
 31. 337598: FS ¿Qué/NC horror/NC !/FS <¿Ese/VLfin> castellano/NC periférico
 32. 340024: P Riotinto/NC /FS -/DASH <¿Esa/NC> excéntrica/ADJ es/VSfin
 33. 343214: García/NC Márquez/NP !/FS <¿Ese/VLfin> fabricante/NC de/PREP be
 34. 382806: REP Álvaro/NC /FS -/DASH <¿Esa/NC> chica/NC aún/ADV está/VE
 35. 402694: /VLfin serlo/VLlinf ?/FS <¿Eso/NC> es/VSfin todo/QU ?/FS ¿A
 36. 405971: S ¿Ah/NC /CM no/NEG !/FS <¿Eso/NC> sería/VSfin demasiado/AD
 37. 423806: depresión/NC /FS -/DASH <¿Eso/NC> ha/VHfin sido/VSadj todo
 38. 424368: in qué/INT veneno/NC ?/FS <¿Eso/NC> importa/VLfin ?/FS De/PR
 39. 425244: /PREP puta/NC /FS -/DASH <¿Eso/NC> es/VSfin todo/QU ?/FS -/
 40. 425975: /NC sospechaba/VLfin ?/FS <¿Eso/NC> es/VSfin todo/QU ?/FS -/

Aun así, quedan siete casos de *ese* no acentuados ni precedidos por signo de exclamación o interrogación que se han etiquetado incorrectamente como sustantivos comunes. Los localizamos mediante la utilización del comando [word="ese" &

pos !="DM"], que arroja todos los casos en los que la palabra indicada ha sido etiquetada con cualquier otra etiqueta que no sea la de demostrativo.

```
# Query: COMP_ESOR_BAR; [lem="ese" & pos !="DM"];
```

```
#-----#  
1. 105192: ./FS ¿Cómo/NC está/VEfin <ese/NC> borracho/ADJ de/PREP tu/  
2. 110124: VLfin quién/INT era/VEfin <ese/NC> Enrique/NP ./FS ./DASH N  
3. 120051: S A/PREP lo/ART mejor/ADJ <ese/NC> Maceió/NP cae/VLfin cerc  
4. 314680: /DASH ¿Quién/NC era/VEfin <ese/NC> Peter/NP ?/FS Menuda/ADJ  
5. 343842: reído/VLadj el/ART tío/NC <ese/NC> ?/FS Acabo/VLfin de/CSUB  
6. 353075: eño/NC ./CM el/ART tío/NC <ese/NC> ./CM el/ART Conesal/NC .  
7. 356187: con/PREP el/ART poeta/NC <ese/NC> licuado/VLadj ./FS Me/PP
```

Dado que dos estructuras abarcan casi la totalidad de estas últimas siete concordancias, podríamos pensar que son estas las que provocan la confusión del programa. La primera es la posposición del demostrativo al sustantivo. Sin embargo, este tipo de casos se encuentra también entre los que TreeTagger ha etiquetado correctamente como demostrativo:

<Vázquez>: ¿Es cierto que le han dado un corte al forrao <ese> , al tragón de Conesal ?

Lo mismo sucede con los tres casos de etiquetado incorrecto del demostrativo delante de nombre propio¹⁵⁵, que se etiquetan correctamente en otras ocasiones y también cuando incluyen otra forma del demostrativo, lo que nos indica una incoherencia en el etiquetado que no resulta fácil de explicar:

<Vázquez>: los discípulos que le salieron al pobre , incluida <esa> Agnes Heller que es una fugitiva del terror rojo

En el caso de *aquel* (v. Tabla 3, tercera fila), la diferencia de 107 apariciones no se explica únicamente por la falta de formas acentuadas, pues estas solo ascienden a 19: los cuatro *aqué* que han sido bien etiquetados, el único precedido por interrogación —que se ha etiquetado como nombre común— y las 14 ocurrencias de *aquélla* que se han etiquetado principalmente como verbos. Las 88 que faltan en el cómputo corresponden a la única forma no acentuada de *aquel* precedida por exclamación —etiquetada también como nombre común— y a todas las apariciones de *aquello* del corpus, etiquetadas alternativamente como verbo o como nombre común, como muestran los siguientes ejemplos:

¹⁵⁵ Leonetti (1999: 804) atribuye a esta estructura funciones estilísticas y la denomina «deixis empática o emocional». A través de ella, el hablante introduce matices evaluativos de accesibilidad o familiaridad del referente.


```
# Query: COMP_ESOR_BAR; [word="aquello" & pos !="DM"];
```

```
#-----
1. 4567: ba/VLfin que/CQUE todo/QU <aquello/VLfin> me/PPX divertía/VLfin ./
2. 7073: FS Según/PREP el/PPX ./CM <aquello/VLfin> había/VHfin acabado/VLad
3. 8083: T libro/NC ./FS Pero/CCAD <aquello/VLfin> ya/ADV no/NEG era/VEfin
4. 8340: ./FS Definitivamente/ADV <aquello/VLfin> no/NEG era/VEfin el/ART
5. 10534: uenta/NC de/PREP que/CQUE <aquello/VLfin> no/NEG podía/VLfin conti
6. 15604: contándose/VLadj todo/QU <aquello/NC> ?/FS La/ART tragedia/NC
7. 16393: ó/VLfin como/ADV si/CSUBX <aquello/VLfin> le/PPC pareciera/VLfin u
8. 19486: L subinspector/NC todo/QU <aquello/VLfin> ./CM estudiar/VLinf junt
9. 24940: en/PREP absoluto/ADJ ./CM <aquello/VLfin> era/VEfin su/PPO rutina/
10. 25947: ¿Quién/ADV firmaba/VLfin <aquello/VLfin> ?/FS Daba/VLfin igual/AD
11. 26035: o/ADJ al/CSUBI leer/VLinf <aquello/NC> ./CM porque/CSUBX lo/PPC
```

Tampoco es así en los diecisiete casos omitidos de *este*, ya que en esta ocasión el programa sí ha recogido todas las formas acentuadas, excepto los tres casos precedidos por signo de exclamación. Además de no recoger el resto de formas precedidas por exclamaciones o interrogaciones (9), mediante el comando [lem="este" & pos !="DM"], podemos ver que el programa ha cometido varios errores al etiquetar demostrativos como sustantivos (etiqueta «NC») pero, a su vez, ha detectado correctamente los tres casos en los que *este* se utiliza efectivamente como sustantivo para indicar el punto cardinal denominado de esta manera (en negrita en las concordancias siguientes). Estos casos tienen que ser detectados manualmente cuando las búsquedas se realizan mediante expresiones regulares.

```
# Query: COMP_ESOR_BAR; [lem="este" & pos !="DM"];
```

```
#-----
1. 19490: estudiar/VLinf juntos/ADJ <este/NC> nuevo/ADJ material/ADJ .
2. 111886: EP la/ART calle/NC a/PREP <este/NC> imbécil/ADJ -/DASH orden
3. 156345: ón/NC ./CM el/ART crío/NC <este/NC> !/FS »/SYM Mamá/NC siemp
4. 194375: CQUE venía/VLfin del/PDEL <Este/NC> y/CC que/CQUE limpiaba/V
5. 218729: n/ART trato/NC como/CSUBX <este/NC> que/CQUE hacemos/VLfin c
6. 343961: ART distinción/NC de/PREP <este/NC> Single/PE Malt/NP ./CM a
7. 378930: PPX interesa/VLfin el/ART <Este/NC> ./CM que/CQUE ofrece/VLf
8. 378971: mente/ADV por/PREP el/ART <Este/NC> y/CC el/ART Oeste/NC ./F
```

Sucede lo mismo en el subcorpus TRAD. La búsqueda sobre el etiquetado mediante el comando [pos="DM"] arroja 3707 concordancias: 1241 formas de *este* (65 *ésta*, 1 *ésta*s, 102 *éste* y 12 *éstos*), 1860 de *ese* (ninguna forma acentuada) y 606 de *aquel* (12 *aqué*l). La búsqueda a través de expresiones regulares, no obstante, nos muestra un total de 3940 formas demostrativas y, por tanto, una diferencia de 233 ocurrencias entre un cómputo y otro; 242 si eliminamos las nueve formas de *este* sustantivo mal etiquetadas como demostrativos del primer cómputo (la diferencia es de 22 casos en *este*, 115 en *ese* y 105 en *aquel*).

"(¿ i)?(e é)st(e o a)(s)?" %c;	1254	(66 <i>ésta</i> , 1 <i>éstar</i> , 102 <i>éste</i> y 12 <i>éstos</i>)
"(¿ i)?(e é)s(o a)(s)?" %c "(e é)se" %c;	1975	(35 <i>ésta</i> , 6 <i>éstar</i> , 27 <i>éste</i> y 6 <i>éstos</i>)
"(¿ i)?aqu(e é)l(l)?(o a)?(s)?" %c;	711	(12 <i>aqué</i> l y 13 <i>aqué</i> lla)

Tabla 65 - Resultados de la búsqueda de demostrativos en TRAD mediante expresiones regulares

De nuevo, el programa ha logrado etiquetar correctamente todas las formas acentuadas de *este* (excepto la única precedida por un signo de exclamación) y varios casos en los que la misma palabra hace referencia al punto cardinal, etiquetándolo como sustantivo:

Query: PAR_TRAD_BARBARA; [word="Este" & pos != "DM"];

- #-----
1. 224561: C ./FS Noonan/VLfin ./COLON un/ART viaje/NC al/PAL <Este/NC> ./DASH furtivo-para/PREP solicitar/VLinf consejo/
 2. 225582: los/ART papeles/NC a/PREP un/ART juez/NC del/PDEL <Este/NC> ./FS ./DASH ¿Por/NC alguna/QU razón/NC en/PREP es
 3. 226327: h/NP Ave/NC ./FS 4024/CODE ./CM Los/ART Ángeles/NC <Este/NC> ./FS »/SYM Recortes/NC de/PREP prensa/NC sueltos/
 4. 242518: e/CQUE unos/ART mañosos/NC de/PREP la/ART Costa/NC <Este/NC> ./FS Yo/PPX diría/VLfin que/CQUE ha/VHfin
 5. 426018: dirección/NC en/PREP la/ART calle/NC Setenta/CARD <Este/NC> de/PREP Manhattan/NC ./FS ./DASH Merci/NP ./CM Ma

Esta vez, no obstante, debido a la utilización más frecuente en inglés de los puntos cardinales para indicar direcciones, ha tenido más oportunidad de equivocarse y ha etiquetado dos demostrativos como sustantivo (en negrita en las concordancias siguientes), además de incluir en el cómputo de demostrativos nueve casos que no lo son:

Query: PAR_TRAD_BARBARA; [word="este" & pos != "DM"];

- #-----
1. 1528: hallaba/VLfin conduciendo/VLger hacia/PREP el/ART <este/NC> con/PREP el/ART seguro/ADJ de/PREP las/ART puerta
 2. 15494: dj hasta/PREP entonces/ADV ./DASH El/ART sótano/NC <este/NC> ./DASH anunció/VLfin Roberts/NP mientras/CSUBX in
 3. 17860: cuerpo/NC ./FS Los/ART Heat/NP vivían/VLfin al/PAL <este/NC> de/PREP la/ART avenida/NC Chamberlayne/NP ./CM de
 4. 18017: X hubiéramos/VHfin seguido/VLadj hacia/PREP el/ART <este/NC> dos/CARD o/CC tres/CARD kilómetros/NC más/ADV ./CM
 5. 28233: Ladj se/SE curvaba/VLfin sobre/PREP la/ART zona/NC <este/NC> del/PDEL centro/NC ./CM donde/ADV las/ART vías/NC
 6. 93211: PDEL Ecuador/NP y/CC tantos/ADJ grados/NMEA al/PAL <este/NC> del/PDEL meridiano/ADJ de/PREP Greenwich/NP ./FS
 7. **96474: P que/CQUE se/SE haya/VHfin aclarado/VLadj todo/QU <este/NC> asunto/NC ./DASH Descolgó/NC el/ART**
 8. 149074: /VLadj ./FS Recorrí/VLfin en/PREP zigzag/NC el/ART <este/NC> y/CC el/ART sur/PE ./FS Ningún/QU Plymouth/NP gri
 9. 159587: stern/NP y/CC Adams/NP con/PREP acento/NC del/PDEL <este/NC> ./FS Su/PPO historia/NC ./COLON Lucille/NP trabaj
 10. 171132: C había/VHfin tomado/VLadj ./FS Probé/VLfin al/PAL <este/NC> ./COLON el/ART picadero/NC de/PREP Hughes/NC en/P
 11. 202428: l/CSUBI sur/PE por/PREP La/ART Brea/NC ./CM al/PAL <este/NC> por/PREP Slauson/NP ./FS Aque/DM chulo/NC negro/
 12. 202442: j de/PREP color/NC púrpura/ADJ ./FS Más/ADV al/PAL <este/NC> ./CM y/CC al/CSUBI sur/PE por/PREP Central/NC Ave
 13. 202471: Más/ADV al/CSUBI sur/PE ./FS Watts/NP ./FS Al/PAL <este/NC> ./FS Luces/NC de/PREP freno/NC ./DASH Avalon/NP y
 14. »/SYM Al/PAL <este/NC> hacia/PREP el/ART Olympic/NP ./SEMICOLON nubes/NC
 15. 228370: alle/NC 3/CARD ./CM Boyle/NP Height/NP ./FS Al/PAL <este/NC> ./CM hacia/PREP Wabash/NP ./COLON Cerrajería/NC B
 16. 232821: ke/NP ./FS Subí/VLfin por/PREP la/ART carretera/NC <este/NC> hasta/PREP el/ART observatorio/NC ./FS Un/ART cla
 17. **234409: ipos/NC filman/VLfin películas/NC porno/NC en/PREP <este/NC> local/ADJ y/CC ha/VHfin pensado/VLadj**
 18. 241205: Tommy/NP ./CM conduciendo/VLger hacia/PREP el/ART <este/NC> en/PREP sendos/NC coches/NC ./FS
 19. 241279: /CM en/PREP el/ART coche/NC ./CM hacia/PREP el/ART <este/NC> ./FS 12.11/CODE ./COLON Madge/NP y/CC
 20. 341346: ./CM cuyos/REL afluentes/NC bajaban/VLfin del/PDEL <este/NC> y/CC el/ART sur/PE ./FS Los/ART lagos/NC y/CC las
 21. 362390: te/ADV dicho/ADJ destacaba/VLfin hacia/PREP el/ART <este/NC> con/PREP su/PPO despliegue/NC de/PREP iluminados/
 22. 367285: ano/NC abierto/ADJ ./FS Se/SE dirigió/VLfin al/PAL <este/NC> por/PREP Portland/NP Avenue/NC ./CM dobló/VLfin p
 23. 462943: Lfin de/PREP la/ART región/NC situada/VLadj al/PAL <este/NC> de/PREP Fontainebleau/NP ./DASH ./FS Más/ADV gran
 24. 503685: NC de/PREP ventanas/NC que/CQUE daban/VLfin al/PAL <este/NC> ./CM un/ART calendario/NC en/PREP la/ART
 25. 578010: aminó/VLfin dos/CARD manzanas/NC hacia/PREP el/ART <este/NC> ./CM torció/VLadj al/CSUBI llegar/VLinf a/PREP la
-
1. 162320: n ser libres » ? Veamos , la geografía de la Costa <Este> no es nuestro fuerte y es evidente que ya está
 2. 227752: y en deuda con él . Del banco , directamente a L.Á <Este> , algo mareado por la leve inhalación de humos .
 3. 234152: la capota . Pete lo siguió , no demasiado cerca . <Este> , norte : el aire fresco me despejó . Seguimiento
 4. 242539: liquidaba sus trapaperras y algún tipo de la Costa <Este> intenta quedarse con el negocio . Puras estupidec
 5. 15528: n la galería de la muerte . A lo largo de la pared <este> se abrían cinco celdas , con una cama de hierro ,
 6. 57056: e West Cary con Windsor Way y aceleró en dirección <este> . Las luces del techo destellaban y las luces del

7. 119128: mientras conducía por la calle Broad en dirección <este> , de camino al tribunal John Marshall . - Ahora v
 8. 154691: ó raro . Salí tras el coche . Gower sur , Franklin <este> . Poco tránsito y nadie siguiéndome a mí . Al nor
 9. 371276: habitación semiprivada de un hospital llamado Baptista <Este> , con el peor dolor de cabeza que recuerdo

Por supuesto, ninguno de los 20 casos de demostrativo no acentuado precedido por exclamación o interrogación se encuentra etiquetado correctamente.

En el caso de *ese*, nuevamente, el hecho de que se hayan detectado 115 casos más mediante la búsqueda con expresiones regulares (v. tabla 4, segunda fila) se debe, además de al etiquetado incorrecto de los 30 demostrativos que forman parte de interrogaciones y exclamaciones, sobre todo al de todas las formas acentuadas. TreeTagger ha etiquetado las 73 formas acentuadas como adjetivos, nombres propios, nombres comunes y verbos. Los 12 casos restantes, como pasaba en ES, son formas de *ese* no acentuadas que se han etiquetado incorrectamente como sustantivos comunes:

Query: PAR_TRAD_BARBARA; [lem="ese" & pos != "DM"];

#

1. 93585: /NC cuando/CSUBX llamó/VLfin Marino/NC ./FS -/DASH <Ese/NC> condenado/VLadj alcalde/NC no/NEG puede/VMfin ten
2. 321743: /NP Lee/VLfin -/DASH dijo/VLfin Ray/NP ./FS -/DASH <Ese/NC> ./FS Y/CC un/ART hombre/NC apellidado/NC McDermid
3. 68086: Y/NP había/VHfin algo/QU significativo/ADJ en/PREP <ese/NC> sobre/PREP ?/FS -/DASH Sólo/ADV contenía/VLfin va
4. 169674: pezó/VLfin a/PREP divagar/VLinf acerca/ADV de/PREP <ese/NC> «/SYM despampanante/ADJ »/SYM (/LP cito/VLfin tus
5. 178310: REP la/ART película/NC demasiada/QU sangre/NC y/CC <ese/NC> incesto/ADJ ./CM así/ADV que/CQUE Mickey/NP teme/
6. 189026: adj ?/FS Hemos/VHfin sacado/VLadj cosas/NC de/PREP <ese/NC> Champ/PE en/PREP cuatro/CARD números/NC ,
7. 200174: thside../NP ./FS MATARLE/VLinf ./FS »/SYM .../DOTS <ese/NC> desgraciado/ADJ de/PREP policía/NC moqueante/NC v
8. 231766: especie/NC de/PREP jugador/NC arruinado/VLadj y/CC <ese/NC> jodido/VLadj ricacho/NC de/PREP Exley/NC le/PPC c
9. 246393: s/ADV ./FS Bueno/ADJ ./CM también/ADV estaba/VEfin <ese/NC> otro/QU chico/ADJ blanco/NC que/CQUE era/VEfin su
10. 247331: luego/ADV ./CM ¿qué/NC ?/FS -/DASH Luego/ADV ./CM <ese/NC> condenado/VLadj Stemmons/NP se/SE entrometió/
11. 472361: ./FS -/DASH ¿Dónde/NC está/VEfin el/ART gancho/NC <ese/NC> ?/FS -/DASH gritó/VLfin Janice/NC en/PREP tono/NC
12. 492955: te/PPX he/VHfin hablado/VLadj de/PREP él/PPX ./CM <ese/NC> al/PAL que/CQUE pillan/VLfin cada/QU vez/NC ./FS

Finalmente, la suma de 13 *aquellas* y 2 *aquel* incluidos en frases interrogativas y exclamativas no alcanza a explicar las 105 formas de más de *aquel* que encontramos mediante expresiones regulares (v. Tabla 4, tercera fila). Las 90 restantes son, como ocurría en ES, todas las apariciones de *aquello* del subcorpus TRAD, etiquetadas como «VLfin» (verbo léxico conjugado) o «NC» (nombre común).

En definitiva, podemos afirmar que el etiquetador reconoce todas las formas acentuadas del demostrativo *este* pero ninguna de *ese* y solo parte de las de *aquel*. Por otro lado, no es del todo eficaz (aunque es sorprendente que los errores no hayan sido muchos) a la hora de distinguir entre dos palabras homónimas como *este* o *ese* en sus funciones de demostrativo y sustantivo. Además, parece que, por algún motivo, *aquello* no ha sido tenido en cuenta como tal a la hora de programar el etiquetado, mientras que *esto* y *eso* sí se etiquetan correctamente. Finalmente, si hay un signo de exclamación o interrogación delante del demostrativo, el programa no lo reconoce en ninguna ocasión

como tal. Hay que mencionar que, en la búsqueda mediante expresiones regulares, tampoco los encontraremos si no añadimos estos caracteres a la expresión, puesto que en Corpus Workbench pasan a formar parte de la palabra que los sigue, con lo que el lema se modifica.

En este trabajo solo hemos analizado una categoría específica, la de los demostrativos, que conforma una pequeñísima parte de las palabras de un corpus entero. En el capítulo de las perífrasis, asimismo, vimos que los verbos no estaban lematizados de manera totalmente correcta prácticamente en ninguna ocasión (no hemos querido entrar a analizar si la categorización en los diferentes tiempos verbales estaba o no bien hecha). Por lo tanto, a pesar de la potencial utilidad del etiquetador para ahorrar tiempo con respecto a las búsquedas menos automatizadas, apreciamos una incoherencia en las pautas de etiquetado del programa que confirma nuestra impresión de que su utilización sin una revisión manual previa bastante exhaustiva es arriesgada.

5.2.3. El criterio del acento diacrítico para distinguir entre determinantes y pronombres

Es importante mencionar que hasta 1952 era obligatorio acentuar los pronombres demostrativos para distinguirlos de los determinantes demostrativos. Ese año la acentuación diacrítica pasó a ser facultativa y, en 1999, la RAE dictaminó que los pronombres ya no debían acentuarse, salvo en casos de ambigüedad. En 2010, la RAE establece en su nueva *Ortografía de la lengua española*¹⁵⁶ que no deben acentuarse nunca, puesto que los casos ambiguos se resuelven sin problemas con ayuda del contexto. Lo más probable es que en nuestro corpus encontremos esta acentuación diacrítica en mayor o menor medida, puesto que los textos pertenecen al período en el que quedaba a criterio de los autores el utilizarla o no, y es difícil deshacerse de este tipo de hábitos en la escritura. Seco (1989: 71) comenta: «Por tradición, sin embargo, sigue siendo muy general en España la costumbre de acentuar los pronombres demostrativos».

¹⁵⁶ Las principales novedades de la misma, entre las que se incluye el apartado «Eliminación de la tilde diacrítica en el adverbio *solo* y los pronombres demostrativos incluso en casos de posible ambigüedad» pueden consultarse en el siguiente enlace:

http://www.rae.es/sites/default/files/Principales_novedades_de_la_Ortografia_de_la_lengua_espanola.pdf [última fecha de consulta: 12 de abril de 2014]

No obstante, es evidente que habrá autores que hayan elegido no utilizar esta marca de diferenciación y, por otra parte, debemos tener en cuenta los errores ortográficos —o, en este caso, incoherencias ortográficas—, que puedan haberse pasado por alto durante la corrección y edición de las novelas. Para comprobarlo, lo primero que hacemos es dividir la búsqueda en dos grupos por cada demostrativo. La primera expresión recoge las formas no acentuadas y la segunda, las formas acentuadas y la neutra; es decir, las que, en principio, se considerarían pronominales por su grafía.

- 1) "(¿|í)?esos" %c|(¿|í)?esa(s)?" %c|(¿|í)?ese" %c;
- 2) "(¿|í)?ésos" %c|(¿|í)?ésa(s)?" %c|(¿|í)?ése" %c|(¿|í)?eso" %c;

Tras separar las formas acentuadas de las no acentuadas, el paso siguiente es estudiar las concordancias obtenidas para comprobar si esta división coincide con la clasificación ideal «determinante/pronombre» (teniendo en cuenta, por supuesto, que las formas neutras nunca se han acentuado).

Como observamos tras un cuidadoso cribado de las concordancias, esto no sucede así. Todos los autores utilizan algún tipo de diferenciación diacrítica, aunque pocos lo hacen de manera coherente. De todas formas, hay que tener en cuenta que, en principio, la acentuación diacrítica se recomendó para resolver casos de anfibología, no para distinguir sistemáticamente entre pronombres y determinantes, aunque esta haya sido la costumbre adoptada por la mayoría.

Manuel Vázquez Montalbán es el más coherente en este sentido. Acentúa los pronombres y no acentúa los determinantes. En sus concordancias, así como en las de los demás autores, existen ejemplos en los que puede haber dudas acerca de si el demostrativo en cuestión debe considerarse pronombre o determinante. Por ejemplo, en el *Diccionario Panhispánico de Dudas* (RAE 2005), dentro del apartado dedicado a la tilde diacrítica, encontramos la siguiente descripción de los demostrativos, que nos indica que los demostrativos pospuestos¹⁵⁷ se consideran determinantes:

Los demostrativos este, ese y aquel, con sus femeninos y plurales, pueden ser pronombres (cuando ejercen funciones propias del

¹⁵⁷ El demostrativo pospuesto, según Leonetti (1999: 811) «difiere de la secuencia correspondiente con demostrativo antepuesto en algunos matices estilísticos, desde el efecto poético de situar a un objeto en la lejanía, hasta el contenido despectivo o irónico, que no es más que una extensión metafórica, de distanciamiento por parte del hablante, del significado básico de mayor grado de lejanía o menor grado de accesibilidad».

sustantivo): Eligió este; Ese ganará; Quiero dos de aquellas; o
adjetivos (cuando modifican al sustantivo): Esas actitudes nos
preocupan; El jarrón este siempre está estorbando.

Sin embargo, en el punto 17.5m de la *Nueva gramática de la lengua española* (RAE 2009: 1306), se afirma que «no existe acuerdo entre los gramáticos sobre la naturaleza categorial del demostrativo pospuesto», aunque, según los autores, parece mejor encaminado el análisis según el cual deben considerarse pronombres. Para nuestro análisis, seguiremos al *Diccionario Panhispánico* y a la opinión más generalizada en las fuentes consultadas, que es la de considerarlos adjetivos o determinantes (p. ej. Álvarez de Miranda 2014; Macías Villalobos 2006: 207).

En cualquier caso, la coherencia hallada en los ejemplos del corpus de Montalbán deja claro que su utilización o no del acento es consciente y que, por tanto, él los considera determinantes.

<Vázquez>: - A mí no me reconvierte ni Dios . ¿Qué se habrá creído el tío <ese> ?

<Vázquez>: Otra poca soltada . ¿Cómo se puede fallar un premio con el aspecto <aquel> tan horrible que tenía ?

Lo mismo sucede en los casos en los que el demostrativo precede una oración de relativo. El carácter gramatical de dichos demostrativos es discutible pero, en principio, los consideraremos pronombres (Macías Villalobos 2006: 201; Álvarez de Miranda 2014; RAE 2009: 1279, 1295). Vázquez de Montalbán no los acentúa en ningún caso, como recomienda Manuel Seco en su *Gramática esencial del español* (1989: 71).

<Vázquez>: Los pequeños hijos de puta no cuentan . Los que cuentan son <esos> que están en condiciones de hundir a los demás , sean quienes fueren .

<Vázquez>: Es cierto que la revolución conservadora es una involución redistributiva , pero <aquel> que no se sume , será devorado por ella .

Sin embargo, también encontramos un único error de acentuación, en el que se ha puesto tilde a una forma determinante del demostrativo:

<Vázquez>: En <ésas> circunstancias la gastronomía es lo de menos y sería una grosería que la cena fuera más buena que la obra

Alicia Giménez-Bartlett, por su parte, también es bastante coherente en la utilización del acento diacrítico, excepto en un caso, en el que omite el acento en el pronombre *ese*:

<Giménez>: Esperaba que <ese> no fuera el caso del subinspector , que su solidaridad conmigo se tratara tan sólo de una anécdota

Dicho acento, por supuesto, no es necesario, pero el hecho de que ella parezca ponerlo de manera sistemática en el resto de las ocasiones, nos lleva a pensar que se trata de una confusión:

<Giménez>: Bien , entonces había que comenzar en el mismo punto en el que acabamos : la púa de plata .
<Ése> era el hilo que debía ser tirado .

<Giménez>: - ¡Vaya pájaro ! - exclamó Garzón - . Un moderno y eficaz , a <esos> me los conozco como si los hubiera parido .

Por algún motivo, sin embargo, no acentúa el demostrativo *aquel*, pero al no hacerlo en ninguna de sus formas, no podemos tachar esta acción, en este caso, de incoherencia.

<Giménez>: No obtendríamos la más mínima cooperación por parte de las víctimas . <Aquel> sería siempre un asunto de familia y la violada era su punto más débil ,

<Giménez>: ¿A qué se debía su metamorfosis ? ¿Habría sido <aquella> la primera vez que veía replicar una orden ? Probablemente sí .

<Giménez>: A lo mejor cualquier día de <aquellos> uno de sus traficantes de tabaco le pegaba un tiro y lo mataba .

A diferencia de Vázquez Montalbán, no es coherente en la pauta de acentuación de los demostrativos seguidos de oración relativa:

<Giménez>: Esto de revivir los geranios me lo contó un tipo que viene por mi bar , uno de <esos> que realiza cultivos sin fertilizantes artificiales .

<Giménez>: Hábleme de la culpa , Garzón . - ¿De la culpa ? - Sí , de <ésa> que siente cuando se extralimita . Se le borró la sonrisa , bebió coñac .

Encontramos también varios ejemplos que son sencillamente errores de acentuación, puesto que se trata claramente de casos en los que el demostrativo acompaña a un sustantivo y, por tanto, hablamos de determinantes:

<Giménez>: La vimos alejarse , evanescente como una pincelada en un cuadro . - ¿Y cómo <ésa> chica ha podido matar a nadie ?

<Giménez>: Lo intenté faltando a una de mis reglas de oro sobre trabajo y vida privada : - ¿Sabe quién es <ése> caballero que acabo de presentarle ?

<Giménez>: En el mismo momento en que a ti se te deshace una onza de chocolate en la boca , en <ése> mismo momento , un niño pobre está muriéndose de hambre

Juan Madrid, Jorge Martínez Reverte y Andreu Martín también acentúan la mayor parte de los casos de pronombre, como vemos en los siguientes ejemplos:

<Madrid>: . ¿Qué te debo , Inchausti ? - Ni lo piense , señor Ventura . Usted no paga aquí . <Ésta> es su casa .

<Martín>: Es un pobre hombre rodando por la pendiente que lleva indefectiblemente al suicidio . <Éste> fue el presentimiento con que ayer Santiago recibió a Lucía .

<Martínez>: Me arrepentí , pero no fui capaz de oponerme . ¿Por qué no le había dicho que <ése> era mi territorio ? Pero el mal ya estaba hecho

<Martín>: Mejor los recuerdos . Puso su mano en la espalda de Rita y era consciente de que <aquella> era la primera vez que tocaba a una de aquellas mujeres

<Martínez>: la misma afabilidad con la que trataba a los semáforos - , aunque supongo que no será usted uno de <esos> .

<Martín>: ¿Con tu propia mujer ? La pregunta desconcierta al hombre cansado . Le vienen ganas de replicar qué tontería es <aquella> , pero calla .

Sin embargo, dejan sin acentuar varios casos que podrían considerarse paralelos a los que acabamos de transcribir (sirvan como ejemplo las expresiones en negrita de la primera frase de ambos listados):

<Madrid>: El japonés sonrió . - Osaka , yo llamar . - Y otro día vienes más temprano . <Esta> **no es la casa** de tócame Roque .

<Martín>: Lo considero tan justo como pagar un peaje de autopista . <Este> es vuestro barrio y vosotros decidís que hay que aforar para visitarlo .

<Martínez>: Anda , que cómo se ha puesto mi madre . Hasta que he podido convencerla de que <ese> no era el vídeo que yo había hecho...

<Martín>: falda corta , zapatos de tacón y jersey escotado , pero virgen , según demostraría la autopsia . <Esa> era la noticia del día , que apareció con letras de palmo en la primera página

<Martínez>: Si ven un taxi , lo cogen sólo para darnos un billete de cinco mil pesetas . Yo , a <esos> , les cogía de los huevos y les daba hasta hartarme .

<Martín>: Malos años <aquellos> , porque las familias venidas a menos rebosan frustraciones y rabia y disparan con facilidad las acusaciones que no se

Por otra parte, los tres son completamente coherentes a la hora de no acentuar los demostrativos que funcionan como antecedente de oraciones relativas. En el caso de los demostrativos pospuestos, además, todos los ejemplos hallados en el texto de Andreu Martín están sin acentuar y todos los pertenecientes a Juan Madrid están acentuados, por lo que, a pesar de que no coinciden en su clasificación de estas formas como pronominales o determinantes, la acentuación en ambos casos es coherente. En Martínez Reverte no encontramos casos de este último tipo.

<Madrid>: - Te he dicho que quiero charlar contigo . Y vamos de aquí , anda . La calle <ésta> me pone nervioso . - ¿Y adonde quieres ir ?

<Martín>: Pocas . Yo sólo he conocido a una . Sí , la chica <esa> , ¿cómo se llamaba ? Era una excepción . Luego hablaré de ella .

Soledad Núñez, por último, a pesar de acentuar todas las formas pronominales del demostrativo *aquel*, es completamente incoherente con el resto de formas, como podemos ver en los siguientes ejemplos¹⁵⁸:

¹⁵⁸ En todos puede observarse que el demostrativo no acentuado está a principio de frase, por lo que, para descartar que la falta de tilde responda a una norma de no acentuación de mayúsculas, buscamos casos de mayúsculas acentuadas en el corpus. A pesar de que son menos que en el resto de textos, hay casos de este tipo. Ejemplos:

<Núñez>: <Él> también estaba en la cal
<Núñez>: - <Ánimo> , querida - me dijo el h

<Núñez>: - ¿Te gusta el besugo ? - ¿Me vas a invitar a comer ? - <Esta> mañana lo he guisado . Tenía la estúpida esperanza de que Echeve apareciera .

<Núñez>: novela . Un tiempo de proyectos . De ambiciones . De alegría compartida . Mi embarazo acababa de confirmarse . <Esta> era una mañana como tantas otras .

<Núñez>: Tengo guardada una botella de Vega Sicilia . Era para una ocasión muy especial . Y <ésta> lo es . Ha sido todo un placer conocerte , Celia .

<Núñez>: La verdad , comisario , es que no sé cómo te puedo ayudar . - Déjalo de mi cuenta . <Este> asunto necesita un toque de sutileza , de ingenio .

<Núñez>: - Ayúdame , Celia . Te estoy protegiendo . ¿Cuándo vas a confiar en mí ? <Este> es un asunto muy grave . Beatriz ha denunciado su desaparición .

<Núñez>: - Antes me has acusado de farsante , pero la realidad es que <éste> es tu único refugio . Te quedarás aquí . - Abrázame , comisario .

<Núñez>: Pero deja que siga con el resto de la historia . <Esa> tarde se iba a hacer una entrega importante . Teníamos todo previsto para detenerles .

<Núñez>: Porque nunca se puede empezar una nueva vida . <Esa> es otra de las muchas mentiras que nos contamos .

<Núñez>: ¿por qué te dejó libre ? - Querida - dije suspirando - , <ésa> es la pregunta del millón de dólares . Le doy mil vueltas y no lo puedo entender .

<Núñez>: que ellas acudieran a mí tan dispuestas a complacerme como el genio de la lámpara a Aladino . Y con <ese> as en mi manga siempre posponía el momento .

<Núñez>: Para llegar a Barcelona había un transbordo en Zaragoza . <Ese> era mi tren . Conocía la ciudad . Ya era una ventaja .

<Núñez>: - Hola , Víctor - saludé . Levantó la cabeza y me miró de arriba abajo . Y <ése> fue su saludo . Igualito que un asno . Seguí insistiendo .

Tampoco muestra una pauta clara de acentuación de los demostrativos ante oración de relativo:

<Núñez>: Un miserable incapaz de albergar sentimiento alguno en su podrido corazón de hiena . Pero era de la raza de <esos> que creen que les está permitido todo , « pero a mi mujer , ni me la toquen » .

<Núñez>: Tuve la impresión de estar agarrando un cepillo de púas , de <esos> que se utilizan para limpiar moquetas .

Por tanto, en el subcorpus ES vemos una presencia de incoherencias bastante importante a la hora de utilizar la tilde diacrítica. Sorprendentemente, los autores cuya

<Núñez>: <Éramos> novios - dijo con firmeza
<Núñez>: <Échate> en el asiento y mantén

acentuación diacrítica parece más lógica son los que han cometido errores ortográficos al acentuar formas que son claramente no pronominales.

En el subcorpus TRAD las incoherencias y errores observados son menos. Sabaté Vargas —el traductor de Ellroy—, por ejemplo, coloca de manera sistemática tilde a los pronombres y no acentúa los determinantes. Lo mismo observamos en la traducción de Mustieles del texto de Cornwell, salvo por dos casos en los que no se acentúa el pronombre y uno en el que se comete el error de acentuar un determinante:

<Cornwell>: - <Esta> es la idea general . - Espero que a alguien se le ocurra echarle un vistazo a sus finanzas .

<Cornwell>: Sé que usted ve muchas clases de lesiones distintas , y se me ha ocurrido que quizá podría hacerse una idea de cómo fueron infligidas <estas> y por qué .

<Cornwell>: - Los comentarios de <ésta> clase no le hacen ganar admiradores . Es precisamente lo que estaba diciéndole .

Ninguno de estos dos traductores utiliza demostrativos pospuestos o precediendo relativas, salvo Sabaté Vargas en un caso (pospuesto), que acentúa.

En los otros tres textos no hallamos ningún error pero las contradicciones apreciadas son mayores. En el de Moya, que traduce a Grafton, se dejan sin tilde nueve apariciones del pronombre demostrativo *esta* y 5 de *este*, pero se acentúan otras 23 y 22, respectivamente, que no pueden considerarse posibles casos de anfibología. Lo mismo ocurre con *esa* y *ese*. Sin embargo, todos los demostrativos no acentuados se encuentran a principio de frase, y hallamos que otras palabras, como *éramos* o *él* tampoco se acentúan en esta posición, por lo que posiblemente nos encontremos ante una política editorial (o personal del traductor) de no acentuación de mayúsculas.

<Grafton>: - Lo intentaré , aunque <ésta> no es mi jurisdicción .

<Grafton>: <Esta> es una caja doméstica contra incendios . Si fuese antirrobo , encontraríamos un blindaje

<Grafton>: Las puse sobre la palma izquierda de Gilbert . <Este> las miró por encima y se las guardó en el bolsillo del chaquetón .

<Grafton>: - Se la dio a Bucky y <éste> me la devolvió . - Sí , pero ¿por qué la guardaba en una caja fuerte contra incendios ?

<Grafton>: - Nunca se me ocurriría . <Esa> no es . Déjame que la vea .

<Grafton>: - ¿Cómo sabes que no nos vamos a querer ? - Estoy segura de que sí - dije - , pero <ésa> no es la cuestión . - ¿Cuál es entonces ?

<Grafton>: - Johnny Lee - dijo Ray . - <Ese>. Y un hombre apellidado McDermid .

<Grafton>: - ¿Dónde está Farley ? - preguntó Laura con voz sombría . - Ah , <ése> . Lo dejé en California .

En este texto no se acentúan nunca los demostrativos que preceden relativas pero observamos nuevamente incoherencias en la acentuación de los pospuestos:

<Grafton>: Puede que su pasado en las Fuerzas Aéreas , la historia <esa> de la Unidad de Voluntarios , fuera sólo una explicación honorable para justificar su ausencia .

<Grafton>: - Ay , Señor , Señor , me habría gustado verlo . ¿Cuántos años tiene la vieja <ésa> ? - Vamos a ver .

La traductora de Highsmith, Núñez, es coherente en la no acentuación de formas demostrativas ante oración relativa o pospuestas al sustantivo, pero, como Moya, acentúa algunos pronombres y otros no:

<Highsmith>: Janice era imprevisible , era una bala perdida , pensó Tom , pero <aquel> era un término absurdo para alguien tan frágil como ella .

<Highsmith>: Había leído que las patatas horneadas en el microondas quedaban como si estuvieran hervidas y no como las hechas al horno . Y <aqué!> era un tema que Heloise , Madame Annette y Tom se tomaban muy en serio .

Lo mismo sucede con de Hériz Ramón, traductor de Leonard, pero, en su caso, además, observamos que a veces acentúa los demostrativos ante oración relativa y otras no:

<Leonard>: Como la del anuncio de cerveza en la tele . <Ésa> que trabaja en un garito en el desierto . Lo has visto .

<Leonard>: De modo que nuestra agente los ha seguido para ver adónde iban y para entrar en contacto con los otros , para alertarlos... - ¿Era <esa> que tenía algún problema con el micrófono ?

<Leonard>: Como le pasó al que me hacía los silenciadores . » <Ésta> que se ve ahora es una MAC-10 . Lleva silenciador .

<Leonard>: Max sabía quién era la chica joven y dónde vivía . Pero no sabía nada de <aquella> que ahora se metía en el coche con la bolsa y se largaba .

Resulta interesante destacar el hecho de que, en general, hemos detectado una uniformidad mayor en los criterios de acentuación, o menos equivocaciones e incoherencias en cuanto a la colocación de tildes diacríticas, en las traducciones (ES: 51; TRAD: 34), a pesar de ser este subcorpus el de mayor tamaño. Esto podría deberse al hecho de que, a fin de cuentas, los traductores son profesionales de la lengua que deben documentarse concienzudamente sobre todo tipo de aspectos lingüísticos y conceptuales a la hora de realizar su labor de traducción y deben rendir cuentas a correctores y editores por su trabajo. Muchas veces ejercen también de correctores, por lo que pueden ser más conscientes de estos detalles a la hora de utilizar el lenguaje.

De todo esto podemos concluir que, a pesar de no ser obligatorio desde los años cincuenta, lo habitual en nuestro corpus es que autores y traductores utilicen la tilde diacrítica para diferenciar pronombres demostrativos de determinantes demostrativos. Sin embargo, las contradicciones apreciadas no permiten fiarse de este criterio a la hora de realizar un cómputo del uso de estas formas, como observan de Kock y Gómez Molina (1992: 17) en su estudio sobre el pronombre demostrativo, en el que concluyen, además, que los signos de puntuación tampoco son excesivamente útiles para diferenciar funciones gramaticales.

Por otro lado, detectamos en esta primera fase de análisis de concordancias una mayor uniformidad en traducciones con respecto a las estructuras formadas por los demostrativos, sean pronombres o determinantes. Por ejemplo, el demostrativo pospuesto al sustantivo, que cuenta con 19 casos en nuestro subcorpus ES, solo aparece en 6 concordancias de TRAD (LL 11,33 $p < 0,001$). También en este subcorpus hay menos casos de demostrativos ante oraciones de relativo (17 frente a 26; LL 5,35 $p < 0,05$); y hay que recordar una vez más que TRAD cuenta con mayor número de palabras que ES. Este hecho podría relacionarse con el universal de traducción denominado normalización, que, como hemos explicado en la revisión teórica, postula que los textos traducidos tienden a exagerar las características típicas de la lengua meta y a variar poco la norma gramatical o estilística. Sin embargo, no podemos hacer más, en este momento, que hipotetizar al respecto, puesto que los casos hallados no son lo suficientemente numerosos para extraer conclusiones de este tipo, que requieren la realización de estudios más amplios y variados.

5.3. Resultados y discusión

En primer lugar, si realizamos el cómputo de concordancias teniendo en cuenta los casos descartados de la última criba de resultados y comparamos el número total de demostrativos que hemos obtenido en cada subcorpus, las cifras finales absolutas son 3670 (ES) y 3940 (TRAD). Al aplicar el test de log-likelihood para la comparación de frecuencias¹⁵⁹, obtenemos el siguiente resultado, que indica que existe una infrarrepresentación muy significativa de demostrativos en los textos traducidos: LL +87,34 $p < 0,0001$. A continuación, desglosaremos estas cifras por autores, formas y clase sintáctica, comparando en todo momento textos originales y traducciones.

5.3.1. Distribución de formas demostrativas por textos

Como en el estudio de las perífrasis, hemos multiplicado cada resultado por 100 000 y lo hemos dividido por el número de palabras del texto de donde proviene. De esta manera, obtenemos las siguientes frecuencias normalizadas, que nos permiten comparar sobre una base común el uso que hacen los autores y traductores de cada forma y comprobar si existen particularidades estilísticas en alguno de los textos que puedan afectar a los cálculos de frecuencias que realizaremos a continuación.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	960	1131,2	Ellroy	1061	966,7
Vázquez	761	866,7	Leonard	803	1032,3
Martín	596	1060,6	Cornwell	776	706,3
Núñez	550	1295,2	Highsmith	716	728,4
Madrid	405	977,6	Grafton	584	629
Martínez	398	732,5			
Total	3670	999,7		3940	806,4

Tabla 66 - Distribución de demostrativos por texto

En el subcorpus ES, vemos que al normalizar las frecuencias, Núñez asciende a la primera posición. Esto nos indica que hace un uso mucho más abundante de demostrativos del que le correspondería si consideráramos el caso ideal en el que a un número mayor de palabras en un texto correspondiera un número mayor de

¹⁵⁹ Recordemos que para aplicar este test no es necesario normalizar las frecuencias previamente.

demostrativos, dado que hablamos de la segunda novela más corta del subcorpus. Madrid también ocupa un puesto bastante más alto del que le correspondería (es el último en número de palabras), mientras que Vázquez Montalbán, el autor de la novela más larga, queda en quinto puesto.

En TRAD llama la atención sobre todo la traducción de Hériz de Ramón del texto de Leonard, que, a pesar de contar con el menor número de palabras del subcorpus, ocupa la primera posición con respecto al uso de demostrativos, superando a las de Ellroy y Cornwell, los dos textos más extensos. A pesar de contar estos últimos con casi el mismo número de palabras, la traducción de Ellroy de Sabaté Vargas despierta un poco, dado que la de Cornwell queda en cuarto puesto, detrás de esta y la de Highsmith. Al parecer, el recorte en número de palabras del texto de Ellroy del que hablamos en el capítulo de las perífrasis atañe, en efecto, principalmente a los verbos y no a otros elementos como los demostrativos. La traducción de Grafton ocupa la última posición, puesto que le corresponde por número de palabras si no contamos la de Leonard.

A continuación, obtenemos las frecuencias por texto de cada forma básica por separado:

➔ ESTE

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Vázquez	329	374,7	Cornwell	377	343,1
Giménez	197	232,1	Ellroy	297	270,6
Martín	175	311,4	Grafton	236	254,2
Núñez	164	386,2	Highsmith	175	178
Madrid	119	287,2	Leonard	169	217,2
Martínez	103	189,6			
Total	1087	296,1		1254	256,7

Tabla 67 - Distribución de *este* por texto

La amplia diferencia que muestra aquí Vázquez Montalbán con respecto a las frecuencias absolutas de los demás autores puede deberse a que la acción de su novela se produce en una noche en la que se celebra una entrega de premios y, en gran parte de las conversaciones que tienen lugar, los asistentes comentan sobre dicha celebración, señalando al tiempo y a la situación presentes con los elementos del discurso. Por ejemplo, de 71 casos de la secuencia «esta noche» presentes en el subcorpus ES, 50 corresponden al texto de Vázquez Montalbán.

De todas formas, al normalizar las frecuencias, vemos que Núñez presenta un uso relativo mayor. Martín se queda en el puesto que le corresponde mientras que Madrid escala hasta la cuarta posición y Giménez y Martínez presentan una frecuencia de uso menor a la que les correspondería por número de palabras.

En cuanto a TRAD, observamos que la traducción de Leonard pasa a ocupar la penúltima posición en frecuencia de utilización de *este*, a pesar de ser el texto con más cantidad de demostrativos según la tabla 6, mientras que el de Cornwell recupera el primer puesto que le corresponde por número de palabras, Grafton asciende un poco y Highsmith queda en último lugar.

➔ ESE

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	479	564,4	Ellroy	713	649,6
Martín	309	549,9	Leonard	484	622,2
Vázquez	283	322,3	Cornwell	335	304,9
Madrid	261	630	Highsmith	223	226,9
Núñez	224	527,5	Grafton	220	236,9
Martínez	217	399,4			
Total	1773	483		1975	404,2

Tabla 68 - Distribución de *ese* por texto

En este caso, quien nos sorprende es Madrid —último en número de palabras y cuarto en la tabla de todos los demostrativos en conjunto—, que es quien más uso relativo hace de *ese*. Núñez se queda atrás en este caso, volviendo a su quinto puesto original. Martínez, como siempre hasta el momento, queda de los últimos a pesar de ser el cuarto en número de palabras, pero esta vez no es superado por Vázquez Montalbán, quien es el autor que menos utiliza este demostrativo proporcionalmente.

En TRAD, el texto de Highsmith, que es el tercero en número de palabras y demostrativos, queda en último lugar, pero el de Leonard vuelve a ascender hasta alcanzar el segundo puesto de la tabla, lo que acerca los resultados de *ese* a los de los demostrativos en conjunto. Vemos más variación en el uso de *ese* con respecto al número de palabras de cada texto: algunos autores y traductores se decantan por un mayor uso de esta forma que otros.

➔ AQUUEL

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	284	334,6	Highsmith	318	323,5
Núñez	162	381,5	Leonard	150	193
Vázquez	149	169,7	Grafton	128	138
Martín	112	199,3	Cornwell	64	58,2
Martínez	78	143,5	Ellroy	51	46,5
Madrid	25	60,3			
Total	810	220,6		711	145,5

Tabla 69 - Distribución de *aquel* por texto

Vemos una diferencia muy importante entre el último y el primer puesto de ambas tablas. Núñez, cuyo texto es el anteúltimo en cuanto a número de palabras totales de ES, vuelve a ocupar el primer puesto, y las traducciones de Cornwell y Ellroy, que son las primeras en número de palabras en TRAD, ocupan los últimos puestos. Una posible causa podría ser el mayor grado de oralidad presente en determinados textos, que suele asociarse con un menor grado de literariedad y, por tanto, de utilización del demostrativo *aquel*, como habíamos mencionado en el apartado 2.1.2. Núñez, por ejemplo, cuenta con un porcentaje de diálogo menor que la mayor parte del resto de los autores (Martínez Vilinsky 2014: 242). Con la traducción del texto de Cornwell también se confirmaría esta hipótesis, puesto que es la que más se vale del diálogo y ocupa uno de los últimos puestos de la tabla 12. Sin embargo, las de Leonard y Highsmith son el segundo y el tercer texto que más utilizan los diálogos como estrategia narrativa y Highsmith, por ejemplo, que se mantiene en el último puesto en las tablas de los otros demostrativos, aquí ocupa el primer lugar. Los traductores, sobre todo, parecen hacer un uso más idiosincrático de *aquel* que de *este* e incluso de *ese*, pero en ambos subcorpus, las diferencias de frecuencias entre uno y otro texto apuntan a una mayor variabilidad de uso.

Sea como sea, algunas preferencias están claras: Núñez sobresale con respecto a los demás autores por su uso de *aquel* (aunque, en realidad, utiliza más cantidad de veces el demostrativo *ese*), la traductora de Highsmith presenta una frecuencia normalizada de casi 100 casos más de *aquel* que de *ese*, mientras que Madrid, Mustieles (Cornwell) y Sabaté Vargas (Ellroy) parecen decantarse claramente por la utilización de *ese* frente a *aquel*.

5.3.2. *Comparación de distribuciones de las formas determinantes y pronominales entre subcorpus*

Tras haber comprobado que el criterio de distinción entre determinantes y pronombres demostrativos basado en la utilización de tilde diacrítica no es eficaz, realizamos una segunda clasificación manual para obtener las cifras de utilización de ambas categorías tanto en el subcorpus de español original como en el de español traducido. En primer lugar, podemos señalar que los porcentajes en ambos subcorpus en cuanto a clasificación en determinantes y pronombres no difieren en demasía:

ES:	61 % determinantes	(2237)	39 % pronombres	(1433)
TRAD:	60 % determinantes	(2358)	40 % pronombres	(1582)

Está claro que hay más determinantes que pronombres. Este fenómeno es común a algunos de los estudios sobre demostrativos a los que hemos tenido acceso, como los de Macías Villalobos (2006: 369) o Labrador (2011: 80), quizás debido a lo que explica López Samaniego (2010) en su estudio sobre los mecanismos referenciales en la sentencia de los atentados del 11M:

Según la teoría de la accesibilidad de Ariel (1990), los pronombres demostrativos, como pronombres que son, presentan un grado de accesibilidad baja: imponen una serie de restricciones de uso mayor que la que requieren otras expresiones anafóricas más informativas. Por este motivo, es frecuente que en los textos escritos —y, aún más, en los especializados— cuando se emplea un demostrativo para realizar una referencia retrospectiva a un elemento previo del texto, el emisor sienta la necesidad de añadir junto a él un nombre que o bien repite el antecedente literalmente, o bien constituye un sinónimo, cuasi-sinónimo o hiperónimo de este.

Como vemos, los pronombres demostrativos se consideran un mecanismo de recuperación de la información menos explícito de lo que pueden serlo, por ejemplo, los determinantes. Dado que según nuestros datos no parece haber grandes diferencias entre los porcentajes de uno y otro subcorpus, no podemos afirmar que se produzca con

respecto a estos elementos un fenómeno de explicitación de la información en textos traducidos.

5.3.2.1. Este: cribado y clasificación manual

En la tabla 70 pueden verse agrupadas las cifras de utilización del determinante *este* y sus formas flexionadas, tanto en el subcorpus de español original como en el de español traducido, junto con las frecuencias normalizadas y los porcentajes correspondientes. Asimismo, la anteúltima columna nos indica la diferencia entre las frecuencias normalizadas de ambos subcorpus. Un número negativo representa una frecuencia menor de aparición de un determinado elemento en el subcorpus de traducciones; dicho de otro modo, nos indica que el elemento en cuestión estaría infrarrepresentado. En la última columna encontramos el resultado del test de log-likelihood para comprobar la significación estadística de estas diferencias. El color de la celda, como en el caso de las perífrasis, nos indica si el elemento está infrarrepresentado (verde) o sobrerrepresentado (rojo) en traducciones con respecto a los originales en la misma lengua, o si presenta una frecuencia mayor (amarillo) o menor (fucsia) en originales pero no existe una diferencia significativa entre subcorpus que avale su clasificación en ninguno de los dos primeros grupos.

Determinante	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Este	276	37,86	75,19	235	32,28	48,10	-27,09	25,42 p < 0,0001
Esta	313	42,94	85,26	366	50,27	74,91	-10,35	2,81 (n. s.)
Estos	70	9,60	19,07	72	9,89	14,74	-4,33	2,35 (n. s.)
Estas	70	9,60	19,07	55	7,55	11,26	-7,81	8,64 p < 0,01
Total Este	729	32,59	198,59	728	30,87	149,01	-49,58	29,95 p < 0,0001

Tabla 70 - Comparación de frecuencias del determinante *este* en ES y en TRAD y significación estadística de las diferencias observadas

El test de log-likelihood confirma la infrarrepresentación de este determinante demostrativo en traducciones: 729 casos en ES frente a 728 en TRAD (LL 29,95 p < 0,0001). Sin embargo, al analizar la tabla con más detalle, vemos que la infrarrepresentación está presente en solo dos de los cuatro determinantes (*este* y *estas*), mientras las diferencias no llegan a ser estadísticamente significativas en el caso de *esta* y *estos*. Las diferencias pueden deberse a multitud de factores relacionados con el idiolecto de los escritores o incluso con las características particulares de cada

novela —temática, personajes, acontecimientos— como vimos en el caso de Vázquez Montalbán y su abundante utilización de la expresión temporal «esta noche». No obstante, las formas que nos interesan sobre todo no son las flexionadas sino los lemas, pues son los que señalan cada una de las tres posiciones del sistema de demostrativos.

Cuando analizamos los casos en los que funciona como pronombre, reflejados en la tabla 71, vemos que, mientras que en su forma determinante encontramos una infrarrepresentación de la forma básica *este* y de la mitad de sus formas flexionadas, ninguna de las formas pronominales de este demostrativo está infrarrepresentada, —aunque tampoco sobrerrepresentada de manera estadísticamente significativa—.

Pronombre	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Este	76	21,23	20,70	110	20,91	22,52	1,81	0,32 (n. s.)
Esta	52	14,53	14,17	78	14,83	15,97	1,80	0,45 (n. s.)
Estos	8	2,23	2,18	15	2,85	3,07	0,89	0,63 (n. s.)
Estas	6	1,68	1,63	2	0,38	0,41	-1,23	3,40 (n. s.) ¹⁶⁰
Esto	216	60,34	58,84	321	61,03	65,70	6,86	1,58 (n. s.)
Total Este	358	25,03	97,52	526	33,25	107,66	10,14	2,10 (n. s.)

Tabla 71 - Comparación de frecuencias del pronombre *este* en ES y en TRAD y significación estadística de las diferencias observadas

Finalmente, si unimos ambos cómputos (v. tabla 72), observamos que, a pesar de que los pronombres cuentan con una forma más que los determinantes, el mayor número de casos en los que *este* funciona como determinante es definitorio, pues existe una clara infrarrepresentación de este demostrativo en traducciones si se agrupan ambas funciones. El resultado del test estadístico disminuye con respecto a las cifras que arroja el análisis del determinante por sí solo, sin embargo, puesto que hay casi diez puntos menos de diferencia en el caso de la forma flexionada *este* y, además influye la sobrerrepresentación no significativa de la forma pronominal *esto*.

Demostrativo	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Este	352	32,38	95,89	345	27,51	70,62	-25,27	16,26 p < 0,0001
Esta	365	33,58	99,43	444	35,41	90,88	-8,55	1,61 (n. s.)
Estos	78	7,18	21,25	87	6,94	17,81	-3,44	1,28 (n. s.)
Estas	76	6,99	20,70	57	4,55	11,67	-9,04	10,86 p < 0,001
Esto	216	19,87	58,84	321	25,60	65,70	6,86	1,58 (n. s.)
Total Este	1087	100	296,11	1254	100	256,67	-39,44	11,85 p < 0,001

Tabla 72 - Comparación de frecuencias del demostrativo *este* en ES y en TRAD y significación estadística de las diferencias observadas

¹⁶⁰ Test exacto de Fisher: 0,081584 (n. s.)

5.3.2.2. Ese: cribado y clasificación manual

Determinante	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Ese	387	46,40	105,42	446	42,84	91,29	-14,13	4,28 p < 0,05
Esa	273	32,73	74,37	333	31,99	68,16	-6,21	1,14 (n. s.)
Esos	83	9,95	22,61	170	16,33	34,80	12,19	10,82 p < 0,01
Esas	91	10,91	24,79	92	8,84	18,83	-5,96	3,45 (n. s.)
Total Ese	834	37,28	227,19	1041	44,15	213,07	-14,12	1,90 (n. s.)

Tabla 73 - Comparación de frecuencias del determinante *ese* en ES y en TRAD y significación estadística de las diferencias observadas

En el caso del determinante *ese*, como puede apreciarse en la tabla 73, la diferencia total entre originales y traducciones no llega a ser significativa estadísticamente. Una única forma flexionada —*ese*— se encuentra infrarrepresentada, las diferencias en el caso de *esa* y *esas* no alcanzan el umbral para ser consideradas significativas, y *esos*, por su parte, se halla claramente sobrerrepresentada en traducciones.

Pronombre	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Ese	77	8,20	20,97	31	3,32	6,35	-14,62	35,58 p < 0,0001
Esa	51	5,43	13,89	43	4,60	8,80	-5,09	4,88 p < 0,05
Esos	20	2,13	5,45	9	0,96	1,84	-3,61	8,01 p < 0,01
Esas	9	0,96	2,45	8	0,86	1,64	-0,81	0,69 (n. s.)
Eso	782	83,28	213,03	843	90,26	172,55	-40,48	17,94 p < 0,0001
Total Ese	939	65,53	255,79	934	59,04	191,17	-64,62	39,58 p < 0,0001

Tabla 74 - Comparación de frecuencias del pronombre *ese* en ES y en TRAD y significación estadística de las diferencias observadas

Cuando hablamos de *ese* en función pronominal, sin embargo, la infrarrepresentación general está clarísima (v. tabla 74), pues solo una forma —*esas*— no presenta diferencias de frecuencia significativas entre subcorpus. Ocurre, pues, lo contrario a lo que sucedía en el caso del demostrativo de cercanía *este*, cuya forma determinante se hallaba infrarrepresentada mientras que su forma pronominal no lo estaba.

Demostrativo	ES	%	Frec. normalizada	TRAD	%	Frec. Normalizada	Dif. TRAD/ES	Significación estadística
Ese	464	26,17	126,40	477	24,15	97,63	-28,77	15,62 p < 0,0001
Esa	324	18,27	88,26	376	19,04	76,96	-11,30	3,25 (n. s.)
Esos	103	5,81	28,06	179	9,06	36,64	8,58	4,75 p < 0,05
Esas	100	5,64	27,24	100	5,06	20,47	-6,77	4,07 p < 0,05
Eso	782	44,11	213,03	843	42,68	172,55	-40,48	17,94 p < 0,0001
Total Ese	1773	100	482,99	1975	100	404,25	-78,74	29,46 p < 0,0001

Tabla 75 - Comparación de frecuencias del demostrativo *ese* en ES y en TRAD y significación estadística de las diferencias observadas

No obstante, tal como ocurría con el demostrativo de cercanía, al agrupar ambas funciones se obtiene una infrarrepresentación muy significativa de *ese*, a pesar de la sobrerrepresentación de la forma *esos* y la falta de significación estadística de *esa*, como se muestra en la tabla 75.

5.3.2.3. Aquel: cribado y clasificación manual

En el caso de este demostrativo podemos apreciar la infrarrepresentación en traducciones a simple vista, puesto que, tanto en pronombres como determinantes, hay menor número de casos en TRAD que en ES, a pesar de que el primer subcorpus cuenta con mayor número de palabras.

Determinante	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Aquel	286	42,43	77,91	248	42,11	50,76	-27,15	24,44 p < 0,0001
Aquella	264	39,16	71,92	241	40,92	49,33	-22,59	17,90 p < 0,0001
Aquellos	65	9,64	17,71	60	10,19	12,28	-5,43	4,17 p < 0,05
Aquellas	59	8,75	16,07	40	6,79	8,19	-7,88	11,12 p < 0,001
Total Aquel	674	30,13	183,60	589	24,98	120,56	-63,04	55,74 p < 0,0001

Tabla 76 - Comparación de frecuencias del determinante *aquel* en ES y en TRAD y significación estadística de las diferencias observadas

Como vemos en la tabla 76, en la que todas las celdas de significación estadística están pintadas de verde, *aquel* se encuentra claramente infrarrepresentado en TRAD en su forma determinante.

Pronombre	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Aquel	20	14,70	5,45	16	13,11	3,27	-2,18	2,32 (n. s.)
Aquella	18	13,23	4,90	14	11,48	2,87	-2,03	2,30 (n. s.)
Aquellos	7	5,15	1,91	2	1,64	0,41	-1,50	4,55 p < 0,05 ¹⁶¹
Aquellas	0	0,00	0,00	0	0,00	0,00	0,00	0,00 (n. s.)
Aquello	91	66,91	24,79	90	73,77	18,42	-6,37	3,98 p < 0,05
Total Aquel	136	9,49	37,05	122	7,71	24,97	-12,08	10,01 p < 0,01

Tabla 77 - Comparación de frecuencias del pronombre *aquel* en ES y en TRAD y significación estadística de las diferencias observadas

En el caso de los pronombres, la diferencia sigue siendo significativa pero el valor del test no es tan alto (v. tabla 77). Aunque *aquel* en función pronominal se

¹⁶¹ Test exacto de Fisher: 0,044237 (p < 0,05)

encuentra infrarrepresentado si sumamos los totales de todas sus formas, lo cierto es que esta infrarrepresentación parece ser producto sobre todo de la menor utilización de la forma neutra (*aquello*), puesto que, en los demás casos, excepto en el de *aquellos*, que aporta menos casos al total, la diferencia de frecuencias no llega a ser significativa.

Demostrativo	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Aquel	306	37,78	83,36	264	37,13	54,04	-29,32	26,71 p < 0,0001
Aquella	282	34,81	76,82	255	35,86	52,19	-24,63	20,01 p < 0,0001
Aquellos	72	8,89	19,61	62	8,72	12,69	-6,92	6,33 p < 0,05
Aquellas	59	7,28	16,07	40	5,63	8,19	-7,89	11,12 p < 0,001
Aquello	91	11,23	24,79	90	12,66	18,42	-6,37	3,98 p < 0,05
Total Aquel	810	100	220,65	711	100	145,53	-75,12	65,72 p < 0,0001

Tabla 78 - Comparación de frecuencias del demostrativo *aquel* en ES y en TRAD y significación estadística de las diferencias observadas

Al unir ambas funciones en la tabla 78, la mayor cantidad de demostrativos en función de determinantes inclina la balanza hacia la infrarrepresentación en todos los casos, y el valor del resultado del test estadístico resulta ser el más alto de todos los que hemos visto hasta el momento.

5.3.2.4. Tablas agrupadas: determinantes

Determinante	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Este	276	37,86	75,19	235	32,28	48,10	-27,09	25,42 p < 0,0001
Esta	313	42,94	85,26	366	50,27	74,91	-10,35	2,81 (n. s.)
Estos	70	9,60	19,07	72	9,89	14,74	-4,33	2,35 (n. s.)
Estas	70	9,60	19,07	55	7,55	11,26	-7,81	8,64 p < 0,01
Total Este	729	32,59	198,59	728	30,87	149,01	-49,58	29,95 p < 0,0001
Ese	387	46,40	105,42	446	42,84	91,29	-14,13	4,28 p < 0,05
Esa	273	32,73	74,37	333	31,99	68,16	-6,21	1,14 (n. s.)
Esos	83	9,95	22,61	170	16,33	34,80	-12,19	10,82 p < 0,01
Esas	91	10,91	24,79	92	8,84	18,83	-5,96	3,45 (n. s.)
Total Ese	834	37,28	227,19	1041	44,15	213,07	-14,39	1,90 (n. s.)
Aquel	286	42,43	77,91	248	42,11	50,76	-27,15	24,44 p < 0,0001
Aquella	264	39,16	71,92	241	40,92	49,33	-22,59	17,90 p < 0,0001
Aquellos	65	9,64	17,71	60	10,19	12,28	-5,43	4,17 p < 0,05
Aquellas	59	8,75	16,07	40	6,79	8,19	-7,89	11,12 p < 0,001
Total Aquel	674	30,13	183,60	589	24,98	120,56	-63,04	55,74 p < 0,0001
Total Det.	2237	100,00	609,38	2358	100,00	482,64	-126,74	62,15 p < 0,0001

Tabla 79 - Comparación de frecuencias de determinantes demostrativos en ES y en TRAD y significación estadística de las diferencias observadas

En la tabla 79 pueden verse agrupadas las cifras de utilización de todos los determinantes demostrativos, tanto en el subcorpus de español original como en el de español traducido. Como en las tablas parciales, incluimos también las frecuencias

normalizadas, los porcentajes correspondientes, la diferencia entre subcorpus y la significación estadística de dichas diferencias.

En este apartado, en el que nos referimos exclusivamente a los determinantes, observamos que, tomándolos a todos en conjunto, existe una infrarrepresentación muy significativa de dichos elementos en el subcorpus de traducciones al español (LL 62,15 $p < 0,0001$). Como hemos mencionado dentro de los apartados correspondientes, esta infrarrepresentación está presente en poco más de la mitad de los determinantes (*este*, *estas*, *ese* y todas las formas de *aquel*), mientras que la menor utilización en textos traducidos no llega a ser estadísticamente significativa en el caso de *esta* y *estos*, *esa* y *esas*, y, por supuesto, *esos*, que se encuentra sobrerrepresentado en las traducciones.

Este y *aquel* se encuentran infrarrepresentados en su función de determinantes, mientras que la diferencia de frecuencias que presenta *ese* no puede ser tomada en cuenta desde el punto de vista estadístico. En el gráfico de barras 1, que compara las frecuencias normalizadas halladas en ES y en TRAD de cada una de las tres formas básicas, podemos apreciar que *ese* es la forma que presenta la frecuencia de aparición más elevada y *aquel*, la más baja, tanto en español original como en español traducido.

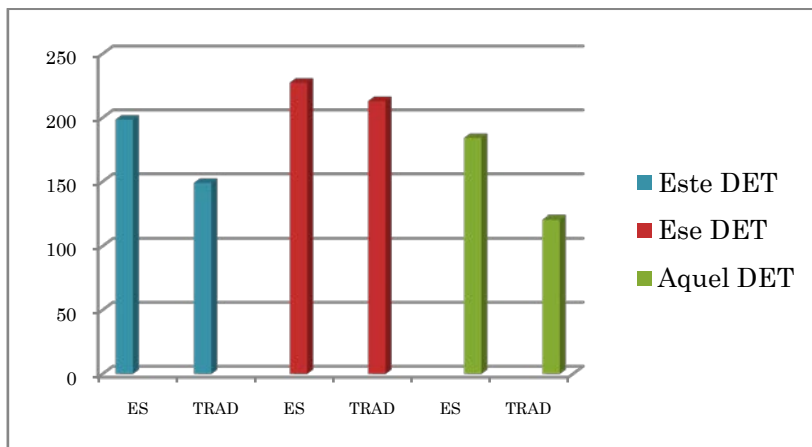


Gráfico 1 - Comparación de frecuencias normalizadas de determinantes demostrativos en ES y TRAD

El gráfico circular 2 muestra el porcentaje que representa dentro de cada subcorpus por separado cada uno de los tres grados del sistema (calculado sobre las frecuencias absolutas). En él podemos apreciar que en TRAD encontramos un 7% más de apariciones de *ese* y todas sus formas que en ES, en detrimento sobre todo de *aquel*, que se reduce un 5%. *Este* también presenta una frecuencia menor en TRAD pero con una diferencia de menos de dos puntos.

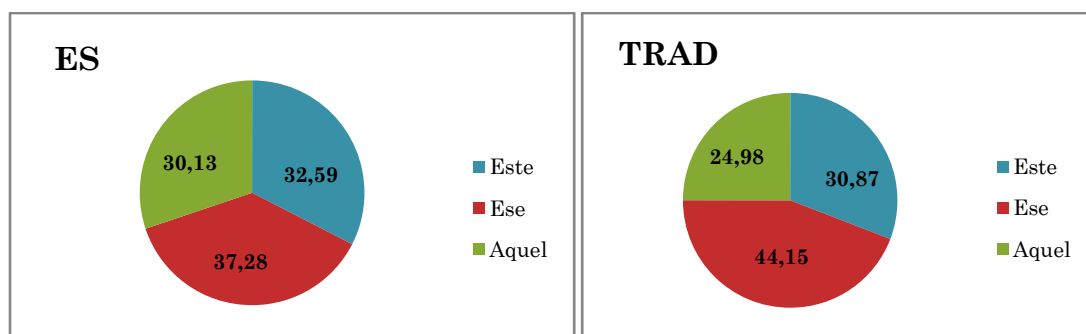


Gráfico 2 - Comparación de porcentajes de uso de determinantes demostrativos en ES y TRAD

Si queremos averiguar si existe una asimetría en el uso de las tres formas de demostrativos entre traducciones y originales, y si esta asimetría es estadísticamente significativa, podemos recurrir al test de chi-cuadrado¹⁶². Este test compara las frecuencias observadas empíricamente con las que cabría esperar según una distribución teórica en la que estas no se vieran afectadas por ninguna otra variable más que el azar, dado que pertenecerían a la misma población hipotética. Consiste en agrupar los datos en una tabla, en este caso de 2 x 3, realizar las sumas marginales y, a partir de ellas, extraer las frecuencias esperadas de los elementos analizados para, posteriormente, compararlas con las frecuencias observadas y determinar si la diferencia entre unas y otras es mayor al umbral de significación establecido por el test para este tipo de tablas al nivel del 5 % (5,99).

El resultado que nos da el test de chi cuadrado para nuestros datos es bastante superior al requerido: $\chi^2=25,40$ (df=2), $p < 0,001$. Podemos deducir entonces que la probabilidad de que esta distribución sea resultado del azar es muy baja. Por lo tanto, las diferencias que observamos entre subcorpus en cuanto a distribución de los determinantes demostrativos son altamente significativas.

5.3.2.5. Tablas agrupadas: pronombres

Cuando aplicamos el test de significación estadística de log-likelihood a cada par de frecuencias, vemos que entre los pronombres demostrativos la infrarrepresentación es menor.

¹⁶² En este apartado hemos seguido las pautas de Stefanowitsch (2005), procedimiento recogido también en Cantos Gómez (2002) y en el capítulo 8 de Lowry (1999-2012), aunque no para el campo específico de la lingüística de corpus. Ejemplos auténticos de la utilización de este test para la comparación de frecuencias en dos corpus diferentes pueden consultarse en Oakes (1998: 26-27) y Corpas Pastor (2008: 174).

Pronombre	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Este	76	21,23	20,70	110	20,91	22,52	1,81	0,32 (n. s.)
Esta	52	14,53	14,17	78	14,83	15,97	1,80	0,45 (n. s.)
Estos	8	2,23	2,18	15	2,85	3,07	0,89	0,63 (n. s.)
Estas	6	1,68	1,63	2	0,38	0,41	-1,23	3,40 (n. s.) ¹⁶³
Esto	216	60,34	58,84	321	61,03	65,70	6,86	1,58 (n. s.)
Total Este	358	24,98	97,52	526	33,25	107,66	10,14	2,10 (n. s.)
Ese	77	8,20	20,97	31	3,32	6,35	-14,62	35,58 p < 0,0001
Esa	51	5,43	13,89	43	4,60	8,80	-5,09	4,88 p < 0,05
Esos	20	2,13	5,45	9	0,96	1,84	-3,61	8,01 p < 0,01
Esas	9	0,96	2,45	8	0,86	1,64	-0,81	0,69 (n. s.)
Eso	782	83,28	213,03	843	90,26	172,55	-40,48	17,94 p < 0,0001
Total Ese	939	65,53	255,79	934	59,04	191,17	-64,62	39,58 p < 0,0001
Aquel	20	14,70	5,45	16	13,11	3,27	-2,18	2,32 (n. s.)
Aquella	18	13,23	4,90	14	11,48	2,87	-2,03	2,30 (n. s.)
Aquellos	7	5,15	1,91	2	1,64	0,41	-1,50	4,55 p < 0,05 ¹⁶⁴
Aquellas	0	0,00	0,00	0	0,00	0,00	0,00	0,00 (n. s.)
Aquello	91	66,91	24,79	90	73,77	18,42	-6,37	3,98 p < 0,05
Total Aquel	136	9,49	37,05	122	7,71	24,97	-12,08	10,01 p < 0,01
Total Pron.	1433	100	390,37	1582	100	323,81	-66,56	26,16 p < 0,0001

Tabla 80 - Comparación de frecuencias de pronombres demostrativos en ES y en TRAD y significación estadística de las diferencias observadas

Si comparamos los gráficos de barras 1 (en el apartado anterior) y 3 (que incluimos a continuación), inmediatamente comprobamos que la cantidad de pronombres es inferior a la de determinantes pero, especialmente, si nos referimos a *este* y *aquel*. El demostrativo pronominal *ese* cuenta con una frecuencia ligeramente superior a la del determinante, pero solo en el subcorpus de originales.

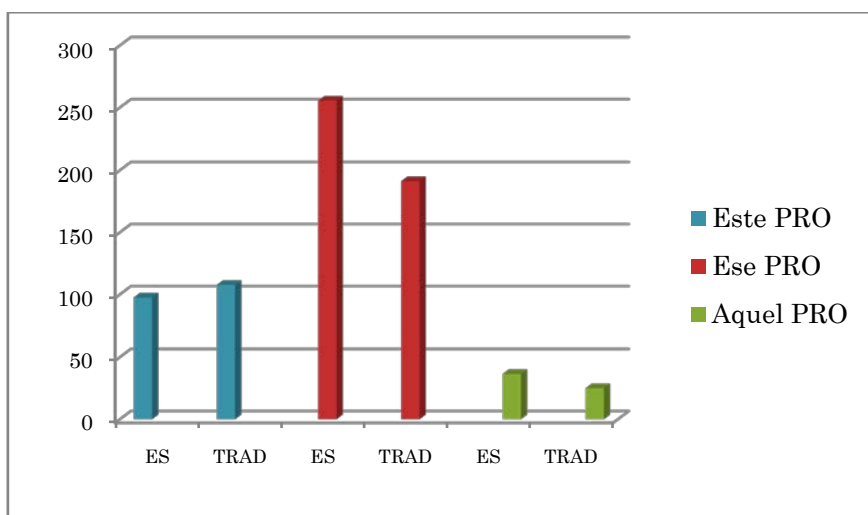


Gráfico 3 - Comparación de frecuencias normalizadas de pronombres demostrativos en ES y TRAD

Una vez más, *ese* es la forma que presenta la frecuencia de aparición más elevada y *aquel*, la más baja pero, en este caso, hay bastante más diferencia entre uno y

¹⁶³ Test exacto de Fisher: 0,081584 (n. s.)

¹⁶⁴ Test exacto de Fisher: 0,044237 (p < 0,05)

otro. La única forma pronominal que exhibe una mayor presencia en traducciones (una vez normalizadas las frecuencias) es *este*, aunque la diferencia no llegue a ser significativa en términos estadísticos. Los otros dos elementos se encuentran infrarrepresentados, como habíamos visto en las tablas correspondientes.

También observamos que los pronombres predominantes, con diferencia, son los neutros; especialmente *eso*¹⁶⁵. No es de extrañar, puesto que los pronombres neutros pueden aludir tanto a entidades físicas (por lo general, no humanas) como a hechos o situaciones o incluso segmentos textuales enteros, por lo que funcionan como un elemento deíctico y anafórico comodín —además de estar incluidos en formulas más o menos fosilizadas como «por eso»—, a lo que se suma el hecho de que *eso* viene a ser la forma no marcada en relación a la situación espacio-temporal, frente a *esto* y *aquello* (Macías Villalobos 2006: 172-73).

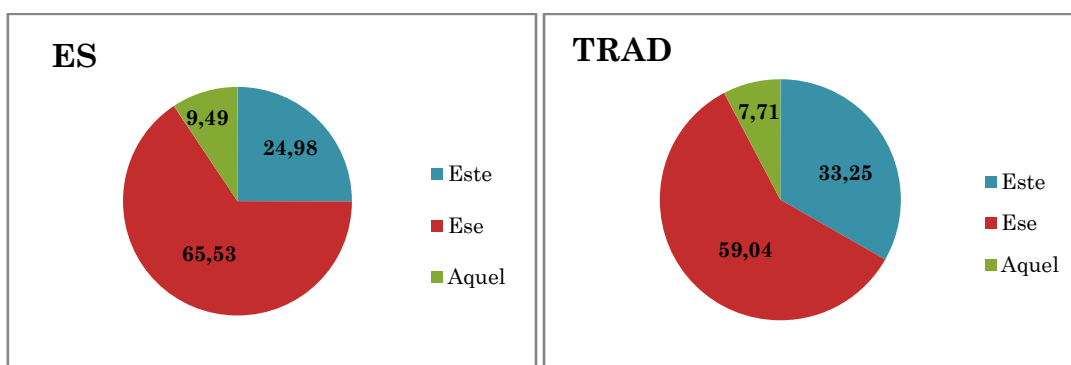


Gráfico 4 - Comparación de porcentajes de uso de pronombres demostrativos en ES y TRAD

Como observamos en el gráfico 4, a pesar de que tanto en el subcorpus de traducciones como en el de originales es más frecuente la forma *ese*, su porcentaje de utilización con respecto a las otras dos formas es más de un 6 % menor en TRAD con respecto a ES, a diferencia de lo que sucedía en los determinantes. Al mismo tiempo, *este* cuenta con una representación un 8 % mayor que en originales. *Aquel*, una vez más, presenta un porcentaje de utilización frente a los otros dos pronombres menor que en los originales (1,78 %). Como ya habíamos apuntado, esta disminución parece ser producto sobre todo de la menor utilización comparativa de la forma neutra.

¹⁶⁵ Hecho que contradice un estudio previo sobre demostrativos (Torres 2012) en el que se sostiene que «dentro del grupo de las formas pronominales neutras, el ‘esto’, en español, resulta la fórmula a la que se recurre más habitualmente en la escritura». Sin embargo, el corpus analizado en dicho estudio era un grupo de síntesis escritas por estudiantes universitarios, y las características deícticas —entre las que contamos las distintas formas demostrativas— de los textos pueden variar mucho según su propósito o el género al que se adscriban (Richardson 1998: 131).

La menor presencia del pronombre neutro en traducciones, al menos en lo que respecta al segundo y el tercer grado del sistema de demostrativos, podría tener que ver con la predominancia en español del proceso conocido como «encapsulación». Álvarez de Mon y Rego (2001) define la encapsulación (*encapsulation*) como una forma de crear cohesión en los textos, conectando una frase determinada con las anteriores; es decir, recuperando (o encapsulando) información aportada previamente en el texto por medio de distintos recursos lingüísticos como sustantivos (p. ej. *problema*, *fenómeno* o *concepto*), nominalizaciones de procesos verbales, atributos o circunstancias —como al utilizar *the expansion* para hacer referencia a la información proporcionada en una frase anterior («causes the spray to *expand* supersonically»)—, pronombres (generalmente *this* y *that* en inglés), adverbios (como *then* o *so* en inglés y *así* en español) y encapsulación implícita mediante una palabra o expresión que implica la existencia de determinada información en el texto anterior (como cuando se utiliza la expresión comparativa «a more effective procedure», que implica que existe un *procedimiento* anterior con el que estamos realizando la comparación, aunque no hayamos utilizado explícitamente esta palabra).

La autora denomina el procedimiento opuesto, que sirve para anunciar al lector determinada información que aparecerá más adelante en el texto, *prospection*. En un artículo posterior (2006: 30), afirma que ambos procedimientos parecen funcionar de manera bastante similar en español y en inglés. Sin embargo, en español predomina la encapsulación de fragmentos de texto previos por medio de sustantivos, adverbios y pronombres, especialmente en grupos preposicionales, como en la expresión «esto es falso» o «por eso», mientras que el inglés se vale en mayor medida de las nominalizaciones, que facilitan la identificación del fragmento de texto «encapsulado», puesto que frecuentemente el verbo nominalizado se encuentra en el texto precedente (como vimos con el ejemplo de *expansion*).

El trabajo mencionado se ocupa de textos de carácter científico, pero podría hipotetizarse que quizás también en otros tipos de texto el demostrativo utilizado en procedimientos de encapsulación es menos frecuente en inglés que en español y esto podría provocar que no se utilizara tanto en las traducciones de este idioma y, en su lugar, se tendiera a usar más sustantivos de lo que sería habitual en español no traducido.

La diferencia de distribuciones que podemos observar en la comparación de gráficos circulares número 4 también se confirma significativa tras aplicar el test de chi cuadrado a los datos: $\chi^2=25,40$ (df=2), $p < 0,001$.

5.3.2.6. Tablas agrupadas: determinantes y pronombres

Hemos comprobado en los dos apartados anteriores que, en el caso de los pronombres, la mayor infrarrepresentación, con diferencia, se da en *ese*. *Aquel* también se encuentra infrarrepresentado pero en menor medida y *este* no lo está. Ahora bien, en el de los determinantes, la infrarrepresentación de *aquel* es aplastante y, curiosamente, *ese* no es el segundo elemento que se encuentra infrarrepresentado, sino *este*.

Demostrativo	ES	%	Frec. norm.	TRAD	%	Frec. norm.	Dif. TRAD/ES	Significación estadística
Este	352	32,38	95,89	345	27,51	70,62	-25,27	16,26 p < 0,0001
Esta	365	33,58	99,43	444	35,41	90,88	-8,55	1,61 (n. s.)
Estos	78	7,18	21,25	87	6,94	17,81	-3,44	1,28 (n. s.)
Estas	76	6,99	20,70	57	4,55	11,67	-9,04	10,86 p < 0,001
Esto	216	19,87	58,84	321	25,60	65,70	6,86	1,58 (n. s.)
Total Este	1087	29,62	296,11	1254	31,83	256,67	-39,44	11,85 p < 0,001
Ese	464	26,17	126,40	477	24,15	97,63	-28,77	15,62 p < 0,0001
Esa	324	18,27	88,26	376	19,04	76,96	-11,30	3,25 (n. s.)
Esos	103	5,81	28,06	179	9,06	36,64	8,58	4,75 p < 0,05
Esas	100	5,64	27,24	100	5,06	20,47	-6,77	4,07 p < 0,05
Eso	782	44,11	213,03	843	42,68	172,55	-40,48	17,94 p < 0,0001
Total Ese	1773	48,31	482,99	1975	50,13	404,25	-78,74	29,46 p < 0,0001
Aquel	306	37,78	83,36	264	37,13	54,04	-29,32	26,71 p < 0,0001
Aquella	282	34,81	76,82	255	35,86	52,19	-24,63	20,01 p < 0,0001
Aquellos	72	8,89	19,61	62	8,72	12,69	-6,92	6,33 p < 0,05
Aquellas	59	7,28	16,07	40	5,63	8,19	-7,89	11,12 p < 0,001
Aquello	91	11,23	24,79	90	12,66	18,42	-6,37	3,98 p < 0,05
Total Aquel	810	22,07	220,65	711	18,04	145,53	-75,12	65,72 p < 0,0001
Total Dem.	3670	100	999,75	3940	100	806,45	-193,31	87,34 p < 0,0001

Tabla 81 - Comparación de frecuencias de demostrativos en ES y en TRAD y significación estadística de las diferencias observadas

Cuando aplicamos el test de log-likelihood a ambas categorías de manera conjunta, sin embargo, se confirma la infrarrepresentación de los tres demostrativos en traducciones: pronombres y determinantes demostrativos se encuentran infrarrepresentados en todos los casos en traducciones, como se observa en la tabla 81 y en el gráfico 5.

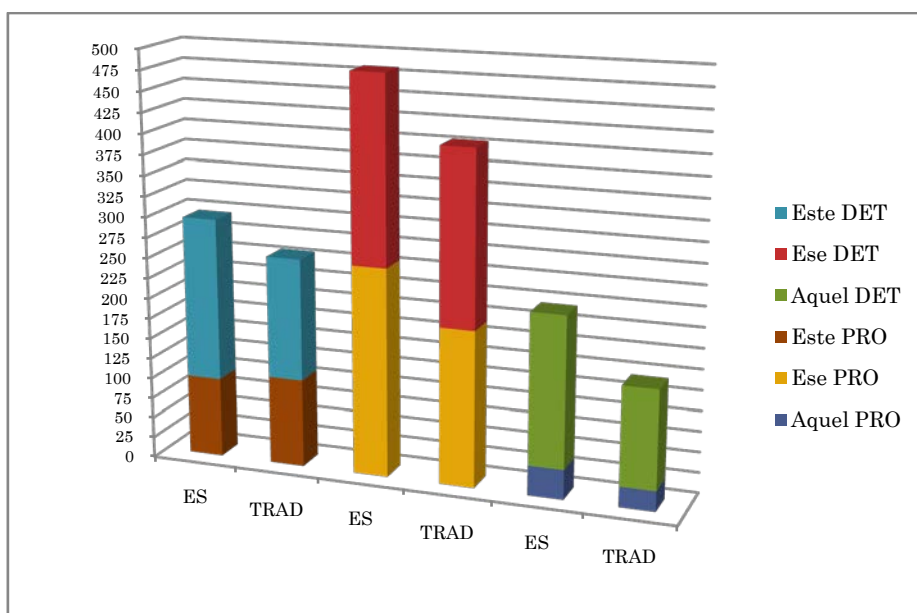


Gráfico 5 - Comparación de frecuencias normalizadas de demostrativos en ES y TRAD

Esto contradice lo que esperábamos, pues pensábamos que solo un elemento correspondiente al grado de lejanía en español —probablemente *aquel*—, lo estaría. Labrador (2011) había encontrado en su corpus de traducciones de inglés a español que existía una sobrerrepresentación del demostrativo sobre todo en sus formas neutras (*esto*, *eso* y *aquello*) y una infrarrepresentación significativa en tres de las cuatro formas del tercer grado de lejanía (*aquella*, *aquellos* y *aquellas*). Nuestros resultados no coinciden totalmente con los suyos, pero podemos confirmar la siguiente conclusión de la autora, con algún matiz: «However, in more than half the variants (15 out of 27), there are no significant differences in use, although in general Spanish demonstratives seem to be used less frequently in translations than they are in nontranslated texts» (Labrador 2011: 82). En nuestro caso son 14 las variantes con diferencias no significativas y, por otro lado, sus resultados generales para todas las formas arrojan una infrarrepresentación no significativa, cuando en nuestro corpus es muy significativa.

La comparación de porcentajes de uso (v. gráfico 6) y el test de chi cuadrado —que arroja un resultado de $\chi^2=19,68$ (df=2) $p < 0,001$ y, por tanto, confirma que la diferente distribución de formas en los subcorpus es significativa— nos indican que, mientras que los otros dos demostrativos cuentan con un porcentaje de utilización mayor en traducciones con respecto al total de formas demostrativas (*ese* es el demostrativo más utilizado de los tres), *aquel* se utiliza en menor medida, si lo comparamos con la distribución de porcentajes del subcorpus de textos originales en

español. Esto podría deberse al menor uso en traducciones de estructuras lexicalizadas que incluyen este demostrativo, a la reducción de la utilización del demostrativo neutro *aquello*, como hemos visto en la tabla 77, o a la naturaleza más oral de las novelas estadounidenses, que favorece la utilización de *ese*. De Kock afirmaba en su estudio sobre los pronombres (1992: 80) que las formas concordantes de *ese* están prácticamente ausentes en lengua escrita y las de *aquel* en lengua hablada.

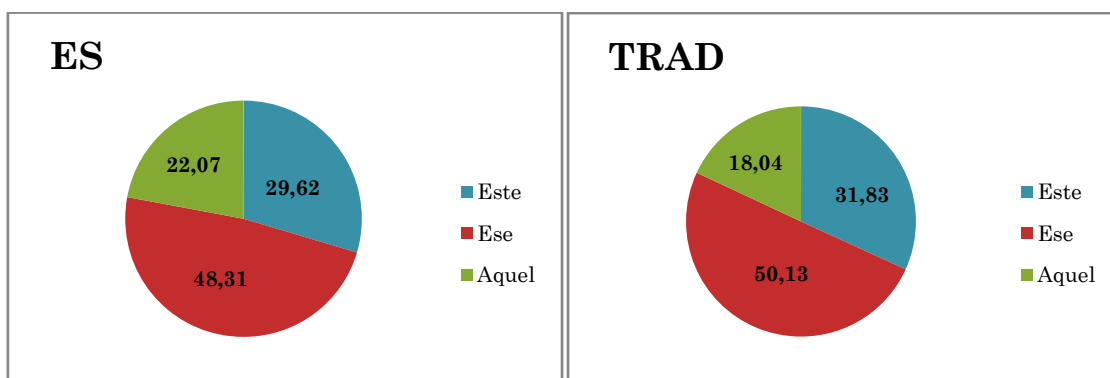


Gráfico 6 - Comparación de porcentajes de uso de demostrativos en ES y TRAD

5.3.3. Análisis del corpus paralelo

Para intentar analizar las causas de la presencia de infrarrepresentación de determinados elementos, o su ausencia, acudimos al componente paralelo del corpus.

Demostrativo	Frecuencia
This	1912
These	236
That	1692
Those	262
TOTAL	4102

Tabla 82 - Frecuencias de aparición de los demostrativos etiquetados como tales en el subcorpus EN

Si realizamos una búsqueda general de los determinantes y pronombres demostrativos en inglés (etiquetados ambos como «DT», *determiner*), obtenemos los

resultados que recoge la tabla 82¹⁶⁶. Una vez más, es necesario tener en cuenta el margen de error comentado anteriormente en relación con el etiquetado de TreeTagger. *That* puede introducir oraciones relativas, casos en los que TreeTagger lo considera pronombre relativo y, por tanto, la etiqueta que utiliza es «WDT» (*wh-determiner*). Por otro lado, también puede estar etiquetado como «IN» (*subordinating conjunction*) cuando el programa considera que introduce complementos del nombre: el ejemplo que se proporciona en la documentación es «the fact *that* you're here». También están los casos de *that* como conjunción completiva («He told me *that*...»), de los que no se habla en las pautas de etiquetado consultadas para el inglés (Santorini 1991) pero que el programa etiqueta igualmente como «IN». Además, también hay que tener en cuenta las concordancias en las que *that* cumple función de adverbio («I wasn't *that* worried about them»). Por lo tanto, podríamos hallar casos de demostrativos etiquetados como conjunciones, pronombres o adverbios (para los que utiliza la etiqueta «RB») y viceversa.

Tras revisar las más de cuatro mil quinientas concordancias que arroja una búsqueda simple de *that* y *those* —y demás variantes como la inclusión de apóstrofes y guiones pegados a la palabra— para comprobar si existen errores en el etiquetado, modificamos el valor del cómputo de 1692 a **2570** y de 262 a **267**. En el caso de *this* y *these*, después de añadir los casos no recuperados por la búsqueda sobre el etiquetado y eliminar 17 ocasiones en las que cumple función de adverbio, el cómputo final queda en **1937** y **238** y la suma total de demostrativos en EN es **5012**. La diferencia a favor de los originales ingleses en cuanto a la utilización de demostrativos es muy significativa cuando se compara con ES (5012 en EN frente a 3670¹⁶⁷ en ES: LL 17,05 p < 0,0001), pero más cuando se lo compara con sus traducciones al español (5012 en EN frente a 3940¹⁶⁸ en TRAD: LL 206,07 p < 0,0001).

A continuación comparamos la distribución de demostrativos de cercanía y lejanía por textos en los subcorpus de originales ingleses (EN) y sus respectivas traducciones al español (TRAD).

¹⁶⁶ Estas cifras confirman una de las conclusiones del estudio de Botley y McEnery (2001: 12): «(...) across all three corpora, the singular demonstratives *this* and *that* tend to be more common for nearly all features than is the case with the plural demonstratives *these* and *those*. This result is not surprising given that singular entities tend to be strongly marked (Croft 1990)». En el corpus utilizado para la redacción de la gramática de Biber et ál. (1999: 274) también resultan más frecuentes las formas singulares. En esta misma obra se afirma que la frecuencia de formas de cercanía y lejanía en los corpus de ficción y conversación (no así en textos expositivos y prosa académica) es bastante pareja, lo que también encaja con nuestros resultados.

¹⁶⁷ Si normalizamos estas frecuencias, los valores serían 1093 para EN y 1000 para ES.

¹⁶⁸ Si normalizamos estas frecuencias, los valores serían 1093 para EN y 806 para TRAD.

EN	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Ellroy	570	611,9	Cornwell	377	343,1
Cornwell	501	488,4	Ellroy	297	270,6
Grafton	446	476,6	Grafton	236	254,2
Leonard	334	438,8	Highsmith	175	178
Highsmith	324	348,5	Leonard	169	217,2
Total	2175	474,5	Total	1254	256,7

Tabla 83 - Distribución de demostrativos de cercanía por texto en EN (*this* y *these*) y en TRAD (*este* y formas flexionadas)

Hay casi el doble de demostrativos de cercanía en EN que en TRAD. Al parecer, los traductores no se dejan influenciar por la abundancia de *this* y *these* existente en los originales ingleses y los omiten si no son imprescindibles o encuentran otras formas de trasladarlos al español. Podemos apreciar aquí las diferencias provocadas por las características particulares de los textos o, en este caso, más probablemente por las preferencias individuales de los traductores, puesto que, a pesar de que las frecuencias de uso se reducen en todos los textos de TRAD, no lo hacen en la misma medida: la traducción de Cornwell de Mustieles sube del segundo puesto en EN al primero en TRAD.

EN	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Ellroy	678	727,9	Ellroy	764	696,1
Leonard	606	796,2	Leonard	634	815
Grafton	601	642,2	Highsmith	541	550,4
Cornwell	500	487,4	Cornwell	399	363,1
Highsmith	452	486,2	Grafton	348	374,8
Total	2837	618,9	Total	2686	549,8

Tabla 84 - Distribución de demostrativos de lejanía por texto en EN (*that* y *those*) y en TRAD (*ese* y *aquel* y formas flexionadas)

En cuanto a la utilización de demostrativos de lejanía, comprobamos que no es cuestión únicamente del traductor, puesto que Leonard, a pesar de contar con el menor número de palabras en su texto, utiliza comparativamente más demostrativos de lejanía que todos los demás autores. Vemos el mismo caso pero a la inversa con respecto al texto de Cornwell, el más extenso, que sin embargo ocupa la anteúltima posición en EN y la última en TRAD. En las traducciones de Leonard y Highsmith se incrementa el uso de demostrativos de lejanía con respecto a los originales, mientras que sucede lo contrario en las de Cornwell y Grafton. A pesar de que aquí las cifras totales de ambos subcorpus son parecidas, dado que TRAD contiene menor número de palabras, el test de log-likelihood nos indica que también existe una diferencia significativa a favor de los originales en el uso de demostrativos de lejanía (LL +19,36 p < 0,05).

Queda claro que en inglés se hace un uso de los demostrativos, sobre todo de los de cercanía que, para la norma española, es excesivo. Por tanto, la tendencia observada es a la baja, los traductores eliminan demostrativos o los reemplazan por otras estructuras y acercan así el lenguaje de las traducciones a la norma que observamos en el subcorpus de originales españoles. Por otra parte, ya hemos comentado que en las traducciones hallamos aún menos demostrativos que en español original, lo que indica que el traductor no se deja influir por el original en cuanto a los usos foráneos del demostrativo pero, al mismo tiempo, no está haciendo uso de estos elementos en la misma medida que los autores en lengua española.

A partir del análisis de las concordancias paralelas, intentaremos hallar claves que nos indiquen qué puede estar causando la menor utilización de los demostrativos en las traducciones al español; es decir, si este fenómeno puede deberse a la influencia de la lengua de los originales o, por el contrario, a determinadas características inherentes al proceso de traducción. Trataremos de determinar si el patrón de utilización de las tres formas se altera de alguna manera debido a la falta de una estructuración equivalente en el idioma original, aunque esto conlleva cierta dificultad, puesto que, como mostraremos a continuación, ni todos los *este*, por poner un ejemplo, son traducción de *this*, ni todos los *this* se traducen por *este*. Cuenca y Ribera (2011: 378) apuntan que la traducción «(...) often implies the underspecification of the subjective perspective of the source text (neutralization), or the introduction of the speaker's subjectivity (overmarking)». Por ejemplo, en su corpus paralelo de novelas escritas en inglés y traducidas al catalán, los autores encuentran que solo en el 58,5 % de los casos un demostrativo traduce a otro demostrativo, ya que los demostrativos ingleses muchas veces se neutralizan, es decir, se omiten (por ejemplo, cuando forman parte de expresiones idiomáticas) o se sustituyen, sobre todo por artículos definidos o pronombres. Lo mismo podría ocurrir al revés, es decir, que un demostrativo del texto meta provenga, por ejemplo, de un artículo en el texto original, aunque estos casos tendrán que ser, por fuerza, menos frecuentes, dada la infrarrepresentación de estos elementos en traducciones hallada en nuestros datos.

Una primera pregunta que podríamos hacernos es cómo se traduce el demostrativo *that* y su plural en todos los casos. ¿Es *ese* una solución más frecuente que *aquel* a la hora de traducir el demostrativo de lejanía?

Elemento equivalente TRAD	Frecuencia	Porcentaje
ESE	975 ¹⁶⁹	34,37
Omisión	489	17,24
Artículo determinado + sintagma nominal o preposicional u oración de relativo	337	11,88
AQUEL	242	8,53
Pronombre personal acusativo	227	8,00
Reformulación, frase hecha o fórmula conversacional	224	7,89
Así	116	4,09
ESTE	87	3,07
Adverbio demostrativo locativo	35	1,23
Posesivo	24	0,85
Artículo indeterminado	21	0,74
Expresión temporal	19	0,67
Dicho, tal, mismo	14	0,49
Ello	12	0,42
Tan, tanto	8	0,28
Quien/quienes	7	0,25
Total	2837	100

Tabla 85 - Elementos equivalentes de los demostrativos *that* y *those* en traducciones

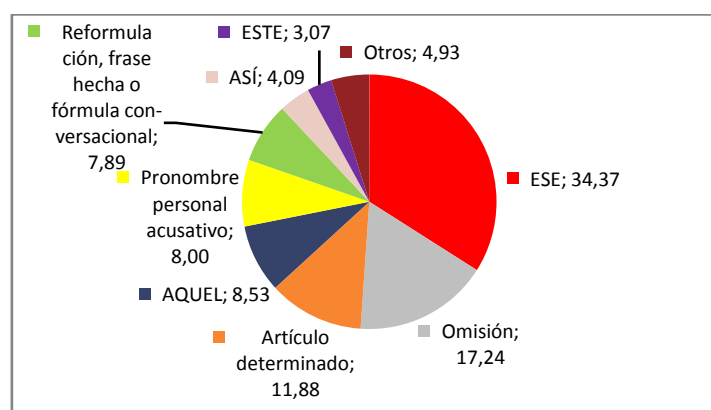


Gráfico 7 - Porcentajes de uso de los distintos elementos equivalentes de los demostrativos *that* y *those* en traducciones

Después de un análisis exhaustivo de todas las concordancias paralelas en las que *that* y *those* se utilizan efectivamente como demostrativos, concluimos que solo en un 42,90 % de los casos (1217 de 2837) se traducen por los demostrativos teóricamente equivalentes en español. A *ese* corresponde el 80,11 % de estas traducciones (975) y a

¹⁶⁹ Las cifras totales de demostrativos ingleses traducidos por demostrativos españoles presentadas en las tablas 85 y 86 en ocasiones no coinciden completamente con la suma de las cifras de demostrativos españoles traducidos por demostrativos ingleses incluidas en las tablas 87 a 92. Por ejemplo, en el caso de *ese*->*that* y *that*->*ese* encontramos una diferencia de 41 casos; la mayor de todas, puesto que en *este*->*this* y en *aquel*->*that* solo es de 4 casos, en *aquel*->*this* de 2, en *este*->*that* de 1 y en *ese*->*this* no existe diferencia. A pesar de haber revisado las concordancias a conciencia, existe la posibilidad de que haya habido algún error de cómputo. Sin embargo, las diferencias pueden explicarse si tenemos en cuenta que la clasificación de los ejemplos en una u otra categoría puede diferir según la dirección en la que consultemos el corpus. Por ejemplo, muchas veces la traducción de *that* es una expresión como «eso es», que no incluimos en la categoría «ese» en esta tabla, sino en la de reformulaciones, pero sí hemos contado como un caso más de *ese* en la tabla que recoge los equivalentes en EN de este demostrativo. También existen casos en los que el demostrativo no está escrito de manera estándar y, por tanto, no se recoge cuando consultamos el corpus en inglés, pero sí cuando buscamos equivalentes del demostrativo en español:

500185: - Tío , <esa> semilla hace que todo el mundo sonría y enseñe los dientes
-->par_en: Ordell said , " Dot ganja , mon , mek everyone smile to show their teet , uh ?

aquel solo un 19,88 % (242). Esto significa que 1 000 casos de *ese* y 469 de *aquel* —el 50,63 % y el 65,96 %, respectivamente, de los casos totales de estos demostrativos en TRAD (v. tabla 81) — aparecen por otros motivos, que son los elementos que veremos a continuación, en el análisis de concordancias paralelas.

Habíamos observado antes que un 73,53 % del total de los casos de demostrativo de lejanía (1975) en traducciones correspondía a *ese* y un 26,47 % (711) a *aquel*. Vemos que hay un 6 % más de utilización de *aquel* que el que surge únicamente de la traducción de los demostrativos ingleses (que, como acabamos de ver, es el 19,88 % del total); quizás por adiciones del traductor, o porque hay elementos en el original que no son demostrativos pero activan el tercer grado de nuestro paradigma a la hora de traducir. Aun teniendo en cuenta este incremento, el porcentaje de *aquel* no alcanza al hallado en ES: de 2583 demostrativos de lejanía encontrados en el subcorpus de originales españoles, el 68,64 % (1773) corresponde a *ese* y el 31,36 % (810) a *aquel*, lo que significa un 4,89 % más que en TRAD y un 11,48 % más que en los casos provenientes exclusivamente de la traducción de *that* y *those*.

En la tabla 86 podemos observar los equivalentes hallados en TRAD para el demostrativo *this* y su plural *these*. En este caso, el demostrativo inglés se traduce por el demostrativo español teóricamente equivalente el 34,62 % de las veces (753 de 2175 concordancias). Por tanto, un 40 % de casos de *este* (501 de 1254) en TRAD tienen un elemento activador distinto a *this*. No se observa una diferencia tan grande entre *este* y *ese*, que ocupan respectivamente el primer y segundo puesto de la tabla, como la que veíamos en el caso de los equivalentes de *that* y *those*. De hecho, si solo tuviésemos en cuenta las concordancias de la forma plural, *these*, veríamos que el porcentaje de veces que se traduce por *ese* y sus formas derivadas es mayor que el de los casos en los que se traduce por *este*, su equivalente más inmediato. *Aquel* es bastante menos frecuente en esta tabla; parece representar un grado de lejanía demasiado extremo para la traducción del demostrativo de cercanía inglés, pero, aun así, está presente en casi un 7 % de casos.

Elemento equivalente TRAD	Frecuencia	Porcentaje
ESTE	753	34,62
ESE	328	15,08
Omisión	254	11,68
Artículo determinado + sintagma nominal o preposicional u oración de relativo	220	10,12
AQUEL	144	6,62
Reformulación, frase hecha o fórmula conversacional	135	6,21
Pronombre personal acusativo	89	4,09
Artículo indeterminado	89	4,09
Así	36	1,65
Adverbio demostrativo locativo	35	1,61
Posesivo	27	1,24
Expresión temporal	23	1,06
Ello(s) y otros pronombres personales	15	0,69
Dicho, tal	15	0,69
Tan, tanto, demasiado, un montón	12	0,55
Total	2175	100

Tabla 86 - Elementos equivalentes de los demostrativos *this* y *these* en traducciones

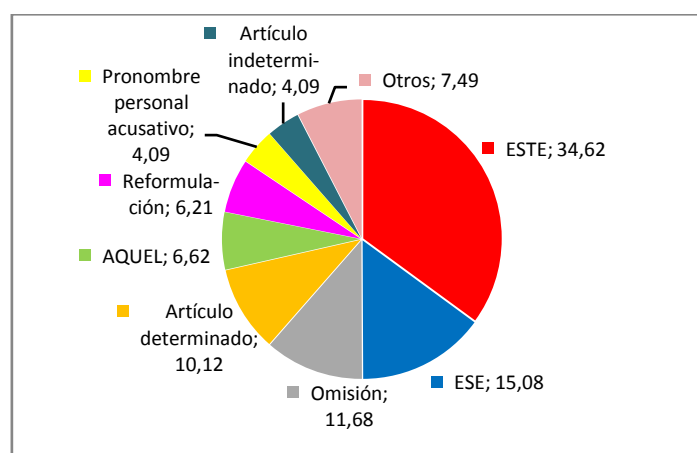


Gráfico 8 - Porcentajes de uso de los distintos elementos equivalentes de los demostrativos *this* y *these* en traducciones

En un 17,24 % de las concordancias de *that* y *those* y un 11,68 % de *this* y *these*, el demostrativo se omite. Esto sucede, sobre todo, a principio de frase, donde en español resulta, a veces, innecesario. Lo mismo sucede en catalán (Ainaud, Espunya y Pujol 2003: 182), en donde es normal elidir el demostrativo en función de sujeto salvo que aporte información imprescindible para resolver ambigüedades.

<Ellroy>: This seller man Pat Orchard , he knows Steve , and he was in jail this afternoon . This policeman strongarmed him - " " <That> was me . "

→par_trad: - Fui yo .

<Grafton>: Ray had pulled on a fleece-lined denim jacket , and as we got in the car , he tossed a sweatshirt to each of us . Gratefully , I pulled the sweatshirt over my head and wore my blazer over <that> .

→par_trad: Me puse aliviada la sudadera por la cabeza y encima me puse la chaqueta .

De la misma manera, en inglés informal, a veces se utiliza el demostrativo *this* delante de sustantivos cuando se introduce a un personaje o elemento por primera vez (Sinclair et ál. 1990: 49; Biber et ál. 1999: 274). Es frecuente utilizarlo para transmitir una sensación de familiaridad con la información que se presenta o para dar a entender que a continuación se nos proporcionará la oportunidad de familiarizarnos con ella (Chalker 1984: 57). Esta función es peculiar del inglés (en comparación con el español) y no se corresponde con ningún demostrativo en nuestra lengua, lo que indica que las diferencias de frecuencias halladas podrían explicarse por el hecho de que una forma determinada realice más o menos funciones en una lengua que en la otra.

<Grafton>: " Anyway , in the late forties , I started writing to <this> woman named Maria I met through a pen pal ad .

→par_trad: - A fines de los años cuarenta - prosiguió - me puse a escribir cartas a una mujer llamada María , que conocí por un anuncio de solicitud de correspondencia .

<Grafton>: He told them he 'd done construction , so they put him to work . Anyway , while he 's crawling around <this> basement one day , he realizes he 's looking at the underside of a bank .

→par_trad: El caso es que estaba recorriendo un sótano cierto día cuando se da cuenta de que está debajo de un banco .

Se observa también la utilización en estos casos del demostrativo de lejanía *ese*, que, como ya hemos dicho, en las concordancias de *these* llega a abarcar más equivalencias que su equivalente más directo, *este*.

<Leonard>: Once in a while Ordell gets into the rough stuff , but usually he plays it safe , has <these> crazy black kids that work for him do the heavy shit .

→par_trad: De vez en cuando , Ordell se mete en algún follón , pero normalmente juega sobre seguro , tiene esos negritos locos que trabajan para él y hacen los encargos duros .

<Cornwell>: He was a prime candidate for a stroke or a heart attack . " Plus , my ankles swell , " he went on . " My feet hurt and I get <these> damn headaches ..

→par_trad: Me duelen los pies y tengo esos malditos dolores de cabeza .

También se lo reemplaza por elementos habituales en español que hacen innecesaria la presencia del demostrativo. Nos referimos, por ejemplo, a las categorías

que incluyen un artículo determinado¹⁷⁰, o un pronombre personal acusativo, dos categorías que conforman el 14,21 % de las concordancias.

<Leonard>: Nineteen years dealing with people who took incredible risks . If he walked over to <that> counter he'd find out what it was like . After , he was to go home and wait for Jackie's call .

→par_trad: Si se acercaba a **la** caja descubriría qué sentía . Luego tenía que irse a casa y esperar la llamada de Jackie .

<Highsmith>: " M'sieur Pritchard had a friend-or in any case a man with him when he was fishing in the canals . Everyone says <that> . "

→par_trad: Todo el mundo **lo** dice .

Aunque en él se trate otra combinación lingüística, podemos referirnos al trabajo de Goethals (2013), quien estudió sintagmas nominales introducidos por determinantes demostrativos, por un lado, y por artículos definidos, por otro, y descubrió que, tanto en una dirección de traducción como en la otra, los demostrativos en holandés corresponden con bastante frecuencia a artículos definidos en español. Hipotetiza Goethals (2013: 544) que, en ciertas ocasiones, el demostrativo español puede resultar «demasiado deíctico» para traducir el demostrativo holandés, que se encuentra más gramaticalizado: cuando el antecedente se repite (es decir, no se hace uso de una anáfora metonímica, hiperonímica ni metafórica), esta repetición literal es suficiente para establecer la conexión; un demostrativo, aunque sería gramaticalmente correcto, tendría demasiada fuerza deíctica. Por otra parte, también sugiere que el artículo definido español tiene suficiente fuerza deíctica y capacidad de identificación del referente; al menos más que en holandés. Si lo fuera también más que en inglés, podría explicarse el hecho de que se utilice tanto en nuestro corpus para traducir demostrativos sin que esto signifique necesariamente una pérdida de fuerza anafórica o deíctica. Esta opinión es apoyada por las conclusiones de Cuenca y Ribera (2011: 275) para el catalán: «Catalan shows a tendency to avoid deictic marking in syntactic contexts where a demonstrative could be interpreted as too focal or somehow emphatic».

¹⁷⁰ Según Leonetti (1999: 800), tanto el artículo definido como los demostrativos «muestran propiedades anafóricas características de la definitud, por lo que pueden remitir a informaciones ya presentes en el contexto». Sin embargo, estos últimos podrían considerarse los elementos marcados, lo que significa que en muchos casos puede sustituirse un demostrativo por un artículo definido, pero no siempre ocurre a la inversa. Esto sucede porque el artículo permite identificar al referente unívocamente, pero el demostrativo se caracteriza por su contenido deíctico y exige, además, que dicho referente pueda ser percibido en la situación comunicativa o recuperado a partir de una mención previa en el discurso. Macías Villalobos (2006: 143-146), por su parte, señala que el artículo definido puede alternar con el demostrativo pero no son intercambiables en cualquier contexto, dado que este último individualiza mucho más a su referente y se asocia a una ruptura en la continuidad referencial, mientras que el artículo puede repetirse en una cadena textual larga.

Por supuesto, también hay veces en que la causa de la omisión podría considerarse menos evidente, como en los casos en los que el demostrativo sirve en inglés para enfatizar algún elemento del contexto o del discurso y en español no vemos un esfuerzo por mantener ese matiz:

<Highsmith>: " Nick-he's been with us barely a year , and didn't know anything about a Murchison who disappeared . " Ed Banbury said it as if Murchison had done just <that> , disappeared , whereas Tom had killed him .

→par_trad: Nick.. . bueno , apenas lleva un año con nosotros , y no sabía nada de un tal Murchison que había desaparecido . Ed Banbury lo dijo como si Murchison simplemente hubiera desaparecido , aunque sabía que Tom le había matado.

<Highsmith>: " Naturally , my housekeeper learned of the Pritchard tragedy this morning . " Murmurs from the officers . Of course , news like <that> would spread quickly ! "

→par_trad: ¡Desde luego , las noticias se extendían tan deprisa !

Sin embargo, también se observan estrategias de compensación en el sentido inverso o de adición de demostrativos debida a la utilización de expresiones más o menos lexicalizadas, por lo que es difícil extraer conclusiones que no resulten parciales sin embarcarnos en un análisis de magnitud considerable:

<Ellroy>: I did n't quite buy it-his hard-on for <that> family played too ugly . Like mine , like Junior 's .

→par_trad: No me convenció del todo : aquella obsesión por la familia tenía un aspecto demasiado feo

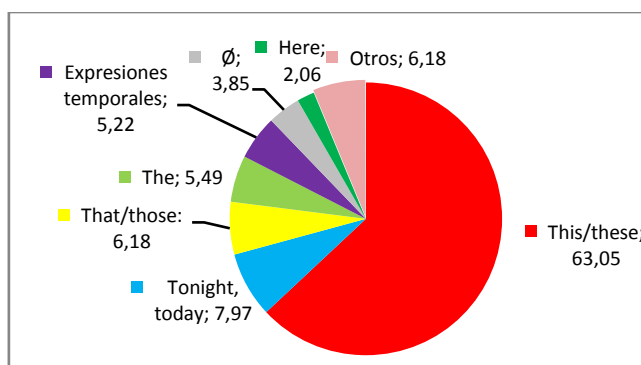
<Grafton>: Because then you have to wonder where he was all <that> time . "

→par_trad: Si no estuvo en la guerra , entonces cambia toda la historia . Porque en ese caso hay que preguntarse dónde estuvo durante todo aquel tiempo .

Con la esperanza, entonces, de llegar a comprender el funcionamiento del sistema de demostrativos estudiado en ambas lenguas, o al menos en las circunstancias que rodean al proceso de traducción, y que esto nos proporcione algunas pistas acerca de los interrogantes que planteamos en este trabajo, lo que hemos hecho es, a partir del análisis cualitativo de todas las concordancias paralelas en las que el traductor ha utilizado un determinante o pronombre demostrativo, intentar clasificar los elementos que han podido sugerir dicha traducción, teniendo en cuenta también los casos en los que no encontramos este elemento. A continuación, mostraremos qué tipo de elementos activan presuntamente la utilización de los determinantes y pronombres demostrativos en las traducciones al español.

5.3.3.1. Determinantes↘ ESTE

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
This/these	459	63,05
Tonight, today	58	7,97
That/those	45	6,18
The	40	5,49
Expresiones temporales: right now, by now, as we speak, currently, again, just (trad.: esta vez, en este momento, a estas alturas)	38	5,22
∅ ¹⁷¹	28	3,85
Here	15	2,06
Pronombres personales y nombres propios	13	1,79
Posesivos	9	1,24
Reformulación ¹⁷²	9	1,24
Such	6	0,82
In ¹⁷³	3	0,41
Ago, for	3	0,41
The latter	2	0,27
TOTAL	728	100

Tabla 87 - Elementos activadores de la utilización del determinante *este* en traduccionesGráfico 9 - Porcentajes de uso de los distintos elementos activadores del determinante *este* en traducciones

¹⁷¹ Nos referimos a los casos en los que no hemos podido identificar un elemento activador del demostrativo en español. El traductor lo añade por diversos motivos que pueden ser estilísticos, de cohesión o por una necesidad de explicitar determinados aspectos relativos al lugar, el tiempo o la causa. En todos los casos se ha intentado ampliar el contexto de la concordancia al máximo, para comprobar si el elemento activador podía encontrarse en otra parte del texto.

¹⁷² Nos referimos a los casos en los que no podemos asignar una categoría determinada al elemento activador, puesto que se utiliza un demostrativo para parafrasear expresiones idiomáticas o sin equivalente directo en español, o simplemente para alterar el fragmento original por diversos motivos como falta de comprensión, razones estilísticas o adaptación al contexto meta.

¹⁷³ En estos casos se utiliza el demostrativo junto con un sustantivo que permite identificar el elemento al que se hace referencia en inglés con la partícula *in*:

<Leonard>: O será que en <este> negocio todos tendemos a hablar el mismo idioma .

-->par_en: Or the business I'm **in** , you tend to speak the same language . "

<PatriciaCornwell>: que se instaló en <esta> casa. Pero, coño, si ni siquiera ha tirado de la cadena

-->par_en: " The damn toilet does n't look like it 's been cleaned since he moved **in** .

<JamesEllroy>: y , si se suma a <este> suceso el ajuste de cuentas entre bandas acaecido horas

-->par_en: and when calculating **in** the gangland slayings in Watts earlier today that also left four dead

Como era de esperar, la mayor parte de las veces que aparece el determinante demostrativo *este*, lo hace como consecuencia de la utilización en inglés de su equivalente más aproximado *this*, ya sea como determinante o como pronombre. Esto es así en todos los demostrativos estudiados y confirma parte de los resultados de Labrador (2010: 80), quien afirma que la traducción más frecuente para el demostrativo inglés es el demostrativo español, al menos en ficción, y que la segunda solución en orden de frecuencia es el artículo definido.

<Cornwell>: <Esta> misma tarde un varón blanco de trece años de edad...

→par_en: “ Earlier **this** evening , a thirteen...

<Cornwell>: <Esta> vez han ido demasiado lejos .

→par_en: **This** has gone too far . “

En gran parte de las concordancias de *esta* observamos elementos temporales que se repiten (como *mañana* o *tarde*). El sentido temporal que puede adoptar este demostrativo en determinados contextos se ve reforzado por el hecho de que la segunda y quinta categorías por orden de frecuencia están formadas justamente por expresiones de tiempo. En casi el 40 % del total de concordancias, sin embargo, *this* no es el elemento activador, como vemos en la tabla 87. Cuando el demostrativo cumple principalmente una función deíctica o fórica temporal, observamos la utilización de adverbios como *tonight*, *today*, *ago*, *now*, *again* y *just*, preposiciones como *for* o expresiones como «as we speak», que sumados todos, conforman un 13,60 % del total.

<Grafton>: He hablado con Chester <esta> mañana y me ha dicho que te habías ido en pos de una liebre .

→par_en: I talked to Chester **this** morning and he said you were off on some kind of wild goose chase .

<Highsmith>: – Sí , <esta> tarde , por teléfono .

→par_en: “ By telephone . **This** afternoon .

<Ellroy> – Hazlo <esta> noche y mañana ya no tendrás problemas con el sindicato .

→par_en: “ Do it **tonight** , and you wo n’t have union trouble tomorrow .

<Cornwell>: - <Esta> tarde , mientras estaba jugueteando en tu ordenador , he visto un fichero llamado Carne .

→par_en: “ While I was roaming around inside your computer **today** , I saw the file called ‘flesh .

Es curioso constatar como en un 6,18 % de los casos es el elemento deíctico teóricamente opuesto en significado (*that*) el que sirve al traductor para utilizar el demostrativo del que hablamos, generando un cambio de perspectiva en el fragmento.

<Cornwell>: Si quieres saber mi opinión , <esta> niña se va a ir al infierno de cabeza .

→par_en: You ask me , **that** child 's going to hell in a hand basket.

<Leonard>: Dice que hay que quitar la tapa trasera y <esta> cinta de aquí .

→par_en: “ Say remove the rear cover and **that** strap there . “

Como ilustran de manera muy adecuada Ainaud, Espunya y Pujol (2003: 214) mediante un ejemplo extraído de una de las últimas escenas de la película *Blade Runner*, «those moments» puede traducirse tanto por «estos momentos» como por «aquellos momentos», según si lo que se quiere potenciar es la remisión a un referente textual próximo —estos momentos que acabo de mencionar— o, de manera anafórica, a las experiencias descritas anteriormente. El cambio también puede tener que ver con la cercanía al emisor o al receptor que busque conseguir el mensaje o, simplemente, con la perspectiva que se tenga del elemento al que se hace referencia; por ejemplo, si hablamos de un elemento que nos resulta más cercano o se encuentra más arraigado en nuestra cultura, como podría ser algún aspecto religioso, y queremos resaltar este hecho (Montiel Ponsoda 1999: 67).

También puede darse el caso de que estos cambios no tengan una finalidad ni gramatical ni afectiva:

Some of the cases in which shifts happen for no grammatical reason may be partly accounted for by preference. Among the available forms within the repertoire of a language that compete as alternatives in a given situation, native speakers will generally come to agree on some forms that will be used most of the time; these will come to be perceived as preferred and unmarked (...) It is suspected, for example, that the proximal 'this' is preferred by Romanian speakers in some cases when an English speaker would opt for the distal 'that' (for instance 'I told you that!' would be equivalent to 'I told you this!' in Romanian) (...) (Mason y Șerban 2003: 274).

El siguiente elemento del que deriva la utilización del demostrativo de cercanía español es el determinante *the*, que en principio no indica nada relacionado con

cuestiones deícticas temporales o espaciales y deja libre al traductor para utilizar el demostrativo (o cualquier otro elemento) que convenga según el contexto.

<Grafton>: Yo no limpio <esta> pocilga ni loca.

→par_en: “ I ‘m not going to clean the damn place .

Cuenca y Ribera (2011: 376) describen en su corpus el fenómeno opuesto a lo que denominan neutralización, el *overmarking*, es decir, la adición del demostrativo, que suele implicar un cambio en el punto de vista narrativo y generalmente afecta a la traducción del artículo definido y del pronombre *it*. Así, aunque en este caso sea en proporción prácticamente anecdótica, encontramos otras equivalencias, como la especificación del referente de pronombres personales como *it* o *he* mediante un sintagma nominal introducido por un demostrativo en español¹⁷⁴:

<Grafton>: No me fiaba de ninguno de los presentes y <esta> convicción me hacía polvo el estómago .

→par_en: I did n’t trust any of them , and it was making my stomach hurt .

<Cornwell>: explicando que las prendas rellenas de <este> material son tan caras que se han convertido en piezas muy buscadas por los ladrones.

→par_en: and how items filled with it are so expensive that they ‘ve become a hot commodity for thieves

<Ellroy>: Además, desconfío de todo lo que aparece en la prensa ; <este> asunto huele a búsqueda de publicidad .

→par_en: And I ‘m wary of anything announced in the press : it smacks of publicity seeking .

<Cornwell>: Al principio , <esta> mujer compraba el silencio de Susan Story .

→par_en: First , she paid for Susan Story ‘s silence .

<Grafton>: No sé a quién habrá salido <esta> hija .

→par_en: I do n’t understand where she ‘s coming from . “

<Grafton>: No me preocupan estas dos , la vieja y ésa - dijo señalándome - . Pero a <esta> otra tengo que ajustarle las cuentas .

→par_en: “ But I got accounts to settle with her .

<Ellroy>: Un oficial de la Policía de Los Ángeles ha escapado a la custodia federal . <Este> oficial tenía que haber testificado contra miembros de la brigada de Narcóticos del LAPD

¹⁷⁴ Dice Richardson (1998: 124) al respecto, aunque refiriéndose a la dirección de traducción inversa: «For example, at the morphosyntactic level, deictic reference may, on occasion, give rise to changes in word class in translation (for instance, when a Spanish pronoun is rendered in English as a noun which has been mentioned previously in the text) (...)».

→**par_en**: A Los Angeles police officer skipped Federal custody. **He** was to have testified against members of the LAPD Narcotics Squad and a criminal family they have long been allied with ,

Por otra parte, es normal encontrar una sustitución de posesivos por demostrativos o incluso artículos, ya que en español los posesivos se utilizan menos. Al respecto podemos citar a Leonetti (1999: 808), quien señala:

La tradición gramatical hispánica recoge explícitamente la intuición de que el español muestra una clara preferencia por el empleo del artículo definido en contextos gramaticales en los que en principio cabría el uso del posesivo, peculiaridad en la que el español contrasta con las lenguas germánicas e incluso con otras lenguas románicas como el francés.

Dado que hemos visto que el artículo definido puede alternar con el demostrativo en determinadas situaciones, es lógico que algunos de los posesivos de nuestro corpus de inglés original se transformen en demostrativos en español. Además, como apuntan Cuenca y Ribera (2011: 377) para el catalán, aparte de sonar poco naturales, los posesivos podrían llegar también a crear ambigüedad a la hora de identificar al referente en español.

<Leonard>: El del pelo oscuro intervino : - Perdón , no he podido evitar observar <esta> situación .

→**par_en**: The dark-haired one said , “ Excuse me , I couldn’t help but observe **your** plight .

<Highsmith>: Espero que disfrutes de <esta> visita .

→**par_en**: We hope you will enjoy **your** visit .

<Highsmith>: Después de <estos> quehaceres , Tom se dio una ducha y se puso a leer la biografía de oscar Wilde.

→**par_en**: After **his** labors , Tom had a shower and read in the Oscar Wilde biography.

Lo mismo podemos decir de los nombres propios, cuya repetición excesiva en español puede considerarse un defecto estilístico:

<Leonard>: Tyler y Nicolet se miraron , y <este> último preguntó : - ¿Te ha dicho sus nombres ?

→**par_en**: Tyler and Nicolet looked at each other again and **Nicolet** said , “ She tell you their names?

A veces las adiciones no parecen responder a ninguna necesidad concreta, pero otras veces pueden responder a la de cohesionar el texto y hacerlo más fluido o menos ambiguo, mediante la explicitación de indicaciones temporales o conectores, fenómeno que conlleva una mayor utilización del demostrativo que nos ocupa. Este hecho es más comprensible, si cabe, en el texto de Ellroy (tercer ejemplo), que es tan escueto en conectores y verbos, y en los casos de lenguaje oral, donde el inglés es más libre de omitir ciertos elementos.

<Highsmith>: – Es curioso <este> sitio – le dijo Tom a Heloise -

→par_en: “ Curious , “ Tom said to Heloise .

<Cornwell>: <Esta> vez no nos tomamos la molestia de añadir hielo .

→par_en: We did n’t bother with ice .

<Ellroy>: <Esta> vez fue él , Dudley Liam Smith , quien me abrazó .

→par_en: Embracing me-Dudley Liam Smith .

<Cornwell>: <Estas> huellas de mordiscos , aquí y aquí , coinciden precisamente con las zonas de carne extirpada en el hombro y el muslo de Eddie Heath.

→par_en: “ Bite marks here and here corresponding closely with the areas of missing flesh on Eddie Heath ‘s shoulder and thigh .

↘ ESE

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
That/Those	475	45,63
This/These	278	26,70
The	130	12,49
∅	42	4,03
Pronombres personales y nombres propios	39	3,75
Posesivos	25	2,40
Reformulación	16	1,54
A/such/so	13	1,25
Expresiones temporales: (right) now/away; just; then; used to	13	1,25
There/here	6	0,58
Tonight	4	0,38
TOTAL	1041	100

Tabla 88 - Elementos activadores de la utilización del determinante *ese* en traducciones

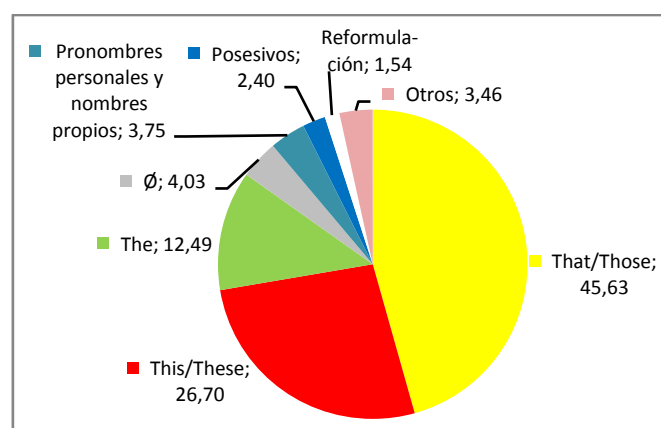


Gráfico 10 - Porcentajes de uso de los distintos elementos activadores del determinante *ese* en traducciones

El equivalente teóricamente más cercano de este demostrativo, *that*, si bien es el que presenta mayor frecuencia de aparición, no llega a aparecer ni en la mitad de las concordancias. Por otra parte, hay una representación importante del elemento contrario, *this*, como ya habíamos mencionado en la página 326 refiriéndonos a las funciones particulares de este demostrativo. En muchos de los casos *this* aparece en inglés en su forma determinante para aportar una nota de registro coloquial, como en el siguiente ejemplo, en el que resultaría extraño traducirlo por un *este*¹⁷⁵:

<Ellroy>: En este momento, el señor Hughes me hace seguir a <esa> actriz que le ha plantado .
 →par_en: Right now Mr. Hughes has got me tailing **this** actress who jilted him .

Mackenzie y Martínez Caro (2012: 189) comentan: «There is no specific equivalent to the Spanish informal (mostly pejorative) construction with the demonstrative following the definite article and noun, as in *el libro ese* ‘that book’» y sugieren, en la página 191, que una posible equivalencia de dicha estructura puede ser la combinación del demostrativo de lejanía con adjetivos valorativos, como en «that awful man Jones» por «el Jones ese». En nuestro corpus hemos comprobado, efectivamente, que el traductor pospone en ocasiones el demostrativo para transmitir esa connotación peyorativa, aunque, como ya hemos mencionado, en menor medida que en español original, probablemente por la falta de un estímulo más directo en inglés.

¹⁷⁵ La RAE (2009: 1294) recoge ciertos usos puertorriqueños y del habla juvenil peruana en los que se hace referencia a alguien que aparece por primera vez en una narración mediante grupos nominales definidos, como en «de repente este hombre se me acerca», que reformula como «se me acerca un hombre». Creemos que estos usos pueden provenir del calco de la estructura inglesa de la que hablábamos en la página 326.

<Grafton>: Puede que su pasado en las Fuerzas Aéreas , la historia <esa> de la Unidad de Voluntarios , fuera sólo una explicación honorable para justificar su ausencia .

→par_en: Maybe this business about the Air Force - **this** AVG **bullshit** - was just a polite explanation for his absence .

<Ellroy>: Esta noche compromete al comunista <ése> , la semana que viene hace escuchas clandestinas para Hush-Hush en una sauna de maricas.

→par_en: Tonight he compromises **this** **Commie** , next week he wires some fag bathhouse for Hush-Hush .

<Grafton>: ¿Cuántos años tiene la vieja <ésa> ?

→par_en: How old is **this** **old bag** ? "

El tercer elemento más representado es una vez más *the*, que, como hemos dicho, funciona como un comodín que permite al traductor seleccionar el determinante que mejor le parezca. De todos los casos que hemos hallado en los que *the* se traduce por un determinante demostrativo, el mayor número de traducciones corresponde a *ese* (128), aunque solo por un pequeño margen, pues *aquel* cuenta con una frecuencia de 126.

<Leonard>: ¿Cómo se llamaba <esa> mujer?

→par_en: " What's **the** woman's name?

Los casos de adiciones responden la mayoría de las veces a un intento de dotar a la frase de una connotación negativa o despectiva, o de una gramática más natural en español cuando el inglés coloquial omite el artículo o demostrativo a principio de una frase:

<Ellroy>: el astro de la producción ,Rock Rockwell , <ese> seductor blandengue

→par_en: with the star of the movie : limpwristed lothario Rock Rockwell.

<Leonard>: <Ese> tío podría salir por la tele y sería gracioso , pero tenía sus reglas .

→par_en: Man could be on TV , funny , but had his rules .

En este caso, la función temporal del demostrativo no es importante: los adverbios y expresiones de tiempo suman únicamente un 1,25 %. Sin embargo, es curioso observar que el adverbio *now*, que indica tiempo presente, se incluye a veces en frases que narran hechos pasados.

<Leonard>: En <ese> momento la ATF andaba muy ocupada con una operación
 →par_en: Right **now** the ATF office was busy working a sting operation

<Leonard>: Le dijeron que en <ese> momento podía hacer una llamada telefónica.
 →par_en: They said she could make her phone call **now**.

Este fenómeno es frecuente en inglés y su finalidad es estilística —por ejemplo, dotar a un relato de mayor vivacidad o intensidad (Mason y Şerban 2003: 285)— pero, por los usos gramaticales y estilísticos propios de nuestra lengua, es más difícil que se traslade al español, lo que obliga al traductor a cambiar el adverbio por expresiones con el demostrativo *ese*. Señalan Mason y Şerban (2003: 271) a este respecto:

In fiction, a writer may use proximals such as here or now either from his or her own perspective as author or from the perspective of one or another protagonist in the narrative. (...) The proximal now is oriented not to the narrator's deictic centre (as are the past-tense verbs [...]) but to the protagonist's experience at the time referred to in the narration.

Los posesivos también constituyen un elemento activador importante, puesto que muchas veces *ese* se utiliza para reemplazar, sobre todo, a los de segunda y tercera persona, que, como ya hemos visto, son más frecuentes en inglés que en español. Para evitar utilizarlos, el traductor recurre al demostrativo:

<Highsmith>: Y escoges un sombrero de paja que haga juego con <esa> chaqueta tan enrollada .
 →par_en: Pick up a hat to go with **your** snappy jacket.

◆ AQUEL

En el gráfico 11 podemos observar que las frecuencias se reparten de manera más equitativa entre los tres elementos principales, lo que indica que no hay un elemento principal, al menos en la misma medida que en los otros casos, que sea el activador directo e inmediato de este demostrativo.

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
That/Those	192	32,60
The	126	21,39
This/These	122	20,71
Posesivos (incl. genitivo sajón)	31	5,26
∅	28	4,75
Today/tonight	25	4,24
(by) now, so far	20	3,40
A/such	14	2,38
Just (then)/(back) then	12	2,04
Pronombre personal, indefinido o nombre propio	7	1,19
(over) there/here	6	1,02
Reformulación	4	0,68
Pronombre relativo	2	0,34
TOTAL	589	100

Tabla 89 - Elementos activadores de la utilización del determinante *aquel* en traducciones

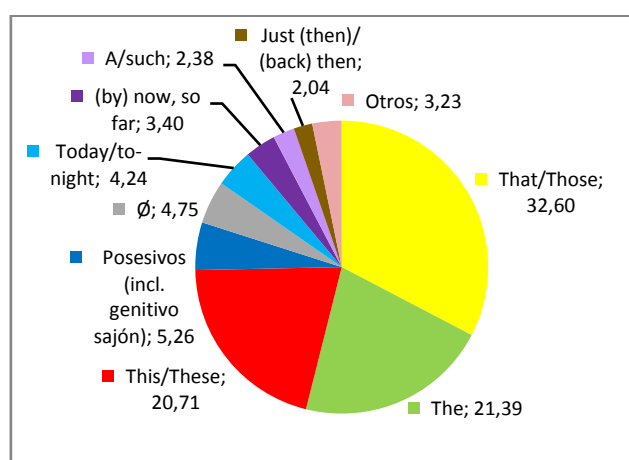


Gráfico 11 - Porcentajes de uso de los distintos elementos activadores del determinante *aquel* en traducciones

De los casos de *that* que se traducen por un determinante demostrativo, un número muy superior corresponden a *ese*. Sabemos, porque vemos el dato en nuestro cómputo de frecuencias (v. tabla 79), que *ese* es más frecuente que *aquel* en español original (un 23,5 % más), pero esta diferencia se incrementa a casi el triple en traducciones (76,7 % más). Ambos elementos son traducciones válidas de *that* y *those* pero, a diferencia de *aquel*, *ese* puede utilizarse en la mayor parte de los casos en los que se usaría *aquel*¹⁷⁶, cosa que no sucede a la inversa, como vemos en el ejemplo, en el que la utilización de *aquel* resultaría extraña, pues indicaría un grado de lejanía que no es apropiado cuando se señala un elemento perteneciente a la esfera del receptor:

<Ellroy>: Y con <esa> misma concisión de abogado , ¿qué dirías tú que hay detrás?

¹⁷⁶ Dice De Kock (1992: 80-82) al respecto de la lengua hablada, que, entendemos, es la que intentan imitar los diálogos de nuestras novelas: «En la lengua escrita los campos de referencia se reducen a dos, uno lejano y otro cercano, para los que se emplean respectivamente *aquel* y *este*, o incidentalmente *ese*. En la lengua hablada estos dos tienden incluso a confundirse, ya que *ese* sirve práctica e incluso preferentemente en cualquier circunstancia».

→par_en: And in **that** concise attorney 's manner of yours , what would you say is behind it?

También, por supuesto, hay que tener en cuenta las motivaciones de naturaleza afectiva o emotiva y la deixis evocadora:

<Leonard>: Ella alzó la cabeza y de nuevo pudo ver <aque!> brillo en sus ojos , aquella mirada que podía cambiar su vida si él lo permitía.

→par_en: She raised her head and he saw her eyes , **that** gleam , that look that could change his life if he let it .

<Cornwell>: Más adelante , la experiencia se había convertido en mi maestra , limpiándose los pies en <aquella> prístina parte de mí que era idealista y analítica .

→par_en: Then experience had become my mentor , wiping its feet on **that** pristine part of myself that was idealistic and analytical .

Sin embargo, una causa de fuerza mayor para seleccionar *aque!* de manera prioritaria frente a *ese* es la narración en tiempo pasado de hechos pasados o el discurso indirecto e indirecto libre. En estos casos, encontramos en inglés deícticos de proximidad temporal (*today, now*) o espacial (*here, this*) junto a verbos conjugados en pasado, pero el español, en cambio, no admite fácilmente esta combinación.

<Grafton>: No sabía adonde iba a llevarnos <aquella> situación

→par_en: I couldn't imagine where **this** situation was heading

<Highsmith>: ¿Por qué razón llevaba aquella blusa de algodón de manga larga , en <aque!> caluroso día de agosto?

→par_en: Was that the reason for the rather quaint long-sleeved cotton blouse on **this** hot August day?

<Highsmith>: Intuía que él se iba a ir - Tom lo sabía - , quizá <aque!> mismo día

→par_en: She sensed that he might be leaving , Tom knew , and **today** .

<Ellroy>: Ser seguido <aquella> noche significaría lamentarlo el resto de la vida

→par_en: A tail **tonight** meant grief forever.

<Grafton>: ¿Qué camino tomaría , el sublime o el mezquino ? ¿Sabía aún a <aquellas> alturas cuál era cada cuál?

→par_en: Was I going to take the high road or the low? **By now** , did I even know which was which?

<Leonard>: al tiempo que Faron afirmaba que no aparentaba cuarenta y cuatro años , al menos a <aquella> distancia .

→par_en: Faron Tyler saying she surely didn't look forty-four , at least not from **here** .

Es interesante observar que no hay una diferencia excesiva entre las frecuencias de *this* y los dos primeros elementos activadores. Esto puede resultar extraño, puesto que, en nuestra escala teórica, no hay nada más alejado del sentido déictico de *this* que el demostrativo *aquel*, pero no tanto, si tenemos en cuenta la explicación anterior.

Goethals (2013: 543) apunta que la razón para introducir este demostrativo cuando hay un artículo definido en el original puede responder más a una estrategia para enfatizar la distancia temporal del acto de habla que para enlazar el sintagma con el contexto anterior.

<Highsmith>: En el caso de que Pritchard volviera con la policía , diría que había tenido que coger <aquel> repugnante saco de huesos - había visto algunos huesos y se había dado cuenta de lo que eran - y quitarlos de la vista de su ama de llaves por una cuestión de decencia .

-->par_en: He would say , in case Pritchard turned up with the police , that of course he had had to put the revolting bag of bones-Tom had seen some bones and had certainly felt them-out of his housekeeper's vision for decency's sake

Lo cierto es que *the* ocupa la segunda posición en la tabla¹⁷⁷, lo que puede indicar que el traductor encuentra en este elemento neutro la oportunidad de «normalizar» el texto (es decir, de acercarse a la norma española que, como hemos visto, se vale en mayor medida que las traducciones de *aquel*) utilizando este demostrativo de lejanía, que quizás no se activa tanto en presencia de elementos de significado más concreto y equivalentes más inmediatos como *this* o *that*.

<Grafton>: ¿Nunca estaba en casa <aquel> hombre?

→par_en: Was the guy never home?

En este sentido, también puede resultar significativo el ligeramente mayor porcentaje de uso de *aquel* (4,75 % frente a 3,85 % en *este* y 4,03 % en *ese*) que proviene de fragmentos en los que no hemos encontrado un elemento activador, y en los que podría haberse utilizado cualquier otro recurso para la traducción, incluidos el demostrativo *ese* y el artículo determinado.

¹⁷⁷ Moropa (2011: 273) comenta que la utilización del demostrativo para traducir el artículo *the* inglés es un patrón común también en los idiomas africanos xhosa y zulú: «It is also sometimes used to translate the personal pronoun “it”. When the demonstrative is used, the sentence becomes more explicit as the demonstrative qualifies that particular noun». El autor asume que los traductores adoptan esta estrategia por la falta de artículos definidos e indefinidos en estas lenguas. No obstante, el español sí cuenta con dichos artículos y aun así utiliza los demostrativos, como hemos visto.

<Cornwell>: <Aquella> « Jenny », ¿sería Jennifer Deighton?

→par_en: Was “ Jenny “ Jennifer Deighton?

<Highsmith>: Quizá iría andando con Noëlle hacia el Hotel Miramare de Casablanca y las dos llevarían bolsas de plástico y cestos recién comprados , llenos de sus compras de <aquella> tarde .

→par_en: Strolling back to the Hotel Miramare , Casablanca , with Noëlle , both carrying plastic bags and newly acquired baskets full of afternoon purchases?

<Highsmith>: Tom no iba casi nunca a la comida semanal en casa de los Plissot , pues <aquellas> comidas le aburrían .

→par_en: Tom almost never went to the once-a-week lunch at the Plissons , as lunches bored him

Comprobamos que este determinante también se utiliza para traducir expresiones de significado temporal, quizás en mayor medida que *ese*, puesto que las tres categorías de este tipo sumadas abarcan casi un 10 % de las concordancias.

<Highsmith>: Pensó que , hasta <aquel> momento , la tumba acuática de Murchison había sido un buen escondite

-->par_en: Until **now** , he told himself , that watery grave of Murchison had been a most successful hiding place

<Cornwell>: Ojalá hubiera avisado a alguien en <aquel> momento . Quizá todavía estaba... . Marino la interrumpió .

-->par_en: " Lord , Lord . I wish I 'd called somebody **then** . Maybe she was - " Marino cut in

Finalmente, podemos mencionar la presencia de los posesivos en una posición relativamente elevada en la tabla. En estos ejemplos vemos que se utiliza el demostrativo para evitar un posesivo que en español resultaría poco natural o superfluo.

<Leonard>: <Aquel> tranquilo tono de voz , mirándole a los ojos

→par_en: In **her** quiet tone , looking right at him...

<Highsmith>: A Tom le hubiera encantado pegarle un puñetazo en <aquel> prominente plexo solar .

→par_en: Tom would have loved to sock him in **his** bulging solar plexus .

Cuenca y Ribera (2011: 377) observan que en su corpus de traducciones al catalán la presencia de *such* en inglés implica sistemáticamente la introducción de un demostrativo en la traducción. En nuestro corpus no sucede lo mismo, pero sí hemos observado en las tres tablas correspondientes a cada uno de los determinantes

demostrativos y en la del pronombre *ese* la presencia, nunca demasiado abundante, de *such* como elemento equivalente.

<Highsmith>: ¿Tenían quizá dos de <aquellos> artefactos tipo rastrillo ?

-->par_en: And maybe they had two such rake like devices with grappling hooks ?

5.3.3.2. Pronombres

➔ ESTE

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
This/These	298	56,65
Pronombres personales, sustantivos y nombres propios	87	16,54
That	43	8,18
∅	28	5,32
Reformulación	18	3,42
The + sustantivo	17	3,23
Here	15	2,85
Pronombres relativos	7	1,33
All, much, many, one	5	0,95
Now, then	3	0,57
The latter	2	0,38
In	2	0,38
Nearby	1	0,19
TOTAL	526	100

Tabla 90 - Elementos activadores de la utilización del pronombre *este* en traducciones

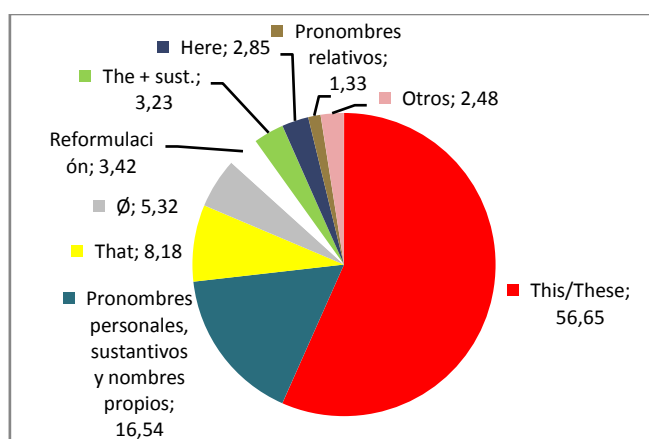


Gráfico 12 - Porcentajes de uso de los distintos elementos activadores del pronombre *este* en traducciones

Cuando *este* funciona como pronombre, las expresiones que hacen referencia al tiempo ya no son tan comunes como en el caso del determinante. Tras el equivalente más evidente, *this/these*, que hallamos sobre todo entre las concordancias de *esto*, lo que encontramos en las formas flexionadas —como es lógico, dado que hablamos de la traducción de pronombres—, es una abundancia de pronombres personales, de sustantivos comunes y de nombres propios. *It* es el pronombre que más frecuentemente aparece (28 veces), sobre todo en las concordancias de *esto*, seguido de 20 nombres propios, 16 *he* y 12 *she*.

. <Cornwell>: - <Esto> es grave – observé . – Podría serlo .

→par_en: “ **This** is serious , “ I said .

<Ellroy>: - Te lo diré si... , quiero decir , cuando <esto> termine .

→par_en: “ I ‘ll tell you if-I mean when **it** ‘s over .

<Grafton>: Aprieta el gatillo y <ésta> morirá . Le retorceré el pescuezo – dijo Gilbert .

→par_en: “ Pull the trigger and **she** ‘s dead . I ‘ll snap her neck , " Gilbert said .

La explicitación necesaria del sujeto en inglés, con pronombres personales como *he*, que en español no se traducen, y la mayor repetición de nombres propios que se da en esta lengua, cuya reproducción en español podría resultar estilísticamente extraña, son factores que pueden influir en la mayor utilización del pronombre *este*. Así, como vemos en los ejemplos, el pronombre se utiliza para hacer referencia al antecedente sin necesidad de repetirlo.

<Leonard>: <Éste> va en serio . Míralo – comentó Louis .

→par_en: Louis said , “ **He**’s serious. Look at him.

<Leonard>: Louis presionó el gatillo y vio que el punto rojo del láser se fijaba en la frente de Gerald . <Éste> sonrió.

→par_en: Louis squeezed the grip and saw the red laser dot appear on Gerald's forehead . **Gerald** grinned at him.

Una vez más, observamos que el tercer elemento activador en la lista es *that*, lo que indica un desajuste entre las estructuras déicticas del inglés y el español, o cierta libertad del traductor a la hora de reproducirlas.

<Highsmith>: <Este> es muy popular . – Nick volvió a sonreír .

→par_en: **That**'s a popular one .

<Grafton>: Tal como yo lo veo , han acabado por descubrirlo y <éste> es el resultado .

→par_en: The way I see it , the boys finally got a lock on him , and **that** 's the result . “

Un elemento que se repite en la categoría «reformulaciones» son las frases hechas, sobre todo en las concordancias del neutro *esto* que, como hemos dicho, puede hacer referencia a conceptos abstractos o fragmentos enteros de texto:

<Cornwell>: Todo <esto> debe de resultarle emocionante .

→par_en: He skates on the edge for kicks .

<Ellroy>: Enteraos , gatitos y gatitas , chicos y chicas : a nosotros nos gusta el pasatiempo nacional tanto como a cualquiera , pero <esto> ya es pasarse .

→par_en: Dig it , kats and kittens , chicks and charlies : we love the national pastime as much as you do , but enough is enough .

<Ellroy>: ¿Qué pinta en todo <esto>?

→par_en: What 's the story on him?

<Cornwell>: Gobernador Norring , <ésta> es una pregunta capciosa .

→par_en: “ Governor Norring , you are baiting me . “

También hallamos dos ejemplos de deixis espacial pura, en los que el narrador del texto en inglés describe el gesto o movimiento de señalamiento que realiza el personaje y, en español, se elige explicitar dicha referencia al elemento señalado mediante el pronombre:

<Grafton>: ¿De qué hablas? – De <esto> – Alargué la mano y golpeé el hinchado vientre de Laura .

→par_en: “ Yes , you do . “ I reached over to Laura 's belly and knocked on the mound .

<Highsmith>: Pero insisto en que te tomes <esto> . Es un buen apoyo moral .

→par_en: “ I insist . Moral support .

Por otro lado, un elemento que se repite entre las apariciones del demostrativo sin elemento activador directo en inglés es el desdoblamiento del nombre propio de un personaje incluido en un complemento circunstancial u objeto directo, que en español se conserva como objeto directo pero, además, para evitar confusiones, se añade una

referencia a dicho personaje mediante un pronombre demostrativo en el complemento circunstancial que le sigue:

<Leonard>: Se quedó mirando a Max , mientras <éste> escogía la Browning 380 automática, la sacaba de la pinza y revisaba el cargador .

→par_en: Watching as Max chose the Browning 380 auto , took it from its peg and checked the slide .

<Leonard>: ...miró a Jackie mientras <ésta> contemplaba a su mujer.

→par_en: He watched Jackie looking at his wife

<Leonard>: Miró a Jackie mientras <ésta> daba una calada al cigarrillo y soltaba una bocanada .

→par_en: He watched Jackie draw on her cigarette and let out a slow stream of smoke .

<Grafton>: Seguí a Bucky mientras <éste> salía por la puerta trasera y bajaba los peldaños del porch

→par_en: I followed Bucky out the back door and down the porch steps .

<Leonard>: Nicolet se acercó a Tyler mientras <éste> sacaba varios fajos de billetes de cien dólares

→par_en: Nicolet stepped closer as Tyler pulled out several packets of one-hundred-dollar bills

<Leonard>: Max se quedó mirando a Winston mientras <éste> se quitaba el pañuelo para ver la herida .

→par_en: Max watched Winston unwrap the handkerchief to look at his wound .

Los siguientes son ejemplos de otros elementos hallados entre las concordancias paralelas de *este*:

<Grafton>: Recogí la llave de la de la bandeja y dejé <ésta> donde estaba .

→par_en: I plucked her key from the room service tray and left **the** tray where it was .

<Grafton>: La tercera rueda mueve la segunda y <ésta> hace girar la primera – dijo

→par_en: The third wheel moves the second wheel , **which** then turns the first wheel , “ he said .

<Grafton>: Ray me condujo hacia <éste> y él se instaló en la silla de madera .

→par_en: Ray steered me toward **the latter** and took the wood chair for himself

<Cornwell>: Personalmente , creo que se trataba de <esto> último .

→par_en: Personally , I think it was **the latter** . “

➔ ESE

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
That/those	541	57,92
Pronombres personales y sustantivos (93 it)	106	11,35
∅	86	9,21
This/these	50	5,35
Reformulación	39	4,18
Relativos: What/which	30	3,21
Sintagmas con so (adv) /such/as much	23	2,46
Conectores: So (conj)/hence/therefore/however/though/then/instead (trad.: «por eso», «a pesar de eso», y «eso que»)	21	2,25
Yeah, right, sure, exactly (trad.: «eso [es]»)	12	1,29
The/one + sustantivo	8	0,86
Meaning (trad.: «¿Y eso?»)«Eso significa»)	6	0,64
There, here	5	0,53
Around, by (trad.:«a eso de»)	4	0,43
Nope, no way, not for a minute (trad.: «de eso nada»)	3	0,32
TOTAL	934	100

Tabla 91 - Elementos activadores de la utilización del pronombre *ese* en traducciones

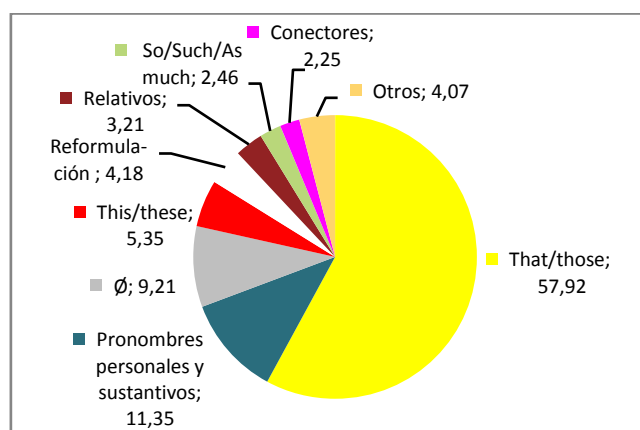


Gráfico 13 - Porcentajes de uso de los distintos elementos activadores del pronombre *ese* en traducciones

En la mayor parte de los casos en los que *that* se traduce por un pronombre demostrativo, *ese* es el elegido, como sucedía también cuando hablábamos de los determinantes. Sin embargo, aunque la diferencia entre la utilización de las formas pronominales de *ese* y *aquel* en ambos subcorpus es mucho más abrumadora que en el caso de los determinantes, la diferencia que encontramos si comparamos dicho uso en español original y traducido no es tan grande como en aquel caso, ya que en ambos subcorpus hay aproximadamente la misma cantidad de apariciones más de *ese* que de *aquel*: aproximadamente siete veces más. Por otra parte, el porcentaje del elemento teóricamente opuesto, *this/these* ya no es tan alto.

It también es un elemento activador frecuente, sobre todo porque suele traducirse por el demostrativo con función anafórica *eso*. Lo incluimos dentro de la

categoría de los pronombres y sustantivos, pero la realidad es que de esos 106 casos, 93 corresponden a *it*.

<Leonard>: - <Eso> va contra la ley .

→par_en: “It’s against the law .

<Ellroy>: - <Eso> dije , y lo mantengo .

→par_en: “ I said it and I meant it .

Aquí encontramos que la tercera categoría en orden de frecuencias es la correspondiente a una falta de elemento activador en el original, y la quinta, la de las reformulaciones. ¿Por qué sucede esto? Como hemos explicado en la descripción de los elementos de análisis, muchas veces se utiliza este demostrativo en expresiones semilexicalizadas como «eso + de» o «¿Qué + SUST + verbo SER + *ese* + de...?», que no encontramos en inglés pero pueden servir para añadir una pincelada de informalidad al registro, hacer más natural la expresión, como en el primer ejemplo, o enfatizar una pregunta, como en el segundo.

<Ellroy>: <Eso> de los pantalones fritos parece cosa de J.C.

→par_en: Fried pedal pushers sounds just like J.C.

<Grafton>: ¿Qué historia es <ésa> de que han entrado a robar?

→par_en: “ What ‘s the story on the break-in?

De hecho, muchas de las otras categorías son muestra de esta capacidad del demostrativo neutro para formar fórmulas conversacionales o expresiones idiomáticas.

<Grafton>: A <eso> de las diez volví a casa de Bucky .

→par_en: By ten A.M. I found myself back at Bucky ‘s.

<Ellroy>: – Sí , <eso> es . Ése era el tipo que andaba siempre con Richie Herrick

→par_en: “ Yeah.. . riiight.. . he was that guy following Richie Herrick around .

<Ellroy>: Sencillamente , me pareció lo más lógico . - ¿Y <eso>?

→par_en: “ Meaning? “

<Leonard>: A lo mejor está limpio . – De <eso> nada . Está metido en armas y a lo grande .

→par_en: “ Maybe he’s been clean . “ “ No way – he’s into guns big-time .

Con respecto a las reformulaciones, encontramos que, además de traducir las habituales frases hechas (p. ej.: *Now you talking* o *No sooner said than done*) el pronombre demostrativo neutro hace referencia a conceptos o proposiciones enteras, que en inglés pueden expresarse a través de un verbo auxiliar o de enlace sin objeto. Esta función, conocida como *coding* (Mackenzie y Martínez Caro 2012: 55, 63) porque el auxiliar codifica una secuencia entera para evitar repeticiones innecesarias, es un mecanismo cohesivo de tipo elíptico sin correspondencia formal en español, lo que explica el uso del demostrativo (aunque también podría haberse optado por otras estructuras, como «*lo* + verbo»):

<Highsmith>: Tom y Heloise le tenían afecto porque ella se lo tenía a ellos o , al menos, <eso> parecía.

→par_en: Tom and Heloise loved her , because she loved them , or seemed to.

<Cornwell>: - ¿El motor estaba parado , las puertas sin seguro ? - <Eso> es .

→par_en: The engine was off , doors unlocked ? " They were .

<Ellroy>: - Pregunte . - <Eso> haré .

→par_en: " Ask them . " " I will .

<Grafton>: ¿Lo ha hecho Gilbert? - <Eso> creo – dije .

→par_en: " Gilbert did that ? " " One assumes , " I said .

También puede servir para explicitar conexiones, como en los siguientes ejemplos:

<Grafton>: Laura ya está traicionando a Gilbert y <eso> que en teoría está enamorada de él .

→par_en: She 's already busy double-crossing Gilbert , and she 's supposedly in love with him.

<Leonard>: Tuve que pensar un poco , por <eso> no lo llamé de inmediato

→par_en: I had to do some thinking , reason I didn't call you right away.

De hecho, la octava categoría de nuestra tabla está formada por conectores de todo tipo:

<Cornwell>: Por <eso> el sábado por la tarde Marino y yo nos pusimos en camino hacia la Penitenciaría del Estado .

-->par_en: **So** it was on Saturday afternoon that Marino and I were en route to the state penitentiary .

<Highsmith>: A Tom se le ocurrió que , a pesar de <eso> , a Georges y Marie no les gustaba la invasión de norteafricanos emigrados a su país .

-->par_en: It crossed Tom's mind that Georges and Marie didn't like the influx of people from North Africa , however .

Habíamos visto que la traducción del texto de Ellroy es de las que más utiliza el demostrativo *ese* (v. tabla 68). En este caso, Sabaté Vargas no es quien más utiliza la forma pronominal, pero observamos que sí es el segundo que más lo añade sin estímulo directo en el original, sobre todo para compensar la falta de conectores en los diálogos, que son cortos, rápidos y presentan abundantes omisiones de elementos (en su caso, las 24 adiciones conforman el 14,81 % de las concordancias del pronombre). En la traducción de Highsmith apreciamos un porcentaje superior de adiciones (22 adiciones, un 18,02 %), pero estas suelen ser de carácter opcional, mientras que las de Ellroy resultan a veces casi necesarias para la construcción de una frase gramatical en español.

Highsmith:

<Highsmith>: Últimamente recuerdas bastante los precios , Nick - le dijo Ed en tono jovial - . <Eso> está muy bien .

→par_en: " You're remembering prices these days , Nick ! " said Ed in a genial manner . " Very good .

<Highsmith>: Las vagas sospechas de Antoine respecto a Tom podían llegar a cuajar y <eso> era lo último que deseaba .

→par_en: Antoine's vague suspicion of Tom would congeal a bit , the last thing Tom wanted .

<Highsmith>: También se dio cuenta de que <eso> significaba arriesgarse , exponerse al peligro .

→par_en: He also realized that he would be sticking his neck out , endangering himself

Ellroy:

<Ellroy>: – Explique <eso> de « presionar » , teniente .

→par_en : « Explain 'lean on , ' Lieutenant . " "

<Ellroy>: Van por ahí intentando resolver muertes de morenos a manos de otros morenos .Se dice que con <eso> intentan contrarrestar toda esa posible publicidad favorable de la investigación federal .

→par_en: Breathless : " They goin ' aroun ' trying ' to solve colored-on-colored killins , word is to forestall all this potential good Federal investigation publicity .

<Ellroy>: Me niego a comentar <eso> . Pero siga preguntando por los Kafesjian .

→par_en : « Refuse to comment . But keep asking me about the Kafesjians . " "

➤ AQUEL

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
That/those	48	39,34
This	24	19,67
Pronombre personal, sustantivo o nombre propio (16 it)	20	16,39
∅	16	13,12
The + sustantivo	7	5,74
Reformulación	4	3,28
A + sustantivo	2	1,64
Which	1	0,82
TOTAL	122	100

Tabla 92 - Elementos activadores de la utilización del pronombre *aquel* en traducciones

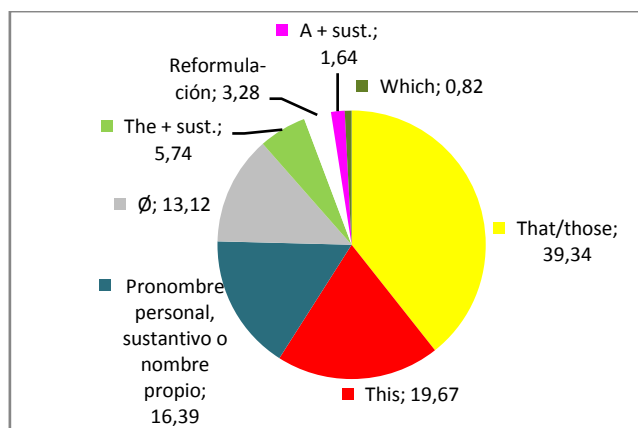


Gráfico 14 - Porcentajes de uso de los distintos elementos activadores del pronombre *aquel* en traducciones

La forma neutra del pronombre *aquel* es la más frecuente con diferencia, en proporción algo superior a la de los originales españoles, lo que puede ser consecuencia de la alta frecuencia que presentan sus equivalentes más evidentes en inglés, *that*, e *it*:

<Cornwell>: Había perdido <aquello> que lo hacía humano , que le hacía pensar y sentir , y no lo recobraría nunca .

→par_en: **That** which made him human , made him think and feel , was gone and was never coming back .

<Grafton>: No hablaba mucho de <aquello>...

→par_en: He didn't talk much about **it**...

En segundo lugar, volvemos a observar un porcentaje alto (mayor que en el caso de *ese* pronominal pero parecido al de *aquel* determinante) del elemento opuesto en la escala deíctica, *this*, que ocupa el segundo puesto de la tabla, lo que confirma las

diferencias existentes en la utilización del sistema de demostrativos en inglés y español que hemos mencionado anteriormente, y la libertad con la que se utiliza al traducir.

<Grafton>: <Aquello> empezaba a parecerse a esas discusiones con los fanáticos de la ufología

→par_en: **This** was beginning to feel like one of those discussions with a UFO fanatic

<Grafton>: Aún no podía creer que estuviera haciendo <aquello> .

→par_en: I still could n't believe I was doing **this** .

<Leonard>: ¿Era <aquella>?

→par_en: Was **this** the one ?

A este respecto, cabe mencionar el trabajo de Mason y Şerban (2003), ya citado anteriormente, sobre deícticos en la traducción del rumano al inglés. El resultado de dicho trabajo indicaba que los cambios de deícticos de cercanía a deícticos de lejanía en traducciones eran más numerosos que a la inversa. Sin embargo, el estudio de Goethals (2007: 99) contradice —en palabras del autor— los resultados de Mason y Şerban, puesto que sus textos meta contienen más deícticos de cercanía que los textos originales, por lo que es muy difícil llegar a generalizar en estas cuestiones. En nuestro caso, los porcentajes que vemos en las tablas indican también que hay más cambios de demostrativos de lejanía a cercanía que viceversa.

Por otro lado, el porcentaje de adiciones de *aquel* es mayor que en el caso de las otras dos formas pronominales (13,12 % frente a 9,21 % en *ese* y 5,32 % en *aquel*), lo que nos hace pensar que el uso de este pronombre es el de naturaleza más facultativa. Aunque lo cierto es que casi siempre las encontramos en fragmentos narrados en pasado:

<Highsmith>: Pero seguían juntos . <Aquello> era enfermizo .

→par_en: Yet they clung together . Sickening .

<Leonard>: O podía usar la que le había dejado Cujo.. . Y <aquello> fue lo que decidió .

→par_en: Or he could use the one Cujo left him – and decided , yeah , he would .

5.3.4. Conclusiones parciales: determinantes y pronombres demostrativos

Si tomamos el conjunto de los demostrativos como un único paradigma, vemos, en primer lugar, que la forma predominante es *ese*; en segundo lugar, que, en general, hay más determinantes que pronombres en nuestro corpus; y, por último, que el demostrativo de lejanía más frecuente es *ese*. Estos dos últimos datos coinciden con las conclusiones del estudio de Macías Villalobos, quien afirma (2006: 196 y 378) que *aquel* sería «una unidad casi reducida a los ámbitos de la lengua literaria, actuaría como variante funcional de *ese*, especializada en la expresión del máximo alejamiento» y que «en los textos que se acercan más al registro de la lengua popular, su presencia es casi testimonial». Señala este autor, sin embargo, que el demostrativo que más frecuencia presenta en su corpus es *este*, lo que no coincide con nuestros resultados.

El panorama cambia un poco cuando dividimos los demostrativos según paradigmas. Los determinantes más utilizados en ambos subcorpus son, en este orden, *ese*, *este* y *aquel* (v. tabla 79). A pesar de que, en cuanto a porcentajes de distribución de formas, *ese* está mejor representado en TRAD que en ES, las frecuencias normalizadas nos indican que se utiliza en menor medida en TRAD (aunque la diferencia no es estadísticamente significativa). En cuanto a *este* y *aquel*, ambos se utilizan menos en TRAD en términos porcentuales y absolutos, pero la utilización de *aquel* disminuye más y podría estar relacionada con el incremento del porcentaje de *ese*.

En el caso de los pronombres, el patrón que hallábamos en los determinantes se repite en ambos subcorpus (mayor porcentaje de *ese*, seguido de *este* y *aquel*), pero aquí la diferencia en la utilización de cada uno es mucho más acusada (v. tabla 80). En cuanto a porcentajes de distribución de formas, *este* aumenta su representación en TRAD, en detrimento de *ese* y *aquel*, y, de hecho, su frecuencia normalizada de aparición en TRAD es mayor que en ES, aunque la diferencia no llegue a ser significativa en términos estadísticos. En cuanto a *ese* y *aquel*, ambos se utilizan menos en TRAD en términos porcentuales y de frecuencia.

En la tabla 93 podemos ver esquematizados los resultados de la comparación entre subcorpus.

Parámetros de comparación	Formas determinantes	Formas pronominales	Ambos paradigmas
Formas infrarrepresentadas en TRAD	ESTE		ESTE
		ESE	ESE
	AQUEL	AQUEL	AQUEL
Orden de formas por frecuencia de uso	1) ESE	1) ESE	1) ESE
	2) ESTE	2) ESTE	2) ESTE
	3) AQUEL	3) AQUEL	3) AQUEL
Incremento o reducción de porcentaje de uso en TRAD ¹⁷⁸	Este -	Este +	Este +
	Ese +	Ese - -	Ese + +
	Aquel - -	Aquel -	Aquel -

Tabla 93 - Comparación de utilización de demostrativos en ES y TRAD

El análisis de concordancias paralelas, por su parte, nos indica que hay más demostrativos en inglés original (*this/these* y *that/those*) que en español original (*este, ese, aquel* y formas flexionadas), pero también que existe una diferencia mucho mayor, según el test de significación estadística, entre los textos en inglés original y sus respectivas traducciones.

Esta diferencia es más acusada en el caso de los demostrativos de cercanía, en los que EN dobla la cantidad de TRAD. En cuanto a los demostrativos de lejanía, hemos comprobado que menos de la mitad de todos los casos de *that* y *those* se trasladan a español en la forma de sus demostrativos equivalentes. El resto se omite o se expresa a través de los diferentes recursos gramaticales y estilísticos disponibles en nuestra lengua. El demostrativo, entonces, quizás se reserve para casos en los que la marca déctica o anafórica que aporta tiene relevancia en el contexto y no se percibe como excesiva.

Si analizamos las concordancias paralelas para hallar los elementos equivalentes a los demostrativos que han servido —o no— de detonadores para su utilización en las respectivas traducciones, observamos las siguientes tendencias para cada una de las formas estudiadas:

- La mayor parte de las veces que aparece el demostrativo *este* (63,05 % en determinantes y 56,65 % en pronombres), lo hace como consecuencia de la utilización en inglés de su equivalente más aproximado *this*. En el —aproximadamente— 40 % restante de las concordancias, observamos en los textos originales una abundante utilización de adverbios y expresiones temporales en la forma determinante y de pronombres personales, de sustantivos

¹⁷⁸ Dos signos positivos o negativos indican mayor aumento o descenso del porcentaje de uso que las otras formas con el mismo signo.

comunes y de nombres propios —ya hemos explicado que la norma estilística en español difiere de la del inglés en este aspecto— en la forma pronominal. Un 6,18 % de los determinantes y un 8,18 % de los pronombres demostrativos traduce al elemento deíctico teóricamente opuesto en significado (*that*). Como explican Mason y Şerban (2003: 271), «(...) selection of the marked member of the set allows the producer to communicate extra meanings», es decir, valores añadidos como la identificación con el referente o el receptor, o el distanciamiento de uno u otro. Otro elemento del que deriva la utilización del demostrativo de cercanía español (en un 5,49 % de los casos de determinantes y un 3,23 % de los de pronombres) es el determinante *the*. En estos casos, el traductor puede querer aportar un matiz más subjetivo o enfático, o simplemente buscar una mayor explicitación del referente. Algunas de las adiciones —que conforman un 3,85 % de las concordancias de determinantes y un 5,32 % de las de pronombres—, también sirven para cohesionar el texto y hacerlo más fluido mediante la explicitación de indicaciones temporales o conectores. Este fenómeno parece darse más en los textos que se acercan al lenguaje oral, como el de Ellroy.

- El equivalente más cercano de *ese, that*, presenta la frecuencia de aparición más elevada en la tabla de equivalentes de este demostrativo, que llega casi al 58 % en el caso de los pronombres pero apenas pasa del 45 % en el de los determinantes. En ambos casos, *that* se traduce más por *ese* que por *aquel*, aunque en los determinantes hay más diferencia entre español original y traducido en cuanto al porcentaje de utilización de las dos formas. Por otra parte, el porcentaje del elemento teóricamente opuesto —*this/these*—, es mucho más alto en determinantes (26,70 %) que en pronombres (5,35 %). Los pronombres personales y sustantivos son el segundo elemento activador en pronombres (11,35 %). No así en determinantes (conforman solo el 3,75 %), puesto que la forma más frecuente de esta categoría es *it*, que suele traducirse por el demostrativo neutro con función anafórica *eso*. El tercer elemento más representado en determinantes es *the* (12,49 %), que, como hemos dicho, funciona como un comodín para el traductor. Sin embargo, en el caso de los pronombres, el porcentaje de aparición de este artículo es anecdótico (0,86 %) y la tercera categoría en orden de frecuencias (9,21 %) es la correspondiente a la falta de elemento activador en el original. Las reformulaciones ocupan el quinto

puesto, y en estas dos últimas categorías hallamos con mucha frecuencia el demostrativo neutro en expresiones idiomáticas, fórmulas conversacionales o frases hechas, para referirse a conceptos o proposiciones enteras o para explicitar conexiones. Los casos de adiciones en determinantes (4,03 %) buscan sobre todo conseguir un tono negativo o despectivo, o naturalizar la gramática. La función temporal de *ese* no es tan relevante: los adverbios y expresiones de tiempo suman únicamente un 1,63 % en los determinantes y un 0,43% en los pronombres.

- Con respecto a *aquel*, vemos que no existe demasiada diferencia, en su forma determinante, entre las frecuencias de *this* (20,71 %) y los dos primeros elementos activadores (*that*: 32,60 %; *the*: 21,39 %). En pronombres también observamos un porcentaje alto del elemento opuesto en la escala déctica, *this*, que ocupa el segundo puesto de la tabla. *The*, por su parte, ocupa la segunda posición en la tabla de determinantes y la quinta en la de pronombres. Asimismo, observamos que el porcentaje de adiciones es mayor que en las otras dos formas, lo que indica que el uso de *aquel* es el de naturaleza más facultativa, sobre todo en el caso del pronombre. *Aquel* presenta una frecuencia menor que *ese* tanto en español original (menos de la mitad, un 54 % menos) como en traducciones (casi un tercio, 64 % menos): el predominio de *ese* en la lengua hablada —los diálogos de nuestras novelas— podría atribuirse a que es la forma con menos precisión distintiva, según de Kock (1992: 82). *Aquel* se selecciona de manera prioritaria frente a *ese* en contextos muy concretos, como la narración en tiempo pasado de hechos pasados, el discurso indirecto o indirecto libre, ciertas expresiones de significado temporal y fragmentos en los que el componente estilístico es importante.

Las principales tendencias observadas en nuestro corpus paralelo son, por tanto, en este orden:

- la traducción del demostrativo inglés por el demostrativo español de significado equivalente (aunque el porcentaje de esta opción disminuye bastante en los casos estudiados de *aquel*, en los que se favorecen otras opciones en mayor medida que en los demás),

- la traducción del demostrativo inglés por otra forma demostrativa española, incluso del grado de cercanía opuesto (el porcentaje es mínimo en el caso de *este*),
- la sustitución de posesivos o pronombres personales ingleses (especialmente por pronombres demostrativos),
- la sustitución del artículo definido inglés (especialmente por los determinantes *ese* y *aquel*)
- y la adición del demostrativo (que se da en mayor medida a través de los pronombres *ese* y *aquel* pero no supera el 8 % en promedio en el segundo y tercer grado y el 5 % en el primer grado de cercanía).

La omisión del demostrativo solo puede estudiarse a partir de la búsqueda de los demostrativos ingleses y, lo que hemos visto es que, en el caso de *that*, solo se traduce por los demostrativos equivalentes *ese* y *aquel* en un 34,37 % y un 8,53 % de concordancias respectivamente; el 17,24 % se omite y el resto se traduce mediante otros elementos (v. gráfico 7). Con respecto a *this* (v. gráfico 8), el porcentaje de traducción por el elemento equivalente *este* es del 34,62 %, el segundo puesto lo ocupa *ese* (15,08 %) y la omisión es menor (11,68 %).

Es evidente que la influencia del texto original no es el único factor que afecta al número de demostrativos utilizados, pues, de lo contrario, encontraríamos más demostrativos en traducciones que en originales españoles. Por el contrario, hemos visto una disminución en la frecuencia de formas demostrativas en traducciones con respecto a sus respectivos originales. El uso del demostrativo parece depender de otros factores más relacionados con las características estilísticas y gramaticales de nuestra lengua que los traductores, consciente o inconscientemente, parecen tener en cuenta. Sin embargo, resulta sorprendente que no solo no se siga la tendencia de la lengua original, sino que, además, los demostrativos se utilicen menos, en conjunto, que en español original.

Este fenómeno podría tener que ver con una tendencia a la normalización del español traducido¹⁷⁹, de la que podría ser muestra también, como observábamos en la página 301, el uso de una menor variedad de estructuras propias de nuestra lengua.

¹⁷⁹ Hay que tener en cuenta, no obstante, que lo esperable en una dinámica de normalización sería que las cifras de TRAD estuvieran a medio camino entre las de EN y las de ES, lo cual no sucede aquí, dado que existe una infrarrepresentación de los elementos estudiados en TRAD; es decir, se supera dicha norma.

Recordemos que, al separar los determinantes de los pronombres, detectamos una menor variación de estructuras en TRAD —reflejada, en nuestro corpus, en un escaso uso de demostrativos pospuestos al sustantivo y antepuestos a oraciones de relativo—, que, al no darse en inglés, no se activan a la hora de traducir, por lo que el sentido termina siendo trasladado mediante otras estructuras más similares a las del texto original o más frecuentes en español. Si contáramos con un componente de direccionalidad inversa; es decir, las traducciones al inglés de originales españoles, podríamos realizar una búsqueda de estructuras de estas características y comprobar si existe una manera de trasladarlas que incluya el demostrativo o si el traductor compensa de alguna manera la carencia déctica en el contexto. En este punto, podría aventurarse la hipótesis —que las limitaciones de este trabajo no nos permiten comprobar— de que los elementos únicos puedan no ser los demostrativos o uno de ellos en concreto, sino ciertas de sus funciones o de las estructuras de las que estos forman parte. De todas formas, este es solo uno de varios factores que podrían influir en las diferencias entre subcorpus; las dos estructuras mencionadas por sí solas no son suficientes para explicarlas. Además, la escasez de estas estructuras en particular podría originarse en diversos factores. Por ejemplo, la mayor oralidad de los textos estadounidenses podría ser la causante de la menor presencia de estructuras de relativo, como sugiere Macías Villalobos (2006: 203): «En el resto de obras, las que consideramos más cercanas a los registros de la lengua hablada, el número de oraciones de relativo es infinitamente menor y, curiosamente, es también donde el demostrativo de tercera está prácticamente ausente».

En este sentido, es interesante señalar que, además de que la infrarrepresentación de *aquel* en traducciones es bastante mayor, proporcionalmente, que la de *ese* (las frecuencias normalizadas indican que hay un 16 % menos de casos de *ese* en TRAD, frente a un 34 % menos de *aquel*), la proporción en la que se utilizan ambas formas demostrativas del grado de deixis de lejanía varía también entre ambos subcorpus de español. Así, hallamos una disminución del 4,02 % del porcentaje de *aquel* en TRAD junto con un aumento de *ese* de 1,82 % (v. gráfico 6).

El uso de *aquel* parece tener que ver —en mayor medida que en el caso de los otros dos niveles décticos—, con las normas o usos de gramaticalidad y estilo, y con los valores décticos o estilísticos particulares que este puede aportar, como mencionábamos un poco más atrás. Se aprecia, por ejemplo, una tendencia mayor en la forma determinante de *aquel* que en la de *ese* a colocar con determinadas palabras

temporales, como podemos observar en las tablas siguientes, en las que —teniendo en cuenta que *ese* supera siempre a *aquel* en frecuencia de aparición—, *aquel* supera a *ese* en 5 de 7 casos en ES (*mañana, tarde, noche, vez* y *día*). Dicha tendencia es aún más acusada en TRAD, en el que *aquel* añade dos casos en los que supera a *ese* (*momento* y *hora*) y con un margen mayor en *tarde* y *día*, aunque *vez* aparece en mayor número de ocasiones con *ese* en este caso.

ES	Mañana	Tarde	Noche	Momento	Vez	Día	Hora	Total
Este	13	11	71	20	51	4	15	185
Ese	0	2	11	42	0	6	11	72
Aquel	10	4	31	35	6	19	10	115
Total	23	17	113	97	57	29	36	372

Tabla 94 - Utilización de demostrativos en expresiones temporales en ES

TRAD	Mañana	Tarde	Noche	Momento	Vez	Día	Hora	Total
Este	54	25	67	42	63	5	6	262
Ese	3	4	15	15	5	5	2	34
Aquel	13	10	39	55	1	28	10	156
Total	70	39	121	112	69	38	18	452

Tabla 95 - Utilización de demostrativos en expresiones temporales en TRAD

Hallamos que existen otros detonadores de su uso, que pueden ser léxicos, como *over* o *back then*.

<Cornwell>: Lucy se volvió hacia él . - ¿Sería capaz de poner una bala en las doce del reloj que hay en <aquella> pared ?

→par_en: Lucy said to him , " Could you hit the twelve on the clock **over there** on **the** wall ?

<Grafton>: En <aquella> época se construían grandes cámaras de seguridad

→par_en: **Back then** big vaults were poured in place

Estos detonantes pueden ser también de tipo emotivo o afectivo, dado que hemos visto que la deixis evocadora se consigue sobre todo con este demostrativo y, de hecho, en el corpus vemos que se utiliza en fragmentos más poéticos y coloca a menudo con *ojos* o *mirada*, sobre todo en la traducción de de Hériz Ramón (Leonard).

<Leonard>: Se detuvo porque ella tenía de nuevo <aquella> mirada , aquel brillo sonriente que le excitaba

→par_en: He stopped because she had **that** look in her eyes again, that gleam with the smile in it that turned him on.

<Highsmith>: con <aquella> mirada que Tom conocía muy bien , como si supieran quién era
 →par_en: with **that** look Tom knew well , as if they knew who he was

<Leonard>: <Aquellos> ojos apagados que parecían no tener vida pero que no se perdían nada.
 →par_en: **Those** dull eyes that didn't seem to have life in them but didn't miss anything .

<Leonard>: recordando cómo lo había mirado con <aquellos> ojos verdes centelleantes
 →par_en: Jackie looking at him the way she did with **those** sparkly green eyes

Por último, pueden ser detonantes de tipo gramatical, como los casos de señalamiento en discursos en estilo indirecto (Macías Villalobos 2006: 171) o la narración de hechos pasados, en la que *aquel* produce un efecto de mayor formalidad y lejanía temporal que *ese*. Por el mismo motivo, *aquel* no se utiliza tanto en diálogo, y esta puede ser una de las causas de la reducción de su frecuencia y porcentaje en traducciones: ya hemos mencionado que las novelas de nuestro corpus de traducciones contienen mayor porcentaje de fragmentos dialogados que las españolas (Martínez Vilinsky 2014: 244). Por supuesto, siempre hay excepciones, como es el caso de este diálogo informal en el que el traductor elige utilizar *aquel*:

<Leonard>: - Ya . Y tú la ayudaste a salir de <aquel> follón.
 →par_en: Yeah , and you helped her out of **that** mess ..

5.3.5. Distribución de adverbios demostrativos por textos

En este apartado no tenemos, en principio, problemas con las búsquedas de elementos ni con el etiquetado del corpus, pues los adverbios demostrativos aquí analizados son palabras que no cuentan con formas homónimas, al menos en español.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	184	216,8	Highsmith	312	317,4
Madrid	136	328,3	Cornwell	269	244,8
Vázquez	135	153,7	Leonard	267	343,2
Martín	99	176,2	Ellroy	205	186,8
Núñez	76	179	Grafton	199	214,3
Martínez	73	134,3			
Total	703	191,5		1252	256,3

Tabla 96 - Distribución de adverbios demostrativos locativos por texto

Una vez más, realizamos una búsqueda de los demostrativos existentes (mediante el comando [word="(¡|¿)?aquí(¡|¿)?acá(¡|¿)?ahí(¡|¿)?allí(¡|¿)?allá" %c]) en nuestro corpus y agrupamos estos resultados por texto en la tabla 96. En ES sorprende Madrid, el autor con menor número de palabras que, sin embargo, se sitúa en el primer puesto en cuanto a uso proporcional de adverbios demostrativos (sobre todo por su frecuente uso de *aquí*). También el hecho de que Martín y Núñez superen a Vázquez Montalbán. En TRAD, las traducciones de Leonard (sobre todo por su alta frecuencia de uso de *ahí*) y Highsmith (que está a la cabeza en *allá* y *allí*) cuentan con una frecuencia mayor a la esperada si tomamos como base el número de palabras de sus novelas. Sucede todo lo contrario con la de Ellroy, que solo remonta posiciones en el cómputo de *ahí*.

A continuación, presentamos los porcentajes de utilización por texto de cada adverbio demostrativo por separado.

5.3.5.1. AQUÍ

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Madrid	80	193,1	Cornwell	145	132
Giménez	68	80,1	Highsmith	130	132,2
Vázquez	51	58,1	Grafton	96	103,4
Martín	35	62,3	Leonard	92	118,3
Núñez	31	73	Ellroy	88	80,2
Martínez	18	33,1			
Total	283	77,1		551	112,8

Tabla 97 - Distribución de *aquí* por texto

En esta primera tabla podemos ver el motivo por el cual Madrid se encuentra más alto de lo esperado en la tabla 96. Evidentemente, su utilización del adverbio *aquí* es muy superior a la de los demás autores, tanto en cantidad total como en cuanto a frecuencia normalizada con respecto a la extensión de su novela, que es la más corta. Núñez también cuenta con una frecuencia mayor de lo esperable, al contrario que Vázquez Montalbán. En el otro extremo, vemos que la traducción de Ellroy se sitúa en la última posición en utilización de este adverbio a pesar de contar con mayor número de palabras que casi todos los otros textos.

En este corpus, como esperábamos, no hemos encontrado ningún caso del adverbio demostrativo de cercanía *acá*.

5.3.5.2. AHÍ

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	36	64,1	Leonard	81	104,1
Giménez	34	40,1	Ellroy	46	41,9
Madrid	22	53,1	Cornwell	42	38,2
Martínez	17	31,3	Highsmith	27	27,5
Vázquez	16	18,2	Grafton	19	20,5
Núñez	11	25,9			
Total	136	37		215	44

Tabla 98 - Distribución de *ahí* por texto

Es lógico que este adverbio presente una frecuencia de aparición menor a la de *aquí*, puesto que comparte la función de expresar los distintos grados de lejanía con *allí* y *allá*. Aun así, la diferencia es algo mayor en TRAD: *ahí* abarca el 32,4 % de los casos totales de *ahí*, *allí* y *allá* en ES (420), mientras que en TRAD desciende al 30,7% (de 701).

En ES, Madrid vuelve a estar más presente de lo normal, Martín pasa al primer puesto (a pesar de ser el tercero en número de palabras) y Vázquez Montalbán se sitúa en último lugar. En TRAD, el texto de Leonard es quien más sorprende por la alta frecuencia que presenta comparado con el resto de traducciones, pues es el último en número de palabras.

5.3.5.3. ALLÍ

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	76	89,5	Highsmith	131	133,3
Vázquez	48	54,7	Leonard	87	111,8
Martínez	37	68,1	Cornwell	74	67,3
Madrid	34	82,1	Grafton	68	73,2
Núñez	32	75,3	Ellroy	60	54,7
Martín	26	46,3			
Total	253	68,9		420	86

Tabla 99 - Distribución de *allí* por texto

La utilización de *allí* es claramente superior a la de *ahí* —hay casi el doble de casos en ambos subcorpus— y se acerca más a la de *aquí*, lo que indica que la contraposición de cercanía y lejanía déctica se da sobre todo entre *aquí* y *allí*.

Los autores de ES utilizan más este adverbio, en general, aunque hay variaciones individuales: mientras que Giménez Bartlett utiliza *allí* más del doble de veces que *ahí* (y se pone con ello a la cabeza de esta tabla), en Martín, por ejemplo, solo encontramos una diferencia del 30 %. Quizás por la gran diferencia de uso existente en los demás

autores, este último pasa del primer puesto en *ahí* al último en *allí*. Una vez más, Madrid lo utiliza más de lo esperable.

En TRAD, los textos de Highsmith y Ellroy casi intercambian posiciones con respecto a las tablas de frecuencias del adverbio anterior, lo que indica que existen ciertas preferencias por uno u otro demostrativo de lejanía en los distintos autores. El traductor de Ellroy, Sabaté Vargas, es el segundo que más utiliza *ahí* y el que menos utiliza *allí* (aunque la diferencia de frecuencias no es muy grande) y la de Highsmith, Núñez, pasa del anteúltimo puesto en utilización de *ahí* al primero en *allí*, con una diferencia de frecuencias, esta vez sí, bastante importante. Los demás traductores también utilizan bastante más *allí*, excepto el de Leonard, en cuyo texto la diferencia es mínima.

5.3.5.4. ALLÁ

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Vázquez	20	22,8	Highsmith	24	24,4
Giménez	6	7,1	Grafton	16	17,2
Núñez	2	4,7	Ellroy	11	10
Martín	2	3,5	Cornwell	8	7,3
Martínez	1	1,8	Leonard	7	9
Madrid	0	0			
Total	31	8,4		66	13,5

Tabla 100 - Distribución de *allá* por texto

Finalmente, hemos encontrado algunos casos de *allá*, variante funcional de *allí*. Es evidente que en español peninsular no es una forma muy extendida, dado el escaso número de apariciones y el hecho de que un autor no la utilice en ninguna ocasión. De hecho, en el caso de Vázquez Montalbán, parece ser una cuestión de preferencia personal, pues su utilización de esta forma es muy superior a la media de los demás autores. Si comparamos estas cifras con las de *allí*, vemos que el uso de esta forma es prácticamente anecdótico en todos los autores menos en Vázquez Montalbán. En su caso, esta forma representa el 30 % del total de apariciones de *allí* y *allá*, y se utiliza más que la opción *ahí*.

En TRAD, Highsmith es quien más utiliza ambas formas pero con una diferencia muy importante a favor de *allí*. Sin que su uso llegue, ni mucho menos, a superar el de *allí*, Grafton y Ellroy escogen *allá* algunas veces más que el resto de autores, que bajan posiciones en esta tabla. Aquí las frecuencias están repartidas de manera más equitativa entre los cinco textos que en ES.

5.3.6. Comparación de distribuciones de adverbios demostrativos locativos entre subcorpus

Reproducimos los resultados de nuestras búsquedas de adverbios demostrativos en el corpus en la tabla 101, que incluye el porcentaje que representa cada adverbio demostrativo sobre el total de adverbios de sus respectivos subcorpus, la frecuencia normalizada de cada uno y la diferencia de frecuencias entre TRAD y ES. La última columna nos indica si las diferencias halladas son significativas desde el punto de vista estadístico (test de log-likelihood). Una vez más, el color rojo en las casillas indica una sobrerrepresentación del elemento en cuestión en el subcorpus de traducciones y el fucsia, una frecuencia más alta pero no estadísticamente significativa.

Lo primero que llama la atención es que en las traducciones encontramos casi el doble de adverbios demostrativos que en los originales españoles. Aunque, tras normalizar las frecuencias, esta diferencia disminuye, lo cierto es que el color de las casillas de la última columna nos indica que hay bastantes más adverbios demostrativos locativos en TRAD que en ES, lo que puede deberse a que en inglés se utilizan más estos elementos o a que los traductores recurren más a ellos. Lo analizaremos en el apartado de búsquedas en el corpus paralelo.

Adverbio locativo	ES	%	Frec. normalizada	TRAD	%	Frec. normalizada	Dif. TRAD/ES	Significación estadística
Aquí	283	40,26	77,09	551	44,01	112,78	-35,69	28,04 p < 0,0001
Acá	0	0,00	0,00	0	0,00	0,00	0,00	0,00 (n. s.)
Ahí	136	19,34	37,05	215	17,17	44,01	-6,96	2,50 (n. s.)
Allí	253	35,99	68,92	420	33,55	85,97	-17,05	7,84 p < 0,01
Allá	31	4,41	8,44	66	5,27	13,51	-5,07	4,89 p < 0,05
Total	703	100,00	191,50	1252	100,00	256,26	-64,76	39,13 p < 0,0001

Tabla 101 - Comparación de frecuencias de adverbios demostrativos en ES y en TRAD y significación estadística de las diferencias observadas

Tras aplicar el test de significación estadística de log-likelihood a cada par de frecuencias, no solo no encontramos ningún elemento infrarrepresentado, sino que, además, tres de los cuatro presentes en el corpus (*aquí*, *allí* y *allá*) muestran el fenómeno inverso: existe una sobrerrepresentación de estos elementos en el corpus de traducciones con respecto a las frecuencias encontradas en el subcorpus de originales españoles (el cuarto, *ahí*, también está más presente en traducciones pero esta diferencia no llega a ser estadísticamente significativa). Por lo tanto, esta sobrerrepresentación de adverbios locativos es muy significativa; los traductores los tienen muy presentes en su

paleta de recursos lingüísticos y hacen un uso superior de ellos, en comparación con los textos escritos originalmente en español.

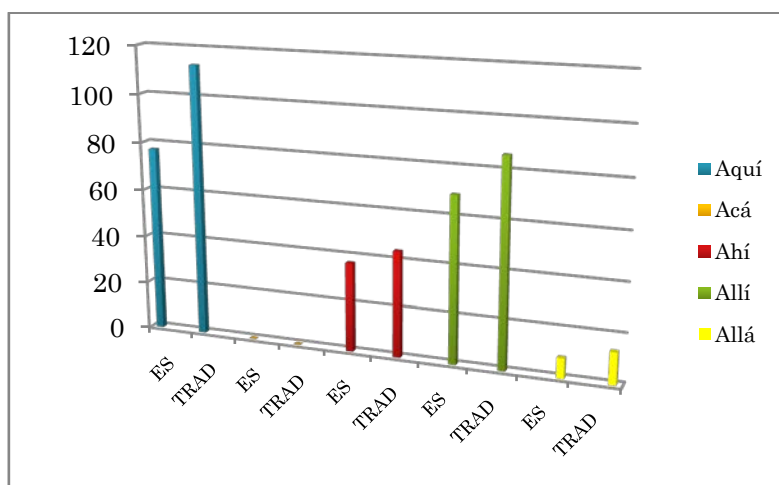


Gráfico 15 - Comparación de frecuencias normalizadas de adverbios demostrativos en ES y TRAD

En general, observamos unos porcentajes de distribución de los distintos adverbios bastante similares en ambos subcorpus. Las frecuencias normalizadas son superiores en todos los casos para el subcorpus TRAD, pero tanto *ahí* como *allí* están dos puntos por debajo de los porcentajes de distribución hallados en ES. Además, se aprecia un uso de *aquí* con respecto al total de elementos superior a ES en un 4%. *Allá* también se utiliza un poco más en traducciones. No obstante, hemos de tener en cuenta que esta diferencia en la distribución no es significativa desde el punto de vista del test de chi cuadrado ($\chi^2=4,11$ (df=3)¹⁸⁰).

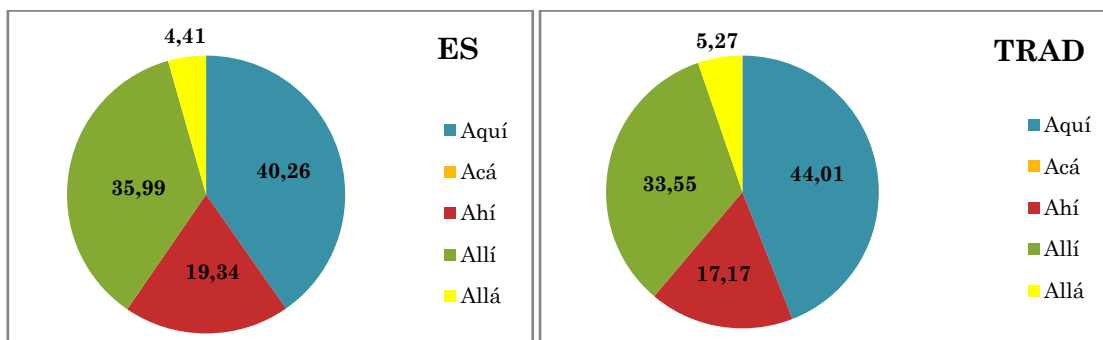


Gráfico 16 - Comparación de porcentajes de uso de adverbios demostrativos en ES y TRAD

¹⁸⁰ El umbral de significación mínimo ($p < 0,05$) está establecido para tablas de 3 grados de libertad en 7,815 (Stefanowitsch 2005).

5.3.7. Análisis del corpus paralelo

El hecho de que los adverbios locativos estén, globalmente, más representados en las traducciones que en los originales podría deberse sencillamente a que en inglés *here* y *there* son de uso más frecuente que toda la gama de sus posibles equivalentes sumados. Para comprobar si esto es así, recurrimos al corpus paralelo y localizamos todas las formas de *here* —848, tras restar los nueve casos en los que la palabra funciona como interjección (p. ej.: «Here, let me help you»)— y *there* —655, tras descartar todos los casos interjección (p. ej.: «There! Anything else you would like me to move?») y de *there* existencial¹⁸¹ (p. ej.: «Is there a phone I can use?»), que están etiquetados como «EX» pero a veces de manera errónea—. En total, 1 503 adverbios demostrativos que, comparados con los 703 hallados en ES y los 1252 de TRAD, arrojan una diferencia muy significativa (146,28 $p < 0,0001$ y 41,69 $p < 0,0001$, respectivamente, según el test de log-likelihood) a favor de EN en ambos casos.

Demostrativo	Frecuencia
Here	848
There	655
TOTAL	1503

Tabla 102 - Frecuencias de aparición de los adverbios demostrativos en el subcorpus EN

En la tabla 103 podemos observar la distribución de *here* y *there* por texto en EN. Como habíamos comentado en el análisis de TRAD, Leonard cuenta con una frecuencia mayor a la esperada tanto en *here* como en *there* (al igual que sucedía con los pronombres y determinantes demostrativos). Sin embargo, suponemos que de Hériz Ramón, su traductor, se deja bastantes *here* por traducir porque, mientras que en EN ocupa el primer puesto, en TRAD es el tercero. Elroy ocupa el último y anteúltimo puesto en estas tablas, a pesar de estar más alto en algunos adverbios en TRAD y de que, en el caso de los pronombres y determinantes, estaba entre los primeros puestos. El texto de Highsmith, por su parte, no ocupa un puesto tan elevado en *here* (el tercero) como sucede en *aquí* (el primero).

¹⁸¹ Según el documento de consulta de TreeTagger, el *there* existencial «is the unstressed *there* that triggers inversion of the inflected verb and the logical subject of a sentence» (Santorini 1991: 3).

HERE	Frec.	Frec. norm.	THERE	Frec.	Frec. norm.
Cornwell	210	204,7	Leonard	249	327,1
Leonard	194	254,9	Highsmith	116	124,8
Highsmith	189	203,3	Cornwell	106	103,3
Grafton	178	190,2	Ellroy	96	103,1
Ellroy	77	82,7	Grafton	88	94
Total	848	185		655	142,9

Tabla 103 - Distribución de *here* y *there* por texto en EN

En cuanto a *there*, Leonard triplica en casos al siguiente en la tabla. Sin embargo, en TRAD se encuentra bastante igualado, si sumamos los casos de *ahí*, *allí* y *allá*, con el texto que ocupa el segundo lugar de la tabla, el de Highsmith (224,9 frente a 185,2). Todo esto nos deja bastante claro que los equivalentes de *here* y *there* en español no son siempre nuestros adverbios demostrativos y, a la inversa, no siempre la utilización de nuestros adverbios viene provocada por un elemento activador en el original. Para comprobarlo, recurriremos una vez más a las concordancias paralelas.

En primer lugar, puede ser interesante partir del inglés, analizando las concordancias paralelas de *here* y *there*, como hicimos en el caso de los determinantes y pronombres demostrativos para averiguar, por ejemplo, si el traductor se decantaba más por uno u otro demostrativo de lejanía.

5.3.7.1. HERE

Elemento equivalente de HERE	Frecuencia	Porcentaje
AQUÍ	395	46,58
Omisión	140	16,51
Omisión por presencia de verbos de estado o movimiento que hacen redundante la inclusión del demostrativo (ej.: venir, llegar, volver, estar, marchar, pasar, entrar, salir, irse, quedarse, traer, acercarse, largarse, llamar, acudir)	114	13,44
ALLÍ	60	7,08
Explicitación de posición o dirección (excepto sintagmas que incluyen un demostrativo)	30	3,54
ESTE (pronombre y determinante)	27	3,18
AHÍ	17	2,00
Sustitución por frase en futuro o imperativo (ej. : «voy a contarte», «te diré», «fíjese»)	16	1,89
Reformulación, frase hecha o fórmula conversacional	16	1,89
Referencia a la primera persona (ej. : «soy», «conmigo», «para mí», «yo», «visitarme», «nos mandan»)	12	1,41
Expresión temporal (ej. : «ahora», «hoy», «al llegar a este punto», «otra vez», «entonces»)	9	1,06
ESE (pronombre y determinante)	7	0,83
AQUEL (pronombre y determinante)	4	0,47
ALLÁ	1	0,12
Total	848	100

Tabla 104 - Elementos equivalentes del adverbio locativo *here* en traducciones

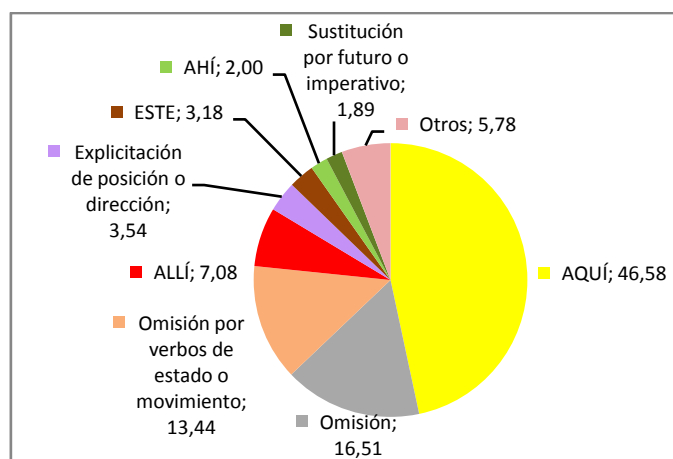


Gráfico 17 - Porcentajes de uso de los distintos elementos equivalentes del adverbio locativo *here* en traducciones

A partir de la tabla de equivalentes de *here* hallados en TRAD y el gráfico de porcentajes de distribución de dichos elementos, vemos que *aquí* presenta en TRAD un porcentaje de utilización del 44,01 %, superior en un 3,75 % a ES (v. tabla 101). Aun así, este porcentaje no llega a igualar el que encontramos en los originales estadounidenses para su equivalente más evidente *here*: 56,42 % (848 de 1503, v. tabla 102) frente al 43,58 % de *there*.

En el 46,58 % de las ocasiones en las que aparece *here* en inglés, se opta por trasladarlo al español mediante su equivalente más evidente, *aquí*. El 30 % de las veces, el demostrativo no se traslada al español. Un motivo bastante frecuente es la utilización de verbos que implican ya una situación o una dirección hacia la que se dirige la acción, sobreentendida en determinados contextos, que equivale a *aquí*. El adverbio podría utilizarse —y, de hecho, se traslada también a veces—, pero no es imprescindible:

<Leonard>: She said , " He ain't <here> , " and closed the door .

→par_trad: - No está - les dijo . Y cerró la puerta .

<Cornwell>: Eddie Heath 's body is <here> . "

→par_trad: - Han traído el cuerpo de Eddie Heath

<Grafton>: I do n't know how you 're going to resolve this little conflict , but I 'm getting out of <here> . "

→par_trad: No sé cómo resolveréis este pequeño conflicto , pero yo me largo .

Las razones para los casos contenidos en la primera categoría denominada «omisión» son múltiples: casos en los que el *here* es una fórmula conversacional que tiene una traducción más o menos estándar (a), casos en los que la utilización de *here*

con valor temporal del inglés sonaría forzada en español (b) o cuando adopta diversos valores para los que en español en principio no utilizaríamos el adverbio (c). De todas formas, no siempre se decide omitirlo, como vemos en los contraejemplos subrayados.

(a) <Cornwell>: Nicholas Grueman <here> .

→par_trad: - Nicholas Grueman al habla .

<Cornwell>: " Dr. Scarpetta , " I answered . " Nicholas Grueman <here> .

→par_trad: - Aquí Nicholas Grueman .

(b) <Highsmith>: A headless corpse , pretty hard to identify , I'd say , fingerprints long ago washed off with the skin . " <Here> Ed forced a laugh that sounded sick . " Let's get out , " Tom said at once .

→par_trad: Ed forzó una carcajada que sonó enfermiza .

<Highsmith>: " He was surprised , it seems , that the bones-the skeleton-seemed to be the objective of Preechard . " <Here> Jean looked hard at Tom . "

→par_trad: - Aquí Jean miró a Tom con firmeza - .

(c) <Grafton>: I have no reason to trust you any more than you trust me . " " And <here> I thought we were such pals , " I said .

→par_trad: - Yo creía que éramos amigos - dije .

<Ellroy>: " Ray , tell Stemmons <here> your chemical angle .

→par_trad: - Ray , cuéntale a Stemmons tu punto de vista químico .

<Cornwell>: " Johnson , how 'bout dusting the door handles so the doc <here> can get in the car . "

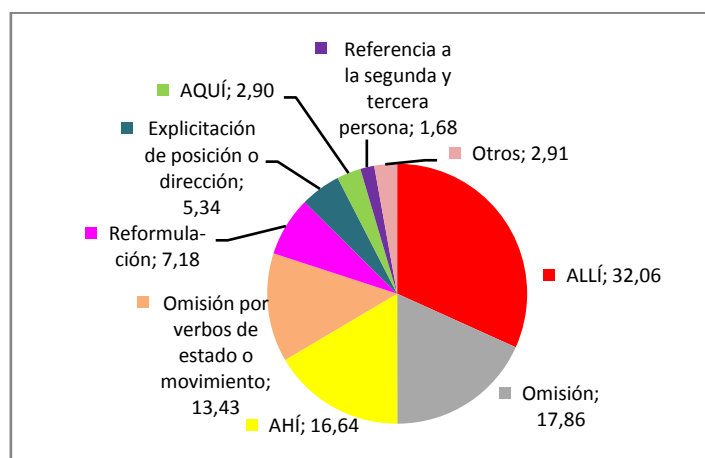
→par_trad: Johnson , ¿qué tal si espolvoreamos los tiradores de las puertas para que aquí la doctora pueda entrar en el coche ?

Un 7 % de concordancias corresponde a *allí*, pero el resto de categorías, entre las que se incluyen la explicitación de la localización a la que el adverbio hace referencia y la sustitución por otros adverbios, por *este*, *ese* y *aquel* en función pronominal y determinante, o por expresiones de otro tipo, no alcanzan porcentajes demasiado elevados. A pesar del alto porcentaje de omisiones y sustituciones, ya vimos en la tabla de comparación de demostrativos que *aquí* estaba sobrerrepresentado en traducciones, por lo que, aunque la diferencia con los originales ingleses sea mayor, no llega a ser tan grande para impedir que se siga utilizando en mayor medida que en nuestro subcorpus ES.

5.3.7.2. THERE

En la tabla 105 enumeramos los equivalentes de *there* hallados en TRAD. Tanto en esta como en el gráfico 18, observamos que *there* se traduce por *allí* casi el doble de veces que por *ahí*, aunque, juntos, ambos adverbios conforman el 48,70 % de las concordancias, un porcentaje que está en la línea de lo que observábamos en el caso de la traducción de *here* por *aquí* y del de los pronombres y determinantes (v. gráficos 7 y 8). También, que la omisión es la segunda categoría más frecuente y, de hecho, si juntamos la segunda y la cuarta, la cifra de omisiones llega casi a igualar a la de la utilización de *allí*.

Elemento equivalente de THERE	Frecuencia	Porcentaje
ALLÍ	210	32,06
Omisión	117	17,86
AHÍ	109	16,64
Omisión por presencia de verbos de estado o movimiento que hacen redundante la inclusión del demostrativo (ej.: estar, quedar, entrar, salir, ir, venir, volver, llegar, haber, llevar, largarse, ser, existir, aparecer, encontrarse)	88	13,43
Reformulación, frase hecha o fórmula conversacional	47	7,18
Explicitación de posición o dirección (excepto sintagmas que incluyen un demostrativo)	35	5,34
AQUÍ	19	2,90
Referencia a la segunda y tercera persona (ej.: «verlo», «los verá», «con vosotros»)	11	1,68
ALLÁ	6	0,92
ESE (pronombre y determinante)	5	0,76
Expresión temporal (ej.: «luego», «ahora» «entonces»)	4	0,61
AQUEL (determinante)	4	0,61
Total	655	100

Tabla 105 - Elementos equivalentes del adverbio locativo *there* en traduccionesGráfico 18 - Porcentajes de uso de los distintos elementos equivalentes del adverbio locativo *there* en traducciones

El resto de categorías, entre las que se cuentan la reformulación, la explicitación del elemento locativo al que hace referencia el adverbio y la sustitución por otros elementos, presenta frecuencias relativamente bajas. Mientras que en la tabla 104 encontrábamos representados los tres grados de lejanía del demostrativo, aquí el primer grado, *este*, no presenta ni un caso (por el contrario, *aquí* sí que aparece, con 19).

Si realizamos una comparación por demostrativos, contraponiendo *here* a *aquí* y *there* a *ahí*, *allí* y *allá*, obtenemos los siguientes resultados:

- 848 casos de *here* frente a los 283 de *aquí* en ES: **LL +183,76 p < 0,0001**
- 848 casos de *here* frente a los 551 *aquí* en TRAD: **LL +83,88 p < 0,0001**
- 655 casos de *there* frente a las 420 formas de lejanía de ES: **LL +12,82 p < 0,001**
- 655 casos de *there* frente a las 701 formas de lejanía de TRAD: **LL -0,01 (n. s.)**

La frecuencia de adverbios demostrativos locativos en los textos en inglés es claramente muy superior a la de los textos en español, tanto original como traducido. Observamos, sin embargo, que la diferencia es bastante mayor en el caso de *here*, puesto que en *there*, la significación estadística es menor con respecto a la diferencia entre EN y ES, y nula con respecto a EN y TRAD. Esto podría sugerirnos que quizás las distintas formas del tercer grado de lejanía estén utilizándose en las traducciones en mayor medida que en originales españoles por influencia de los originales ingleses. Por otro lado, mientras que en inglés hallamos más adverbios de cercanía que de lejanía, en español, tanto original como traducido, sucede lo contrario; la influencia de los textos originales, por tanto, parece tener un límite.

En las siguientes cuatro tablas, hemos partido de los textos meta para agrupar en categorías los elementos hallados en sus respectivos textos originales que podrían haber dado lugar a la utilización del adverbio locativo correspondiente o, por el contrario, la falta de cualquier elemento activador de este tipo.

5.3.7.3. AQUÍ

Como vemos en la tabla 106 y el gráfico 19, el equivalente más frecuente de *aquí* es, previsiblemente, *here*. Casi todos los casos de *here* cumplen una función deíctica espacial, puesto que se encuentran en las frases de diálogo directo de las novelas, por lo que el referente al que se hace alusión es un elemento del contexto cercano de los participantes en la interacción.

<Highsmith>: - ¡Qué campos tan bonitos hay <aquí> ! - dijo Ed mientras se alejaban de la región situada al este de Fontainebleau -

→par_en: " What beautiful fields here ! " Ed said as they rolled away from Fontainebleau eastward.

<Grafton>: Tengo una moradura desde aquí hasta <aquí> - dijo con orgullo .

→par_en: I got a bruise goes from here to here , " she said with satisfaction .

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
Here	398 ¹⁸²	72,23
∅	40	7,26
This/these	32	5,81
Phrasal verbs ¹⁸³ (move/bring/come/stop/look/get/walk/show/pass/sit/stick/nose + in/out/by/up/around)	20	3,63
There	19	3,45
Sintagma nominal, preposicional o adjetival de significado locativo no incluido en las demás categorías ¹⁸⁴	14	2,54
Adverbios de lugar o movimiento: Around, over, back, nearby	12	2,17
Expresiones temporales (today, now, so far)	6	1,09
That	5	0,91
Reformulación	5	0,91
TOTAL	551	100

Tabla 106 - Elementos activadores de la utilización del adverbio *aquí* en traducciones

¹⁸² Como sucedía con los pronombres y determinantes y, seguramente, por los mismos motivos, las cifras en una dirección y en otra no coinciden por completo tampoco en el caso de ciertos adverbios. Observamos una diferencia de 3 casos entre *here->aquí* (v. tabla 104) y *aquí->here* (v. tabla 106), de 2 en el caso de *there->ahí*, de 1 en el de *here->allí* y ninguna diferencia en el caso de *here->ahí*, *here->allá*, *there->aquí*, *there->allí* y *there->allá*.

¹⁸³ Según listado y clasificación de Heaton (1979).

¹⁸⁴ Ejemplos:

<PatriciaHighsmith>: Su mujer se quedó en casa.. .. <aquí> .

-->par_en: His wife stayed at home-in my town .

<Leonard>: - No , éste es de <aquí> - respondió Nicolet - .

-->par_en: " No , this one's a homey , " Nicolet said .

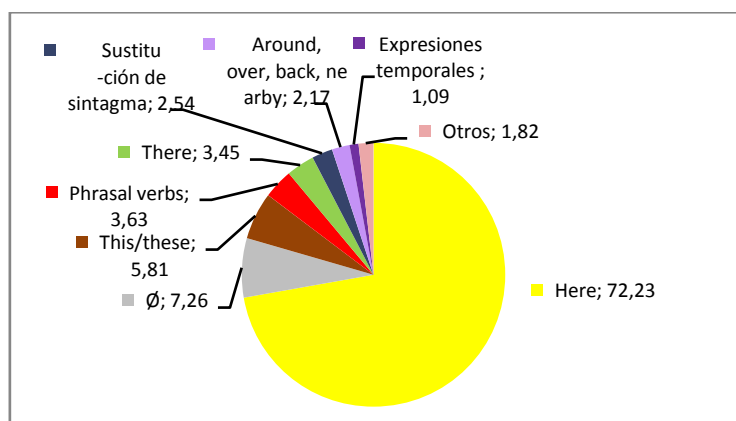


Gráfico 19 - Porcentajes de uso de los distintos elementos activadores del adverbio *aquí* en traducciones

También encontramos algunos casos de catáfora, en los que el hablante informa posteriormente al interlocutor de a qué lugar exactamente se refiere con el adverbio demostrativo:

<Ellroy>: Puede llamarme <aquí> , al Bel-Air .

→par_en: You can call me **here** at the Bel-Air .

En estos casos, a veces el consecuente no es un referente espacial sino textual:

<Cornwell>: - He <aquí> lo que vamos a hacer - prosiguió -

→par_en: " **Here** 's the way we 're going to work this , " he said .

<Ellroy>: Y.. . bueno , Dave , <aquí> es donde la cosa se pone extraña y graciosa

→par_en: David , **here** 's where it gets strange and funny

En algunos casos, el *here* posee función deíctica temporal en lugar de espacial:

<Highsmith>: - <Aquí> Jean miró a Tom con firmeza - .

→par_en: **Here** Jean looked hard at Tom .

<Highsmith>: Quizá vengamos . - <Aquí> soltó una risa gutural - .

→par_en: He gave a full-throated chuckle **here** .

En otras ocasiones, la deixis temporal del original, expresada mediante adverbios temporales, se transforma en espacial en español:

<Highsmith>: Es difícil ver el camino desde <aquí>...

→par_en: Hard to see the lane now.

<Cornwell>: Pero hay otras personas a las que sí podrán escuchar <aquí>

→par_en: But there are others you will hear from today .

Habíamos visto, partiendo de los originales, que el traductor omitía muchas veces el adverbio. Como se ve en la tabla 106 y el gráfico 19, el segundo grupo más importante aquí lo constituyen los casos en los que no encontramos un equivalente lingüístico claro que pueda haber inducido al traductor a utilizar el adverbio demostrativo. El porcentaje, sin embargo, es bastante más bajo en este caso (7,26 % frente al 16,63 % que observábamos en la tabla 104 de equivalentes de *here*), por lo que entendemos que la omisión es más abundante que la adición en el caso del primer grado.

En dichos ejemplos, a veces vemos que a través del contexto se entienden los motivos argumentales o estilísticos por los que el traductor ha decidido compensar y añadir este demostrativo.

<Leonard>: Esos dos vienen hacia <aquí> - anunció

→par_en: He said , " Those guys are coming,"...

<Cornwell>: Ahora mismo está andando por <aquí> cerca .

→par_en: He 's off on foot somewhere .

<Highsmith>: Estamos <aquí> charlando.. . y yo me estoy vistiendo . - Se rió -

→par_en: We are talking-and I am dressing . " She laughed .

Sin embargo, encontramos muchos otros en los que parece que este no resulta imprescindible, puesto que la frase se entendería igual con demostrativo o sin él, pero observamos una necesidad de enfatizar, especificar o explicitar por parte del traductor, como sucedía en el caso de los pronombres y determinantes:

<Cornwell>: Le sugería que viniera <aquí> , a Richmond .

→par_en: She suggested he come to Richmond . "

<Leonard>: Se sentía tan ajeno al ambiente como en el hotel de los colombianos , pero <aquí> el espectáculo era mejor

→par_en: It was no more Louis's scene than the hotel where the Colombians hung out , but the show was better

<Cornwell>: El que tenemos <aquí> en el departamento es una porquería .

→par_en: The one we got in the squad room sucks . "

A veces incluso se añade información que el texto original no proporciona, al menos en ese punto de la historia:

<Ellroy>: Voy a encontrar a Stemmons y a secuestrarle <aquí> , donde los federales no puedan encontrarle .

→par_en: I 'm going to wrap that Watts shootout up , find Stemmons and sequester him where the Feds can't find him .

El tercer puesto de la tabla lo ocupa el demostrativo *this*, en función pronominal o como determinante acompañado de sustantivo, lo que nos indica que los tres paradigmas, adverbios locativos y pronombres/determinantes, son a veces intercambiables en sus funciones deícticas y fóricas.

<Cornwell>: ¿Hay alguien más <aquí> , aparte de usted y yo ?

→par_en: Somebody else in **this** office besides you and me ? "

<Ellroy>: - <Aquí> dice que has intentado vender una vajilla robada en la casa de empeños Happytime .

→par_en: " **This** says you attempted to sell stolen silverware to the Happytime Pawnshop .

Vemos también, en el caso de los adverbios, que podemos encontrar el demostrativo de significado teóricamente opuesto (en este caso, *there*) entre los elementos activadores en inglés. A pesar de que de las concordancias se desprende que el elemento activador principal de *aquí* es *here*, en algunos casos, el traductor cambia este tipo de deícticos según el punto de vista que adopta en el contexto.

<Cornwell>: Haga el favor de venir por <aquí> , señora .

→par_en: If you'll step over **there** , ma'am . "

<Leonard>: - ¿Y qué pone <aquí> arriba ? - preguntó Dulzón - .

→par_en: Sweatman said , " What's it say on top **there** ?

Las últimas dos categorías, los *phrasal verbs* y los adverbios de lugar o dirección, indican un desplazamiento hacia un lugar determinado, que en español se expresa mediante *aquí*, o un movimiento sin rumbo fijo, especialmente en el caso de *around*, que se traduce como «por aquí».

<Highsmith>: - ¿Probamos <aquí> ? - preguntó Tom.

→par_en: " Shall we look **in** ? " Tom asked .

<Cornwell>: No he visto ningún alborotador por <aquí> .

→par_en: " I have n't seen any troublemakers **around** .

Richardson explica, al respecto de esta abundancia de partículas adverbiales en los originales (1998: 134):

Slobin (1995) has shown how English tends to pay more attention to the concepts of path and direction than Spanish. (...) Thus, English has a greater facility for concatenating a series of statements about path and direction ('up the hall', 'along the corridor', 'down the stairs', etc.) while directionality in Spanish is normally expressed within the verb itself (*subir*, [to go up]; *entrar*, [to go in], etc.).

5.3.7.4. AHÍ

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
There	111	51,63
∅	21	9,77
Here	17	7,91
That	17	7,91
Around, somewhere	16	7,44
Reformulación	14	6,51
Sintagma nominal, adjetival o adverbial con sentido locativo	9	4,19
Along, over, yonder	5	2,32
This	3	1,39
Expresiones temporales (now, so far)	2	0,93
TOTAL	215	100

Tabla 107 - Elementos activadores de la utilización del adverbio *ahí* en traducciones

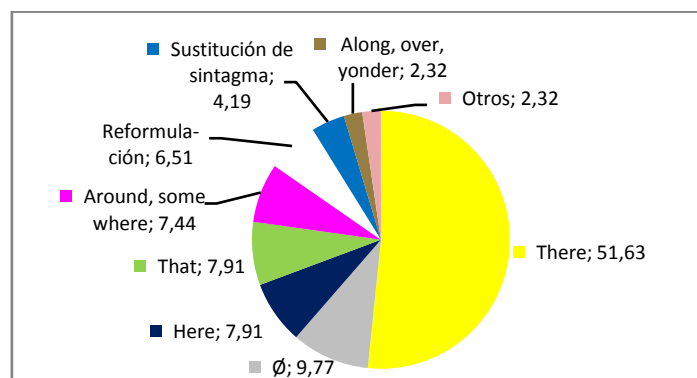


Gráfico 20 - Porcentajes de uso de los distintos elementos activadores del adverbio *ahí* en traducciones

There también aparece sobre todo en diálogo, con función deíctica espacial o anafórica:

<Leonard>: De nuevo actuaba el chico bueno - ¿Ves ese aparcamiento de <ahí> abajo ?

→par_en: The good guy again , saying now , " That parking lot you see right there ?

<Leonard>: No, espere. Tengo el coche en el aeropuerto . - Puedo dejarla <ahí>

→par_en: "No, wait. My car's at the airport." " I can drop you off there . "

Como en casos anteriores, aunque el elemento activador principal de *ahí* es *there*, este demostrativo puede derivar también de la traducción de muchas expresiones distintas, incluido el demostrativo que, en principio, consideraríamos opuesto en significado (*here*), como vemos en los ejemplos.

: <Leonard>: Sin mover la boca , le dijo : - ¿Quiere salir de <ahí> ?

→par_en: Without moving her mouth she said , " Will you get out of here ? "

<Cornwell>: A la mañana siguiente empezó a caer una ligera nevada mientras Lucy y yo nos vestíamos con unas prendas de esquí que podían divisarse desde <ahí> hasta el Eiger

→par_en: The next morning a light snow began to fall as Lucy and I dressed in ski clothes that could be spotted from here to the Eiger .

Por otra parte, volvemos a observar ese paralelismo del que hablan las gramáticas entre determinantes y pronombres demostrativos y adverbios demostrativos, puesto que *that* ocupa el cuarto lugar en importancia de la tabla.

<Cornwell>: <Ahí> no se equivoca - asintió .

→par_en: " You got that right , " he said.

<Ellroy>: <Ahí> es dónde envió el dinero .

→par_en: That 's where I send the gelt .

Lógicamente, entonces, también encontramos tres casos de *this*:

<Cornwell>: - El cuarto de los invitados está por <ahí> - le indiqué - .

→par_en: " The guest bedroom is back this way , " I said .

Casi todos los ejemplos de *around* y *somewhere* se traducen mediante la expresión «por ahí», puesto que, según la RAE (2009: 1316), esta expresión lexicalizada o semilexicalizada se emplea «para designar el lugar impreciso en el que está algo o alguien, o bien la zona por la que se mueve», y estos dos adverbios transmiten justamente ese matiz de imprecisión locativa.

<Leonard>: - Anda por <ahí> .

→par_en: " He's **around** . "

<Highsmith>: - ¿Y suele ir por <ahí> molestando a la gente ?

→par_en: " He makes a habit of going **around** annoying people ? "

Las reformulaciones provienen sobre todo de expresiones idiomáticas y frases hechas:

<Ellroy>: las cosas no terminarían <ahí>

→par_en: and lots of other shit would hit the fan

<Ellroy>: Por <ahí> va bien , teniente .

→par_en: " You 're on track .

Finalmente, muchas de las adiciones, categoría que ocupa el segundo puesto en nuestra tabla, obedecen claramente a razones estilísticas, para cohesionar más las frases o para remarcar el hecho de que una acción es continua:

<Cornwell>: Su mirada se desvió de mi cara al montón de papeles financieros , y de <ahí> a la libreta de notas cubierta con mi caligrafía indescifrable.

→par_en: Her eyes moved from my face to the stack of financial files to the legal pad filled with my indecipherable penmanship .

<Cornwell>: Nos hizo pasar al comedor , donde había un montón de ropa por planchar sobre la mesa , y de <ahí> a la sala de estar , en la parte delantera de la casa .

→par_en: He led us through the dining room , with laundry piled on the table , then into the living room at the front of the house .

<Ellroy>: No les gustaría tenerle por <ahí> haciendo preguntas .

→par_en: They would n't appreciate you asking questions .

<Highsmith>: pero Pritchard va detrás de la pista de Murchison... o, al menos, va diciendo su nombre por <ahí> .

→par_en: but Pritchard's on to Murchison-or at least mentioning his name . "

Una vez más, sin embargo, el porcentaje de adiciones que muestra la tabla —9,77 %— es inferior al de las omisiones que hallábamos partiendo de los originales —un 18,26 % (v. tabla 105)—, incluso si juntáramos las adiciones de los tres demostrativos de lejanía (73 de 701 casos totales, un 10,41%).

5.3.7.5. ALLÍ

Este demostrativo se encuentra, en más ocasiones que los demás, fuera de los pasajes dialogados, es decir, del discurso directo. Además, el fenómeno que encontrábamos en los casos anteriores se da en mayor proporción con *allí*, puesto que el demostrativo opuesto en significado, *here*, ocupa la segunda posición en la tabla.

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
There	210	50,00
Here	61	14,52
∅	49	11,67
Sintagma preposicional, nominal o adverbial	19	4,52
Reformulación	17	4,05
Adverbios de lugar o movimiento: around, over ¹⁸⁵ , in, out, nearby, back, off	16	3,81
Phrasal verbs (move/blow/be/go/stick/stroll/prowl/dawdle/drop/pass/walk + in/out/by /through/on/around/about/past/over)	16	3,81
That	15	3,57
Where(ver) (trad.: casi siempre «allí donde»)	13	3,10
This (far/direction)	4	0,95
TOTAL	420	100

Tabla 108 - Elementos activadores de la utilización del adverbio *allí* en traducciones

¹⁸⁵ Además de los 7 *over* que se traducen por *allí* encontramos 8 casos en los que este adverbio se combina con *there*. No los hemos contado en esta categoría puesto que hemos seguido nuestra pauta de incluirlos en la más general o con mayor número de concordancias.

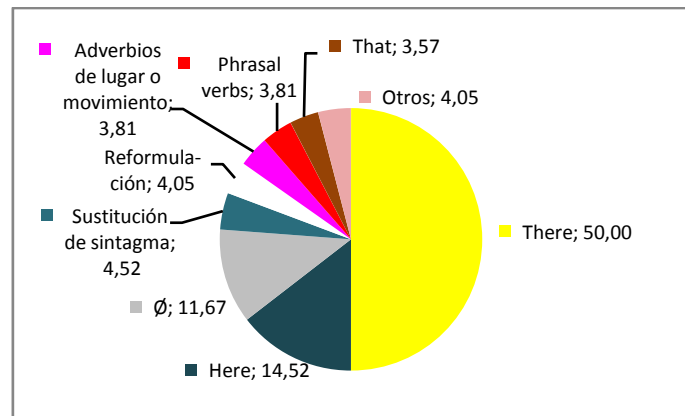


Gráfico 21 - Porcentajes de uso de los distintos elementos activadores del adverbio *allí* en traducciones

La explicación es la que ya hemos mencionado más de una vez. Como vemos en los ejemplos incluidos a continuación, en el corpus de originales ingleses encontramos que se utilizan sin problemas los deícticos de momento presente y espacio cercano para referirse a hechos en la narración, el discurso indirecto o el indirecto libre, cuando en español, las normas de estilo y gramaticalidad nos inducen a cambiarlos.

El frío era <allí> más pronunciado . El aire estaba más enrarecido y el viento era más cortante .
 →par_en: **Here** , the cold was more pronounced . The air felt thinner and the wind more cutting .

<Cornwell>: No quería que Lucy estuviera <allí> , y yo sabía que ella se daba cuenta.
 →par_en: He did not want Lucy **here** , and I knew she sensed it .

<Leonard>: ¿Qué estaba pasando <allí> ?
 →par_en: What was going on **here** ?

<Cornwell> : se identificó como enfermera jefe cuando Trent explicó por qué estábamos <allí> .
 →par_en: identified herself as the head nurse after Trent explained why we were **here** .

Por supuesto, no siempre sucede así, como hemos podido comprobar partiendo de los originales ingleses, en donde observamos también omisiones, reformulaciones y transformaciones a discurso directo (en las que se conserva el adverbio de cercanía):

<Leonard>: Now young guys in the crowd were calling them racist motherfuckers , telling them to go home , go on , get out of <here> , and it looked like the show was over .
 →par_trad: Ahora los jóvenes de la multitud los llamaban cabronazos racistas y les decían que se largaran a casa , **que desaparecieran** , y pareció que se acababa el espectáculo

<Leonard>: They come in yelling down , down , down , screaming it at you , and the next thing you had was a shotgun against your head as you're saying what is this , man , what's going on <here> .

→**par_trad**: Entran gritando « al suelo , al suelo » , e inmediatamente ves que tienes un rifle apoyado en la cabeza mientras dices « qué es esto , tío , qué pasa **aquí** » .

Muchos de los casos de adición del demostrativo, la tercera categoría en orden de frecuencia, obedecen una vez más a cuestiones de estilo o ritmo, puesto que no parecen responder a una necesidad de mayor claridad o explicitación:

<Cornwell>: Hice girar la llave en la cerradura sin saber qué esperaba encontrar <allí>

→**par_en**: As I unlocked the door , I did not know what I expected to find

<Highsmith>: ¿Cuánto tiempo piensa quedarse Heloise <allí> , Tom ?

→**par_en**: How long does Heloise intend to stay , Tom? "

<Ellroy>: Lo recorrí hasta llegar a la cumbre . Desde <allí> , observé el panorama : Provocadores cargando contra la línea de contención de los uniformados .

→**par_en**: I walked it , hit the top , caught the view: Hecklers bucking bluesuit containment-the main road cordoned off .

<Cornwell>: Ella bostezó y se alejó hacia el borde del patio para aprovechar un esquivo rayo de sol . Una vez <allí> , alzó la cara hacia él y cerró los ojos .

→**par_en**: She yawned and wandered to the edge of the yard to catch an elusive ray of sun . Turning her face to it , she closed her eyes .

Vemos más variedad de elementos que indican un movimiento hacia una dirección determinada o una ubicación, agrupados en las categorías de adverbios y de *phrasal verbs*. En *ahí*, estos elementos se distribuían en apenas cinco verbos y la mayoría correspondía a *around* y *somewhere*, es decir que hacían referencia, sobre todo, a un tipo de movimiento circular o errático, a diferencia de los que vemos en los ejemplos a continuación.

<Ellroy>: Corrí hacia <allí> . La puerta número 3 estaba cerrándose .

→**par_en**: I sprinted **over**-door number 3 was just closing .

<Ellroy>: Mensajes pendientes : Meg , para decir que se pasaría por <allí> ; Welles Noonan , para recordarme el turno de vigilancia ,

→**par_en**: Messages waiting : Meg said **drop by** ; Welles Noonan said the guard gig , six sharp .

<Grafton>: Ya hay uno que se quiere instalar <allí> en cuanto la vivienda esté lista

→**par_en**: We got a guy who wants to **move in** once the place is ready .

5.3.7.6. ALLÁ

Finalmente, la frecuencia de aparición de *allá* es bastante baja aunque, a diferencia de *acá*, sí lo encontramos en 66 ocasiones en nuestro subcorpus de traducciones. Esto significa que, a pesar de ser una alternativa preferida en otras regiones hispanohablantes, no está completamente fuera de los usos peninsulares.

ELEMENTO ACTIVADOR	Frecuencia	Porcentaje
Beyond	13	19,70
Away	10	15,15
Down	7	10,61
There	6	9,09
Over	6	9,09
Past, after	6	9,09
On, ahead, further	6	9,09
∅	3	4,55
Across, through	2	3,03
Outside, other	2	3,03
Reformulación	2	3,03
Back	1	1,51
Around	1	1,51
Here	1	1,51
TOTAL	66	100

Tabla 109 - Elementos activadores de la utilización del adverbio *allá* en traducciones

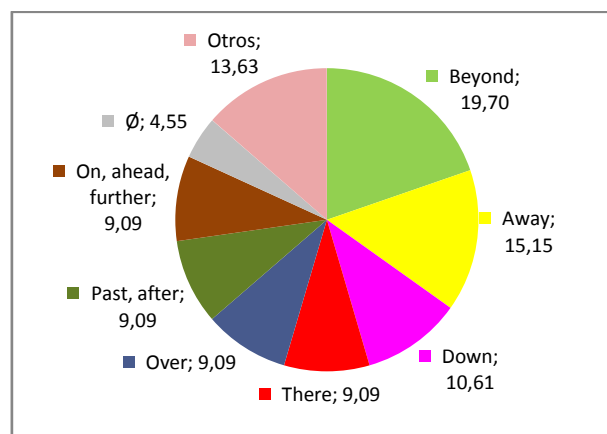


Gráfico 22 - Porcentajes de uso de los distintos elementos activadores del adverbio *allá* en traducciones

Algunos autores opinan que este adverbio aporta un matiz de movimiento o direccionalidad. Nosotros hemos encontrado que no traduce a *here* prácticamente en ninguna ocasión, a diferencia de los adverbios anteriores, y que en pocas ocasiones reemplaza a *there*. Además, como vemos en el gráfico 22, los elementos que propician su traducción en español son variados y sus porcentajes están distribuidos de manera

bastante más equilibrada que en los casos anteriores, lo que indica que no hay uno o varios elementos específicos que activen su utilización.

Al parecer, tiene que haber una palabra o expresión, como las que vemos en la tabla, que aporten ese matiz extra de lejanía o movimiento, para que el traductor se decida a utilizar este adverbio (y, al menos en nuestro corpus, lo hace, casi siempre, como parte de la expresión «más allá»). A este respecto, señala Nilsson (1984: 234) que:

(...) *allá* parece distinguirse de *allí* no solamente por indicar un lugar menos circunscrito o determinado, sino por tener en todos sus aspectos un sentido más vago y débil. Sólo en unos enunciados de carácter translocal se usa *allá* de modo independiente con una referencia nítida. En los demás casos donde *allá* indica un lugar concreto, lo hace generalmente con el apoyo de adverbios o complementos accesorios que le precisan el sentido.

<Highsmith>: El azul del Atlántico se extendía más <allá> , visible a través de un hueco del muro .

→par_en: The blue Atlantic Strait lay **beyond** , visible through a break in the wall .

<Cornwell>: Un par de calles más <allá> hay una iglesia metodista .

→par_en: " There 's a Methodist church two blocks **away** .

<Grafton>: Vive dos casas más <allá> con su amiga Minnie Paxton , pero ahora están fuera .

→par_en: She lives two doors **down** with her friend , Minnie Paxton , but they 're out of town right now .

<Leonard>: - Tres mesas más <allá> , con un vestido azul .

→par_en: " Three tables **over** , in the blue dress .

<Leonard>: unos kilómetros más <allá> de la prisión de Loxahatchee , torció por Southern para meterse luego entre los matorros

→par_en: a few miles **past** the Loxahatchee Road Prison , he turned off Southern to head through open scrub

<Highsmith>: Tom se detuvo otra vez y susurró : -Unos nueve metros más <allá> , ya lo verás.. . Podemos meternos por la hierba .

→par_en: Tom paused again and whispered , " About ten meters **on** , you see-we can get onto the grass .

De siete ocasiones en las que aparece *there* como elemento activador, en cinco lo hace formando parte de la expresión *here and there*, y en una acompaña al adverbio *down*:

<Cornwell>: incrustados aquí y <allá> había fragmentos de lo que supuse debían ser datos .

→par_en: embedded here and there were fragments of what I assumed were data .

<Cornwell>: ¿Es que <allá> en Sudamérica os dan fiesta todo el mes de enero?

→par_en: Down there in South America they give you January off ? "

Aun así, aparece mejor representado en las traducciones que en originales, quizás por la abundante presencia de estos elementos activadores en el texto original —como hemos comentado, el inglés parece prestar más atención a las nociones de dirección y trayectoria—, que aportan un matiz de movimiento que hace, quizás, que al traductor no le parezca suficiente utilizar *allí* y se decante por esta alternativa.

También encontramos un único caso de deixis temporal, traducida por la expresión «allá por»:

<Grafton>: Durante la Depresión , <allá> por 1932 , eran pocos los que podían ganar mil dólares al año.

→par_en: During the Depression - back in 1932 - most guys were n't even pulling in a grand a year .

5.3.8. Conclusiones parciales: adverbios demostrativos.

Como ya lo había hecho Nilsson (1984: 240) en el estudio sobre el uso de los adverbios deícticos en español, portugués y catalán que hemos citado un poco antes, observamos que la distribución en español no es del todo equilibrada: la mayor parte de las concordancias se concentran en los adverbios *aquí* y *allí*, con predominancia del primero, sobre todo en traducciones.

Todos los elementos analizados en este apartado están mejor representados en traducciones que en originales, a excepción de *ahí*, en el que la diferencia no llega a ser estadísticamente significativa y *acá*, debido a que no hemos encontrado ningún caso en ninguno de los subcorpus. Por otro lado, no podemos afirmar, dado el resultado del test de chi cuadrado, que haya diferencias significativas entre subcorpus en cuanto a distribución de adverbios demostrativos, a pesar de que hayamos observado un incremento en TRAD de los porcentajes de utilización de *aquí* y *allá* en detrimento de *ahí* y *allí*.

Habíamos aventurado al principio que el hecho de que los adverbios locativos estuvieran mejor representados en las traducciones que en los originales podría deberse

a que en inglés *here* y *there* son de uso más frecuente que toda la gama de sus posibles equivalentes sumados. Lo que encontramos es que, efectivamente, hay más adverbios demostrativos locativos en el subcorpus de originales en inglés. Pero no parece ser esta la única causa de la sobrerrepresentación en las traducciones, pues también hemos visto que estos adverbios no siempre se traducen.

Demos- trativo	Equivalente en TO <i>here</i> (frecuencia)	Equivalente en TO <i>here</i> (porcentaje)	Equivalente en TO <i>there</i> (frecuencia)	Equivalente en TO <i>there</i> (porcentaje)	Total	Porcentaje con respecto al total de concordancias
Aquí	398	95,44	19	4,56	417	75,68 (de 551)
Ahí	17	13,28	111	86,72	128	59,53 (de 215)
Allí	61	22,51	210	77,49	271	64,52 (de 420)
Allá	1	14,28	6	85,71	7	10,61 (de 66)

Tabla 110 - Porcentaje de utilización de los adverbios demostrativos españoles en el texto meta provocada por la presencia en el texto original de sus equivalentes ingleses

En la última columna de la tabla 110, vemos que los adverbios demostrativos españoles que surgen de la traducción de sus adverbios equivalentes ingleses siempre están por debajo del 80 %; lo que significa que los traductores se valen de ellos para traducir otras expresiones y estructuras cuya función no tiene por qué ser únicamente la de indicar lejanía o cercanía espacial, como hemos visto en el análisis del corpus paralelo. Esto se aprecia sobre todo en los adverbios de lejanía, y podría explicar el mayor porcentaje de adverbios de lejanía con respecto a los de cercanía presentes en español tanto original como traducido. De hecho, en el caso de *allá*, los casos que traducen a *here* o *there* son prácticamente inexistentes, puesto que la mayor parte de su presencia en traducciones es en la forma de expresiones (semi)lexicalizadas, como «más allá». Además, muchas veces, dependiendo del contexto y de si el discurso es directo o indirecto, las referencias espaciales pueden cambiar, incluso del extremo más cercano al más lejano de la gradación deíctica, dando lugar a casos como los que hemos analizado, en los que un adverbio de cercanía como *here* se traduce por uno de lejanía como *allí*. De hecho, *allí*, según la tabla, es el que más se presta a traducir el elemento opuesto en la escala (en un 22,51 % de las ocasiones). Esta puede ser otra de las razones por las que en traducciones vemos un porcentaje mayor de adverbios de lejanía que de cercanía, a pesar de que en inglés suceda a la inversa.

El uso total de adverbios demostrativos locativos en traducciones es, por supuesto, objetivamente mayor que en originales españoles. Ciertas cuestiones de contraste lingüístico o gramatical hacen que los traductores tiendan a utilizar estos elementos para explicitar conexiones o significados del texto original, para cohesionarlo,

o para trasladar ciertas frases hechas o expresiones metafóricas del inglés a través de fórmulas estereotipadas del español que incluyen estos demostrativos. Todo esto, sumado al hecho de que estos elementos —principalmente *aquí* y *allí*— son muy frecuentes en español original, influye en el mayor uso de estos elementos en traducciones.

De hecho, un promedio de un 8,31 % de concordancias son adiciones que surgen de motivos de cohesión o estilo y un 3,62 % de reformulaciones (en total, 11,93 %). No obstante, también hemos visto que las adiciones conforman un porcentaje bastante menor al de las omisiones de estos adverbios.

Como apunta Goethals (2013: 529-530), con un cuarto componente de traducciones al inglés de nuestros textos originales españoles podríamos averiguar si sucede lo mismo en la dirección contraria. De esta manera podríamos observar si se produce un incremento o una reducción de la presencia de los demostrativos y, por tanto, decidir si la sobrerrepresentación detectada es un fenómeno independiente de las lenguas implicadas —lo que confirmaría su estatus de candidato a universal de traducción— o está causada por diferencias contrastivas entre ambas, es decir, porque en español no se utilizan los adverbios tanto como en inglés y, por tanto, esta característica de los textos originales están influyendo en las decisiones de traducción.

En un trabajo anterior (2007: 91), Goethals utiliza esta metodología para discutir los resultados de un trabajo del año 2004 de Whittaker. Esta autora, habiendo hallado una sobrerrepresentación del demostrativo en traducciones (del francés al noruego con respecto al noruego original), lo calificó de fenómeno inherente al proceso de traducción: una estrategia de explicitación de la relación entre el elemento anafórico y su antecedente para facilitar la lectura. Sin embargo, los resultados de Goethals, que, a diferencia de Whittaker, utiliza un corpus bidireccional para la combinación holandés-español no apoyan esta hipótesis sino la de que se trata de diferencias estilísticas entre las lenguas.

Sea como sea, los datos nos indican que los adverbios demostrativos son un elemento tan esencial del lenguaje que su uso no está supeditado únicamente a la influencia de determinados elementos activadores en inglés, sino que depende también de las normas de la lengua meta. Por tanto, creemos que la sobrerrepresentación de adverbios demostrativos locativos en las traducciones puede estar motivada por el uso frecuente de estos elementos en español, a lo que se suma la potencia del estímulo que supone su alta frecuencia de aparición en inglés.

5.4. Conclusiones generales: demostrativos

El principal objetivo de nuestro trabajo era averiguar si alguna de las formas que componen los tres grados de cercanía y lejanía deíctica del paradigma de los pronombres y determinantes demostrativos y del de los adverbios demostrativos locativos se encontraba infrarrepresentada en textos traducidos al español debido a la falta de una equivalencia estructural en la lengua de partida capaz de activar la utilización de dichos elementos en proporción natural, es decir, similar a la que se daría espontáneamente al escribir un texto original en español. Nuestra hipótesis era que encontraríamos una menor frecuencia de uso del tercer grado (*aquel*) en el caso de los pronombres y determinantes y de alguna de las dos formas más habituales del grado de lejanía en el caso de los adverbios locativos. A continuación resumimos las principales conclusiones extraídas del análisis:

- Las pruebas estadísticas indican que las diferencias halladas entre el subcorpus de originales españoles y el de traducciones de inglés a español son significativas. Los determinantes y pronombres demostrativos en conjunto están infrarrepresentados en TRAD con respecto a ES y además muestran una distribución distinta dentro de cada subcorpus.
- En cuanto a los tres grados del sistema del demostrativo en español, la comparación arroja los siguientes datos:
 - **Este:** la proporción de su forma determinante aumenta cuando se trata de textos traducidos. A pesar de que no sucede lo mismo en su forma pronominal, puede afirmarse que aumenta en proporción si tomamos como referencia la suma de ambas categorías. A fin de cuentas, y a pesar de que la categoría gramatical de la palabra puede ser interesante desde el punto de vista de la organización del análisis, no nos interesa tanto a la hora de extraer conclusiones como la división existente en la mente del traductor entre los distintos grados deícticos a los que alude cada forma.

- ✦ **Ese:** lo mismo es aplicable en este caso, aunque a la inversa: aumenta su proporción en el caso de los pronombres pero no en el de los determinantes. Sin embargo, tomando ambas categorías en conjunto, se incrementa su utilización en traducciones.
 - ✦ **Aquel:** es el único elemento que no varía cuando diferenciamos entre formas adjetivales y pronominales. En ambos casos encontramos una reducción del porcentaje de apariciones con respecto a las otras dos formas en el subcorpus de traducciones.
- El comportamiento de *aquel* es el que más se acerca a la categoría de elemento único, ya que la diferencia de frecuencias entre ES y TRAD no sólo es estadísticamente significativa, sino mayor a las que presentan los otros dos grados. Además, como acabamos de mencionar, es el único de los tres grados cuya proporción en TRAD disminuye con respecto a los otros dos grados tanto en su forma pronominal como determinante. Por otro lado, en el apartado de distribución por autores, habíamos observado que su uso parecía ser el más idiosincrático de las tres formas analizadas.

Este hecho no se aleja demasiado de nuestras primeras intuiciones sobre las frecuencias de uso de los demostrativos: parece indicar que el elemento que se tiende a dejar de lado al traducir *that* es *aquel*, pues la forma en la que, al parecer, piensan primero los traductores en textos literarios con un alto componente de oralidad es *ese*, aunque después, como hemos visto en el análisis de concordancias paralelas, la utilización de *aquel* no decrezca en exceso debido a que surge entre las soluciones traductorales por otros motivos o en contextos específicos en los que la costumbre lingüística o estilística así lo exige.

- El uso de los demostrativos en el subcorpus de originales ingleses es muy superior al que muestran los originales españoles. Sin embargo, la diferencia entre el componente en inglés y las traducciones al español es todavía mayor. Lo que esto indica es que el comportamiento de los traductores tiende más hacia la norma del español (normalización) que hacia las frecuencias del inglés (interferencia), y que, además, incluso supera dicha norma, en la dirección contraria a la sugerida por los textos originales ingleses (lo que provoca la infrarrepresentación de los demostrativos en traducciones).

- La explicación de lo indicado en el punto anterior la proporciona el análisis de las concordancias bilingües. Como cabía esperar, ni todos los demostrativos ingleses se traducen por demostrativos en español, ni todos los demostrativos españoles proceden de demostrativos ingleses. Ahora bien, los que aparecen en los textos traducidos sin tener como estímulo un demostrativo en inglés (adiciones) no compensan los que se «pierden» en la traducción (omisiones), que son muchos más. Dicho de otro modo, el número de DEM EN > NO DEM ES es superior al del caso opuesto.

Existen diversas diferencias contrastivas que podrían explicar este hecho. En primer lugar, es común en español la elisión del demostrativo en funciones en las que no aporta información imprescindible, como por ejemplo a principio de frase en función de sujeto. Hay funciones del demostrativo muy particulares del inglés, como su uso delante de sustantivos para presentar hechos o personajes, que no pueden trasladarse de la misma manera a nuestra lengua. Otros motivos pueden ser la mayor fuerza deíctica que presenta el demostrativo español en determinados contextos, en los que a menudo se sustituye por el artículo definido, o la amplia variedad de recursos gramaticales del español que lo hacen innecesario para identificar el elemento al que se pretende hacer referencia (por ejemplo, el pronombre personal acusativo).

- En cuanto a los adverbios demostrativos locativos, en conjunto se encuentran claramente sobrerrepresentados en TRAD. La hipótesis de los elementos únicos no se cumple para ninguno de los grados de lejanía de los adverbios, ya que *allí* y *allá* están sobrerrepresentados y en el caso de *ahí* la diferencia no es significativa.

A lo largo del análisis hemos ido exponiendo las diferentes hipótesis que podrían explicar este hecho pero, dado que el número de adverbios demostrativos en los originales ingleses es muy superior al de los originales españoles, se puede afirmar que la influencia de los originales ha sido un factor relevante en la sobrerrepresentación, al que se ha sumado la prominencia de este recurso en lengua meta —a pesar de que, como cabía esperar, por otra parte, no todos los adverbios demostrativos del inglés se han traducido por adverbios demostrativos en español—.

Si nos referimos otra vez a las características que menciona Hareide en la definición de *unique item*, podemos señalar que tampoco en este caso se cumplen en su totalidad:

- Has no direct counterpart in the source language: en ambos casos, determinantes y pronombres, por un lado, y adverbios, por otro, existe un grado con un equivalente directo inmediatamente identificable, que es el de cercanía (*this/these* para *este/a(s)* y *here* para *aquí/acá*) y dos grados en el que la situación no es tan evidente, pues corresponden a un único elemento en la lengua de partida (*that/those* para *ese/a(s)* y *aquello/a(s)*, y *there* para *ahí* y *allí/allá*). Por tanto, la falta de equivalente directo se da únicamente en estos últimos casos.
- Has no salient or prototypical source language structure that may impact its choice in the target language: en este caso podríamos decir que los demostrativos son una categoría muy frecuente en inglés y prototípica también, pues son un grupo cerrado que se utiliza para expresar una función específica, que es la de marcar relaciones deícticas o fóricas. Ainaud, Espunya y Pujol (2003: 180), cuando mencionan que los elementos que más destacan por su capacidad deíctica espacial, temporal y textual son los demostrativos determinantes y pronombres y los adverbios de lugar y de tiempo, no parecen hacer distinciones entre inglés y español. Por lo tanto, esta condición no se cumple, creemos, ni en el caso de los pronombres y determinantes ni en el caso de los adverbios, dado que en este último grupo de elementos la influencia del texto original afecta claramente a los resultados.
- It is characterized by being infrequent and non-prototypical: no podemos decir que los demostrativos en español no sean frecuentes: las frecuencias halladas en el corpus, mucho mayores a las de todos los otros elementos estudiados en este trabajo, lo demuestran. Con respecto a la prototipicidad, creemos que el caso es el mismo que con los demostrativos ingleses; de hecho la RAE (2009: 1269) define a los demostrativos como «los representantes más característicos del paradigma de las categorías deícticas», que incluye también los pronombres personales, los posesivos, los morfemas verbales de tiempo y persona, ciertos verbos de movimiento y algunos adjetivos locativos y temporales. Sin embargo,

sí es cierto que en el caso de *aquel*, las frecuencias de uso son inferiores en originales a las del otro deíctico de lejanía, *ese*, por lo que, en este sentido, sí podría considerarse que la primera forma no es prototípica dentro de la categoría más reducida de los pronombres y determinantes demostrativos, y eso hace que se encuentre todavía más infrarrepresentada que los otros dos grados en traducciones.

Con respecto a la Gravitational Pull Hypothesis de Halverson, observamos que existen dos resultados completamente opuestos: por un lado, los determinantes y pronombres han resultado infrarrepresentados; por otro, los adverbios han resultado sobrerrepresentados.

- ✓ El primer factor que influye en la posible sobre o infrarrepresentación de un elemento dado es, recordemos, la frecuencia de uso o prototipicidad de dicho elemento en la lengua meta. Es probable que un elemento más destacado de la red semántica del traductor, debido a su mayor frecuencia, convencionalidad, familiaridad o estatus en nuestra lengua (Giora 1997: 185), se active antes que otros al traducir. Esto explicaría el mayor uso proporcional que se da de *ese* en detrimento de *aquel*, una diferencia que es mucho más marcada en el caso de las traducciones.

En el caso de los determinantes y pronombres, esta prototipicidad no parece ejercer fuerza suficiente en el equilibrio entre factores (dado que todos los elementos se encuentran infrarrepresentados), mientras que sí lo hace, aparentemente, en el caso de los adverbios. Esto puede significar que los primeros son elementos que pueden omitirse o sustituirse más fácilmente en español, aunque aparezcan en el original (sobre todo, como hemos visto, los de cercanía), mientras que los segundos son en cierta manera elementos comodines que pueden encajar en diferentes contextos y la influencia del original los hace aparecer aún en mayor medida.

- ✓ Lo que acabamos de decir enlaza con el factor 2, que es la frecuencia de uso o prototipicidad del elemento de la lengua de partida. Aquí ocurre lo mismo: no es suficiente para garantizar una representación equivalente a la de los originales en el caso de los determinantes y pronombres, probablemente porque, como decíamos cuando hablábamos del primer factor, los demostrativos cuentan con

una amplia variedad de elementos en español que, quizás por su menor fuerza deíctica, pueden ser útiles o más adecuados en muchas circunstancias en las que en inglés se utilizaría el demostrativo. Sin embargo, sí es suficiente en el caso de los adverbios y, de hecho, la gran cantidad de adverbios en inglés, debida, quizás, a la mayor propensión de este idioma a expresar matices de dirección y trayectoria, como hemos explicado, influye sin duda en la mayor utilización de sus equivalentes españoles en las traducciones.

- ✓ El factor 3, por último, se refiere a los patrones de conexión presentes en el lexicón mental del traductor. En los adverbios observamos que se traducen por el elemento equivalente más inmediato (p. ej. *here* por *aquí*) aproximadamente en el 48 % de los casos, mientras que en determinantes y pronombres, esto sucede, en promedio, en un 39 % de las ocasiones (p. ej. *that* por *ese* o *aquel*), lo que podría indicar que la conexión entre adverbio inglés y adverbio español equivalente está un poco más clara que en el caso de estos últimos elementos.

Un estudio que apoya nuestras conclusiones es el de Labrador (2010). Esta autora partía de una hipótesis de trabajo opuesta, en cierta medida, a la nuestra: que en traducciones se usarían más los demostrativos por influencia del inglés en detrimento de otras opciones más idiomáticas. En su trabajo confirmó la hipótesis solo en parte, pues había formas que se encontraban sobrerrepresentadas y otras que estaban infrarrepresentadas; aunque en más de la mitad de los casos no había diferencias de uso significativas. Sin embargo, la tendencia general era hacia el menor uso en traducciones. Ella no estudiaba los adverbios demostrativos, sobre los cuales hemos comprobado que existe una sobrerrepresentación general en traducciones, pero en el caso de pronombres y determinantes, la tendencia observada en conjunto en nuestro corpus es, como en el estudio de Labrador, hacia la infrarrepresentación, y mucho más marcada que en su caso. De todos modos, hay que mencionar que en el corpus que ella utiliza se incluyen otros géneros aparte del narrativo, lo que puede ser determinante en la diferencia de resultados.

En definitiva, podemos decir que, aunque la solución de traducción principal para los demostrativos es la utilización del demostrativo equivalente en la lengua meta, en este estudio nos hemos encontrado ante toda una gama de soluciones de traducción que dependen mucho de la situación y también de diversos factores, incluidas las

preferencias personales de los traductores. En el caso de los pronombres y determinantes, los traductores tienen muy en cuenta el contexto y las normas lingüísticas y estilísticas del español y no «sobreutilizan» este recurso, a diferencia de lo que ocurre con los adverbios demostrativos. Sin embargo, la falta, en inglés, de una categorización paralela a la estructura tripartita del español hace que el tercer grado de lejanía de pronombres y determinantes no se utilice en traducciones con la frecuencia y distribución que cabría esperar si habláramos de un texto escrito originalmente en español.

Según lo que hemos ido observando en este trabajo, la utilización de los distintos demostrativos se ve más afectada, al parecer, por preferencias estilísticas que gramaticales, lo que no resulta extraño, puesto que se trata de elementos que pertenecen al ámbito de la deixis, que, después de todo, es un fenómeno que depende de la perspectiva de la situación o del mundo que adopte cada autor en un momento determinado. Cambiar estos elementos puede suponer mantener o, por el contrario, alterar aspectos importantes de la construcción del universo literario que pretende ofrecernos el autor de la obra en cuestión. Como explica Richardson (1998: 125-127):

One consequence of this [translation] process is that the deictic perspective which pervades the TL text must be structured in such a way that is deemed coherent by the TL reader. (...) decisions about how much adaptation is needed, if any, will depend on a variety of factors, such as the purpose of the translation, the intended readership or the characteristics of the text being translated. In a literary translation, there will often be a tendency to avoid such contextualisation (...) Such a translation, more literal and less ‘communicative’ (in the sense of being adapted to the TL reader’s perspective) may be felt to convey something essential about the author’s construal of his or her message and about the uniqueness of the ‘world-view’ which is being communicated.

De todos modos, es necesario recordar, como comenta Álvarez de Mon y Rego (2006: 35) que cada texto representa un estilo y un idiolecto y refleja las preferencias estilísticas o la inclinación por un recurso en concreto del autor del mismo. Por este motivo, es difícil hacer generalizaciones sobre cuestiones como esta que nos ocupa, que puede tener mucho que ver con el estilo personal de cada escritor o traductor.

6. Sufijos apreciativos

6.1. Sufijos apreciativos: definición y caracterización

En su obra *Procedimientos de formación de palabras en español*, Almela Pérez (1999: 18) distingue entre dos procedimientos para la ampliación del repertorio léxico de un sistema lingüístico: la formación de nuevos vocablos a través de los mecanismos propios de dicho sistema y la incorporación al mismo de términos de otras lenguas. Dentro del primero de estos dos procedimientos es donde se ubica el objeto de este estudio, la sufijación apreciativa.

Como deja claro el autor (1999: 25), la clasificación de los distintos tipos de procesos de formación de palabras es una cuestión que ha sido y sigue siendo muy debatida entre los estudiosos del tema. Sin embargo, la mayoría menciona tres procedimientos principales: sufijación, prefijación y composición, que a veces se recogen bajo el paraguas de *derivación*, un término que, según Almela Pérez, es de todo menos preciso, pues puede significar cosas distintas dependiendo de a quién consultemos. Para unos se opone a flexión, para otros a composición; para ciertos autores equivale a sufijación, o bien al conjunto de sufijación, prefijación y composición; también puede significar simplemente que «una palabra procede de otra por evolución» (Almela Pérez 1999: 34-35).

Por ejemplo, Gómez Torrego (1997: 18) considera que las palabras derivadas son aquellas que añaden a la raíz un sufijo —las que llevan prefijo se denominan prefijadas—. La *Nueva gramática de la lengua española* (RAE 2009: 23) divide la morfología léxica en las subdisciplinas de composición (unión de dos bases léxicas) y derivación (unión de una base léxica y un afijo), que incluye no solo la sufijación, sino también la prefijación y la parasíntesis (una mezcla de composición y derivación o definida también en algunos casos como la adición simultánea de un prefijo y un sufijo). La propuesta de Almela consiste en la división en cuatro tipos: adición (que incluiría los procedimientos de prefijación, sufijación, interfijación, composición y parasíntesis),

modificación (conversión, sustitución, suplección y repetición), sustracción (regresión y abreviación) y combinación (siglación y acronimia). En su opinión, la sufijación es el más productivo, el que dispone de más unidades (los sufijos) y el más complejo de todos estos procedimientos.

La sufijación consiste en la adición de un afijo pospuesto a una base léxica¹⁸⁶ para modificar su significado en mayor o menor medida. Normalmente se distingue entre sufijación apreciativa y sufijación no apreciativa. La primera (también denominada *potestativa* por Gómez Torrego¹⁸⁷) es la que incluye los sufijos diminutivos, aumentativos y peyorativos o despectivos, que son precisamente los que nos interesan en este estudio. Estos sufijos denotan «una idea cuantitativa (cantidad, intensidad o tamaño) de lo significado por la raíz, o una actitud afectiva del que habla ante el objeto significado» (Seco 1989: 207), sin afectar sustancialmente el significado de la palabra ni alterar su categoría gramatical. La sufijación no apreciativa (llamada también *significativa* por Seco) engloba todos los demás sufijos, aunque, como explica Lang (1990: 20), muchas de las funciones de estos últimos se solapan con las de los sufijos apreciativos.

Algunos autores realizan este agrupamiento basándose en motivos nocionales o conceptuales (como Gómez Torrego 1997: 20-21, o como acabamos de ver en la definición de Seco); otros, en su comportamiento lingüístico, como Almela Pérez (1999:

¹⁸⁶ La base léxica es la voz de la que se parte en un proceso morfológico (RAE 2009: 23). Se distingue de la raíz, que es el segmento morfológico que aporta el significado léxico. Siguiendo el ejemplo proporcionado por la RAE, la palabra *arenoso* se formaría a partir de la adición del sufijo *-oso* a la raíz de *arena*, que surgiría de la cancelación de la vocal átona final de la base en dicha palabra. Manuel Seco (1989: 203-206) distingue entre la base, «que es la parte “fija” de una palabra y que contiene la significación general de esta» y la raíz, que, junto a los afijos, forman la primera. Uno de los ejemplos que pone el autor es el de *cantar* y *canturrear*: en *cantar*, *cant-* actúa como base, pero en *canturrear*, solo como parte de la base (la raíz); el fonema añadido, *urre*, modifica y concreta la idea general de la raíz y forma la base de *canturrear*, a la que posteriormente se añade el formante *ar* (también llamado desinencia o morfema flexivo). Lang (1990: 13) también afirma: «(...)it is apparent that the true source of the derivation is the base lexeme (*nadar, gordo, casa, espeso*) and not the lexical morpheme itself». Considera que existen distintos tipos de base (simples, compuestas o derivadas) pero «Both simple and derived bases are reduced to stems where required in order to enter into a complex form»; por ejemplo, en el caso de la palabra derivada *regular*, la base de la que surgiría sería *regular*, pero la raíz a partir de la que empezaría la derivación, sería *regul-*.

¹⁸⁷ Gómez Torrego (1997: 21) distingue entre sufijos de significado gramatical (los de infinitivo, gerundio y participio) y de significado no gramatical, que, a su vez, se dividen en obligatorios si aportan un significado nuevo (pueden cambiar la categoría de la palabra, como es el caso de *ajustar* > *ajuste*) y potestativos si solo aportan a la raíz un matiz semántico afectivo o expresivo (en este caso, no cambian la categoría de la palabra). Estos últimos son los denominados *apreciativos* por otros autores y Gómez Torrego añade dos clases a las tres que estudiaremos nosotros: superlativos y familiares.

Algunos autores incluyen los superlativos *-ísimo* y *-érrimo* entre los sufijos apreciativos. En este estudio no los hemos incluido, pero sería interesante analizarlos en el futuro. Con respecto a los familiares, afirma Santos Muñoz (2006: 21) lo siguiente: «son aquellos que encierran significados de ironía, desprecio, etc., pero tienen un marcado carácter familiar, y las terminaciones que presentan no corresponden con los aumentativos y/o diminutivos, sino que forman palabras del tipo: *vivales, guaperas, finolis* (Gómez Torrego, 1997), donde en ningún momento puede interpretarse la expresión de la cuantificación, sino más bien de la cualidad».

98). Este autor (1999: 102-103) atribuye a los apreciativos cuatro características fundamentales: homocategorización (normalmente mantienen la categoría de la base), riqueza alomórfica (poseen más variantes que los demás), oscilación entre diptongación y monoptongación (a veces mantienen el diptongo de la base; otras lo monoptongan), determinación semántica de la base (unas veces toman el significado y el género de la base; otras, toman el significado pero imponen su género) y ordenación (suelen colocarse detrás de los sufijos no apreciativos).

Para Lang (1990: 97) no está tan clara la frontera entre sufijos apreciativos y no apreciativos, puesto que en casos como los de *empujón* o *quejica*, por ejemplo, hay un cambio de categoría gramatical, de verbo a adjetivo o sustantivo. Por lo tanto, salvo que se considere que el *-ón* de *empujón* es un sufijo distinto al de *fortunón*, afirma este autor que existiría un solapamiento de funciones entre ambas clases. LA RAE sin embargo, tiene más clara esta cuestión. Si hay un cambio de categoría gramatical, como veremos más adelante, la palabra no se considera fruto de un procedimiento de sufijación apreciativa.

Por las limitaciones de este estudio y la naturaleza de nuestros objetivos, no nos interesa adentrarnos en la descripción de las reglas de formación de palabras mediante la adición de sufijos en español, que Almela Pérez (1999) sintetiza en las páginas 90 a 98 de su obra, utilizando el término *reglas* en un sentido muy laxo, pues están sujetas a multitud de excepciones y, como bien aclara este autor, dependen, en última instancia, de la voluntad del hablante. Gooch (1970) menciona también en las páginas 31 a 34 de su obra las excepciones a algunas de estas reglas de formación. Lang, por su parte (1990: 42-62) resume distintas teorías que han intentado explicar los diversos procedimientos de formación de palabras. Sin embargo, considera que

(...) it is apparent that lexical morphology theory falls seriously short in coping with the many peculiarities of formation in this area [of diminution, augmentation, and pejoration], especially in explaining the often arbitrary application of allomorphy and in the overlaps and unpredictability of the pejorative suffix group» (1990: 206).

Así, una misma palabra puede tener distintas formas diminutivas o aumentativas: en determinadas regiones se preferirá la forma *solecito*, mientras que en otras podremos oír

más frecuentemente la variante *solcito* y, muy a menudo, será complicado explicar la razón de una u otra preferencia.

Existen otros debates abiertos entre los lingüistas, como el referido a la cuestión de si los sufijos apreciativos deben considerarse morfemas flexivos o derivativos. Gómez Torrego (1997: 16) define los morfemas flexivos —o desinenciales— como «terminaciones de una palabra que indican los significados gramaticales de género, número, persona, tiempo, modo y aspecto», y los morfemas derivativos —o afijales— como «ciertos segmentos que no son desinencias, que preceden o siguen a una raíz o morfema léxico y que no tienen autonomía fuera de la palabra». Lang (1990: 16) resume las diferencias entre los procesos flexivos y derivativos en cuatro puntos: la derivación puede cambiar la categoría gramatical de la base, mientras que la flexión, no; los morfemas derivativos van antes que los flexivos en la palabra; varios morfemas derivativos pueden encadenarse en una misma palabra, cosa que no sucede con los flexivos, y la derivación, a diferencia de la flexión, no es completamente productiva; es decir, que su paradigma derivativo no es completo¹⁸⁸. Zacarías Ponce de León (2008: 224), por su parte, los considera una categoría intermedia entre la derivación y la flexión, dado que, en su opinión, comparten rasgos con ambas categorías y, a la vez, poseen características y comportamientos lo suficientemente diferenciados para constituir un grupo aparte¹⁸⁹.

En la gramática de la Real Academia (2009: 627) se afirma que los sufijos apreciativos se añaden a sustantivos, adjetivos y otras clases de palabras (algunos adverbios e incluso numerales y otros cuantificadores como *todo* y *poco*) «para expresar tamaño, atenuación, encarecimiento, cercanía, ponderación, cortesía, ironía, menosprecio y otras nociones —no siempre deslindables con facilidad— que caracterizan la valoración afectiva que se hace de las personas, los animales o las cosas». A pesar de atribuirles toda esta variedad de valores, la gramática separa también dichos

¹⁸⁸ El ejemplo que proporciona el autor para explicar esto último es que no todos los adjetivos cualitativos bislabos se nominalizan mediante el mismo sufijo (*pobre* > *pobresa*, *grande* > *grandeza*, pero *lento* > *lentitud*), mientras que, por el contrario, todos los verbos de la tercera conjugación sin excepción forman el plural de indicativo de primera persona en *-imos*) (1990: 17).

¹⁸⁹ «Si bien es cierto que la mitad de las propiedades de los apreciativos son derivativas, no hay que desdeñar las características flexivas que presentan y que son básicas en la caracterización de esta categoría. Parece evidente que los apreciativos presentan comportamientos muy particulares, como la gran iteratividad (*chiquitititito*), su muy alta productividad y su posición, cercana a la flexión, después de cualquier otro morfema derivativo. Estas características permiten afirmar que los apreciativos, más que derivativos o flexivos, son una categoría fronteriza entre estos dos grandes procesos morfológicos» (Zacarías Ponce de León 2008: 234).

sufijos en las tres clases tradicionales. Sin embargo, advierte Lázaro Mora (1999: 4648) que «esta repartición es aproximativa, porque los límites de tales grupos son poco nítidos a veces. Por un lado, los sufijos peyorativos también implican tamaño: *pajarucho* y *mujeruca* están coloreados, en efecto, con desestima; pero un *pajarucho* es necesariamente grande, mientras que una *mujeruca* deberá tener una estatura limitada. Por otro lado, los diminutivos y aumentativos no siempre aminoran o agrandan: junto a la idea de lo pequeño suelen asociarse connotaciones afectivas positivas, y, a la de lo grande, negativas». Matizaremos esta observación a continuación.

6.1.1. *Los diminutivos*

Se trata del grupo más estudiado en la bibliografía a la que hemos podido tener acceso, quizás porque es también el más productivo. Rojo (2009: 230-231) se basa en el inventario de Erichsen (2007a) para resumir de la siguiente manera los usos más comunes del diminutivo en español: indicar que algo es pequeño y compacto (*cucharilla, gatito*); expresar afecto o simpatía (*abuelita, maridito*); indicar que algo es casual o temporal (*rollete, amiguete*); restar importancia a algo, especialmente a cosas que nos resultan desagradables o incómodas (*necesitar un dinerillo, estar rellenito*); expresar cuantificación de grado en ciertas expresiones adverbiales (*prontito, poquito*); expresar molestia o descontento con respecto a cosas que suceden durante un período específico de tiempo (*¡vaya una semanita!*); dotar a nuestras palabras de un tono familiar y amable (*Un momentito, por favor. ¡Despacito!*); en el lenguaje dirigido a niños o en el de los enamorados (*pillín, cielito*); formar nuevas palabras, a veces sin conexión con el significado del término original (*libreta*).

Alonso (1982: 163) señala, respecto a la significación disminuidora de este grupo, que «es con mucho la función menos frecuente», y menciona como más importantes las funciones emocional, representacional y activa. Afirma también que «cuando el sentido central es realmente el de disminución, se suele insistir en la idea de pequeñez con otros recursos (*una cajita pequeña, una cosita de nada, etc.*)».

Gooch (1970: 2) también opina que no es la función principal de los diminutivos la de transmitir idea de pequeñez, sino la de expresar afecto o conseguir una reacción favorable en nuestro interlocutor. Con respecto a esto último, Gómez Torrego (1997:

21-22) afirma que pueden ser apelativos o conativos cuando intentan mover la voluntad del oyente, como cuando decimos «momentito» con el objeto de que este no se impacienta. Gooch también menciona el uso de diminutivos por parte de los niños al hablar de cosas o personas pequeñas y atractivas y, por extensión, por parte de los mayores, para rebajar la importancia, gravedad o magnitud de alguna cosa

in cases where they wish to reduce something to proportions which are for them *emotionally* manageable. Here the diminutive suffix gives an insight into what is a deep-rooted human characteristic: fear of things that are big and a concomitant desire to reduce them to friendly proportions (Gooch 1970: 3).

Considera que son polifacéticos y que es el contexto el que dará la clave para interpretar el significado que se quiere transmitir, como se desprende del siguiente fragmento de la obra de Alonso (1982: 186), que también cita Gooch:

Parece como si la constante en el diminutivo fuera ese destacar la representación del objeto (...) Luego, la situación, los consabidos y el contexto dan los *indicios* de cuál es el motivo de esa atención privilegiada, si la ternura o el desamor por el objeto, si el acercamiento o el apartamiento de él, si la complacencia o la displicencia, si el saboreo o el disgusto, si la insistencia enfática en el objeto o la detención en el interés con que lo vemos.

Lang (1990: 102), por su parte, afirma que todos los diminutivos —aunque algunos más que otros— pueden ser utilizados también de manera peyorativa, puesto que dicha idea está implícita en la noción de pequeñez y es atenuada, a la vez, por esta. La RAE (2009: 652-653) también considera que el diminutivo, además de expresar cortesía o matices afectivos como familiaridad o cercanía, puede aportar connotaciones negativas de ironía o menosprecio. Proporcionan el ejemplo de algunos diminutivos derivados de sustantivos de profesiones, como *abogadito* o *politiquillo* y también de otros sustantivos de persona que pueden adquirir dichas connotaciones a través del contexto, como *soldadito* o *ancianito*. Se menciona además el uso eufemístico de algunos diminutivos en el lenguaje publicitario o en el dirigido a los niños (*la cosita*, *el culito*) y también la función atenuadora de determinadas cualidades, como en *rojito* (un poco rojo) o *pesadito* (un tanto pesado). Por otra parte, también parece utilizarse como intensificador en casos como *loquito por ella* o *coche nuevecito* (completamente loco o

nuevo), aunque se aclara que la interpretación puede venir dada más por el significado de la base léxica que por el sufijo en sí (RAE 2009: 654).

6.1.2. *Los aumentativos*

Según Gómez Torrego (1997: 22) «los aumentativos aportan valores de aumento, de admiración o de engrandecimiento, acompañados o no de otros como rechazo, desprecio, cariño, etc.». Lang (1990: 109) afirma que los aumentativos se solapan especialmente con los peyorativos, pues la idea de tamaño grande suele combinarse con la de fealdad, pero, dado que un tamaño grande puede resultar también una característica positiva, en algunos casos se da un efecto meliorativo (como en *¡qué piscinaza!*). Prieto (2005: 62) comenta lo siguiente al respecto:

«Some have argued that pejorative connotations may be inherent in augmentative functions. (...) Admittedly, as Pharies (in personal communication) suggests, anything that is smaller or bigger than the norm can inspire negative or pejorative connotations because it is too small or too big. Thus, if bigness is fine or positively viewed, then the AUG would not create pejorative connotations (and if smallness is positive, then DIMs would not produce pejoration). For some, what is extraordinarily big may be what is desired and smallness may be undesired or funny; in which case, DIMs would very likely become pejoratives. AUG -azo in modern Spanish, for example, may mean “great” (car+azo).

Como hemos mencionado, su productividad es menor comparada con la de los diminutivos, y sus alternativas sintácticas, «...adjectives such as *grande*, *enorme*, *estupendo*, *impresionante* - are much more likely to be resorted to, especially in literary style, whereas the augmentatives are likely to be preferred in colloquial speech» (Lang 1990: 110).

Según Rojo (2009: 232-233), nuevamente basándose en un artículo de Erichsen (2007b), las principales funciones de los aumentativos son las siguientes: indicar que algo es grande, alto, potente o poderoso, entre otras cualidades asociadas con el tamaño (*perrazo*, *casona*, *muchachote*); aportar mayor intensidad a las cualidades de los objetos, dando lugar a matices negativos o positivos (adjetivos como *solterona* o *cabezón*

expresan desagrado o desprecio, palabras como *pisazo* o *artistazo* sugieren aprobación o admiración, mientras que las connotaciones de términos como *papelón* o *pasote* dependerán del contexto en que se utilicen); indicar golpe con el instrumento o sobre la parte del cuerpo que designa la palabra a la que se añaden (*codazo*, *plumazo*, *pisotón*); y, como en el caso de los diminutivos, formar nuevas palabras lexicalizadas (*tablón*, *lamparón*).

6.1.3. *Los peyorativos*

El repertorio de peyorativos es mucho más extenso que el de los dos grupos anteriores, pero la productividad de cada sufijo, menor. Normalmente expresan lo desagradable o ridículo del objeto al que designan y suelen clasificarse según los matices diminutivos, aumentativos o humorísticos que presentan (Lang 1990: 117). Soler Espiauba (1996) comenta que «Pueden vehicular sentimientos de desamor, hostilidad, odio, desprecio, pero la mayoría de las veces son desvalorativos de dominante intelectual, con escasa emoción». Gómez Torrego, por su parte, afirma que el matiz de desprecio puede ir acompañado a veces de afecto (1997: 22). En los inventarios de algunos autores se incluyen sufijos no apreciativos que, unidos a determinadas bases, pueden funcionar como apreciativos y también sufijos apreciativos de los dos otros grupos que son capaces de aportar matices peyorativos.

Rojo (2009: 234-235) resume sus funciones principales de la siguiente manera: indicar que se comparten en cierta manera las cualidades o comportamientos negativos típicos de la palabra a la que se añaden (*caballuno*, *diablejo*); indicar que algo se hace de manera rudimentaria o sin pericia (*pintarrajo*, *palote*); señalar a alguien o algo caracterizado por las características negativas inherentes a la palabra a la que se añaden (*ceporro*, *bestiales*); burlarse de alguien sin más (*Perico*, *Pericales*); señalar a alguien que realiza alguna acción objetable en exceso (*acusica*, *llorica*); hacer alusión a las condiciones deficientes de algo, por ejemplo, porque es pobre, viejo, feo o no funciona correctamente (*poblacho*, *cacharro*); sugerir que algo o alguien posee la cualidad descrita por la raíz solo hasta cierto punto (*feúcho*, *negruzco*); indicar falta de importancia o estatus secundario (*escritorzuelo*, *populacho*); expresar que algo es en cierto modo «extrínseco» (*padraastro*, *madrastro*); intensificar las cualidades negativas

de la raíz (*brutote, bicharraco*); y expresar desprecio degradando la categoría de la raíz (*pringadete, listillo*).

6.1.4. *Los sufijos apreciativos como elementos únicos*

A continuación, presentaremos los argumentos a favor de considerar los sufijos apreciativos un elemento único del español si lo comparamos con el inglés. Lang (1990: 33) afirma, al respecto de esta cuestión:

Many observers have commented on the flexibility of Spanish derivational morphology, especially on its strong tendency towards synthetic affix structure as opposed to analytical or compound structure. In the words of Rivadó, an early commentator on neologisms in the Academy Dictionary, “Nuestra lengua posee un vigor derivativo poderoso... Las lenguas romances son más ricas en derivaciones que las otras y la castellana es la más rica entre sus hermanas en este sentido”. Similarly, Manuel Seco speaks of “la vitalidad y fecundidad que la sufijación y la composición conservan en nuestra lengua”.

También comenta Lang, más adelante (1990: 207), lo siguiente:

Firstly, the continued exuberance of emotive suffixation is very striking. Not only the vigour of the obviously great favoured suffixes such as *-ito*, *-illo*, or *-azo*, but clear signs of revival of the sometimes classified as stagnant or even archaic, namely *-ico*, *-ete*, or *-ín*). This vitality of the emotive suffix system, reflected not only in their extensive use but in the very multiplicity of morphemes involved, must be counted as a very distinctive feature, since it dominates the whole derivational system and distinguishes Spanish from sister Romance languages such as Italian and Portuguese, where the emotive suffixes are by no means insignificant in their expressiveness and popularity.¹⁹⁰

¹⁹⁰ Otros autores son de la misma opinión. Santibáñez Sáenz (1999: 173) afirma, por ejemplo: «If compared with other European languages (such as English, French or German), Spanish enjoys a relatively rich and productive set of diminutive suffixes, *-ito* and *-illo* being the most frequently employed». Según Soler Espiauba (1996), «Una de las oposiciones más contrastadas entre lenguas germánicas y lenguas románicas es la mayor flexibilidad de estas últimas en la formación de nuevas palabras mediante una sufijación rica en matices, constituida por los llamados *apreciativos*». Lo expresa de manera algo más efusiva Cantera Ortiz de Urbina: «En repetidas ocasiones, sobre todo al tratar de la problemática de la traducción, hemos traído a colación nuestras expresiones agüita fresca y fresquita el agua. Unas expresiones que cobraban especial fuerza expresiva cuando eran exclamadas por una aguadora o un

A continuación, este autor (1990: 34) proporciona algunos ejemplos con el objetivo de comparar los mecanismos de formación de palabras del español con los del inglés y el francés y concluye que, en general, en la formación de familias de palabras, el español prefiere generar nuevos lexemas a partir de una única base mientras que el inglés se vale de etimologías distintas o palabras compuestas (*puño* > *puñal*, *empuñadura* frente a *fist* > *dagger*, *sword hilt*), y el francés designa los mismos conceptos mediante estructuras de tipo «sustantivo + preposición + sustantivo» (*comer* > *comedor*, frente a *manger* > *salle à manger*).

Royo (2009: 197), por su parte, es más contundente:

One of the most obvious morphological divergences between English and Spanish is, for instance, the Spanish use of suffixes that point to a reduction or an increase in the size of things: *piso* (flat) v. *pisito* (small flat) v. *pisazo* (big flat). These suffixes are also used to express affection or/and admiration, as in *abuela* (grandmother) v. *abuelita* (granny) v. *abuelaza* (great granny).

Esta autora (2009: 229) explica que tanto el inglés como el español poseen morfología diminutiva, aunque los sufijos que encontramos en inglés (-*ey*, -*y*, -*ie*, -*o*, -*ette* o *let*) tienen una productividad bastante limitada en comparación con la de los sufijos españoles. Suelen utilizarse en nombre propios (Peter > Petey), para indicar tamaño pequeño o corta edad, como en *piglet*, en el vocabulario infantil (*potty*) y como marca de afecto o confianza (*underwear* > *undies*). Por otro lado, el inglés, a diferencia del español, no cuenta con morfología aumentativa ni peyorativa, por lo que tiene que recurrir a otros medios para expresar estos valores (Royo 2009: 233).

Royo nos proporciona un interesante inventario de estrategias para conseguirlo cuando no existen sufijos equivalentes en inglés. Algunas de las que menciona son la omisión del elemento sufijado con compensación del tono en el contexto circundante; el uso de sintagmas preposicionales (*prontito* > *in just a few days*); la adición de adjetivos (*manita* > *tiny hand*) o adverbios (*recibían un bofetón* > *got their ears soundly boxed*); la separación de la información en más de un elemento, como en *bigotudos* > *mustached*

aguador de aquéllos que en el siglo XIX tanto llamaban la atención de los viajeros extranjeros (...) Por más que nos esforcemos en buscar correspondencia a esas expresiones en otras lenguas, sólo lo conseguiremos en algunas como el catalán, el gallego, el provenzal y el griego; pero fracasaremos en la mayoría de las demás lenguas, incluso en algunas de muy gran riqueza como el francés» (2012: 97).

critters, en donde el adjetivo transmite el significado de la raíz y el sustantivo de connotación peyorativa aporta la información del sufijo; la utilización de un término de la lengua meta que contiene ya la información evaluativa en su raíz, como en *golpecitos* > *tapping* (que significa tamborilear, dar golpecitos suaves y rápidos); la utilización de un equivalente funcional que aporte naturalidad estilística aunque el significado cambie ligeramente (*impropia para señoritas* > *not suitable for minors*); o la traducción por compuestos funcionalmente equivalentes (*flaquita* > *sweetheart*) (Rojo 2009: 236-241).

De la misma manera, Schneider (2003: 2) argumenta que el término *diminutivo* no debería considerarse una categoría morfológica —en la que los diminutivos prototípicos son los sufijos o los sustantivos sufijados—, sino semántica, pues se refiere a todas las posibles formas de expresar disminución. Cada lengua utilizará los medios de los que dispone para hacerlo y la sufijación es solo uno de ellos. Así, en inglés encontraremos tanto formas sintéticas (menciona los sufijos *-ie*, *-ette*, *-let*, *-kin*, *-een*, *-s*, *-er*, *-poo* y *-pegs*, que serían los únicos equivalentes formales directos de los sufijos españoles) como analíticas (por ejemplo, la utilización de adjetivos del campo semántico de lo pequeño, como *small* y *little* antes del sustantivo). Asimismo, el autor explica que en inglés no existen aumentativos sintéticos¹⁹¹, pero eso no significa que no existan aumentativos en esta lengua, pues su significado puede expresarse a través de otros recursos formales como estructuras de tipo «*large/big* + sustantivo», adjetivos que transmitan connotaciones peyorativas, prefijos como *maxi-*, *macro-* o *mega-*, o «inherent augmentatives», como denomina el autor a términos cuyo significado aumentativo no tiene relación con la morfología. Lo mismo sucede con los diminutivos: por ejemplo, *hill* sería un «inherent diminutive» en contraposición a *mountain* (2003: 16-19). Aun así, como hemos dicho anteriormente, los sufijos apreciativos expresan muchos más matices que únicamente el aumento o la disminución y, de hecho, si solo es ese el significado que quiere transmitirse, lo más probable es que en español también se utilicen las fórmulas mencionadas por Schneider.

Por otra parte, en español también existen otros mecanismos para expresar los matices que transmiten los sufijos apreciativos, algunos más próximos a aquellos de los que dispone el inglés, que pueden hacer que la sufijación apreciativa pase a un segundo

¹⁹¹ Cruzado (1982: 70) también afirma que no existen los sufijos aumentativos en inglés. Con respecto a los sufijos apreciativos en general, esta autora comenta: «I have always been fascinated by the emotive quality of these suffixes in Spanish, which is sometimes lost when going from Spanish into English because the Spanish suffix must usually be translated by adjectives and adverbs that, if not carefully chosen, do carry less than the full emotional impact projected in the suffixed Spanish word» (1982: 67).

plano en la traducción. Gooch (1970: 15-16) menciona, por ejemplo, los adjetivos y adverbios *pequeño, reducido, diminuto, minúsculo, un poco, un tanto, algo*, como elementos paralelos a los diminutivos; *grande, inmenso, gigantesco, ingente, descomunal, desmedido, desmesurado, desorbitado, extremado, muy, completamente, sumamente, totalmente* y el superlativo *-ísimo*, como equivalentes de los aumentativos y *malo, miserable, atroz, espantoso, horrible, horrendo, infame, nefando, desagradablemente, despreciablemente y miserablemente* en el caso de los peyorativos. Y esto únicamente refiriéndose al lenguaje formal o literario, con lo que puede suponerse que en el lenguaje popular habrá bastantes más alternativas. No obstante, menciona el hecho de que es frecuente hallar una combinación de elementos, en la que un adjetivo de los mencionados refuerza un sufijo (p. ej. «un duendecillo minúsculo»). También habla de otras combinaciones, como los adjetivos *estupendo, imponente, impresionante, soberbio, bárbaro, bestial y brutal*, el adverbio *algo* seguido de un diminutivo o expresiones como *de padre y señor mío, de miedo, de mil pares de demonios, de tomo y lomo, de mala muerte* o *de tres al cuarto*. Sin embargo, considera todas estas expresiones tautológicas, pues, a pesar de enfatizar el sentido del sufijo, vienen a repetir la función que este ya cumple. Cantera Ortiz de Urbina (2012: 102) habla también de las marcas prefijadas *pedazo de* o *cacho de*, que constituyen para él una especie de aumentativo que puede expresar significados negativos (*pedazo de alcornoque*: una persona muy necia) pero también positivos (*cacho de pan*: una persona muy buena).

Por lo tanto, muchos de los recursos mencionados para expresar en inglés los valores de los apreciativos están también disponibles en español. Eso significa que los sufijos apreciativos españoles son solo uno de los recursos al alcance del traductor para transmitir el sentido diminutivo, aumentativo o peyorativo expresado por el texto original. Nuestro objetivo será averiguar si la lengua de los textos traducidos de nuestro corpus se asemeja a la de los originales en relación con la frecuencia de los sufijos apreciativos utilizados o, por el contrario, la falta de recursos formales equivalentes en el original ocasiona que su número disminuya.

6.2. Corpus utilizado y selección de sufijos

Es una opinión generalizada que los diminutivos, aumentativos y despectivos son propios del lenguaje oral y coloquial (Santos Muñoz 2006: 14)¹⁹² y, por tanto, podría pensarse que no será fácil encontrarlos en textos escritos. Desde luego, no suele ser habitual utilizarlos en textos o traducciones técnicas u otro tipo de géneros especializados. Sin embargo, es evidente que los textos literarios son una excepción¹⁹³, puesto que constituyen un género en el que no es infrecuente hallar este tipo de lenguaje, especialmente en los diálogos. Gooch (1970: 14-15) considera que «the most fertile sources of written examples are to be found in those novelists and playwrights who depict scenes and describe situations of ordinary Spanish life». Las novelas policiales de nuestro corpus no describen hechos habituales de la vida cotidiana, pero sí transcurren en ambientes populares y contienen una buena cantidad de diálogos entre personajes de lo más variopintos.

Lang (1990: 41), por otra parte, considera que mecanismos como la derivación son fundamentales en el lenguaje coloquial espontáneo y comenta: «All this is confirmed by an analysis of any present-day popular writer such as Cela, Vázquez-Montalbán, Marsé, etc., who give their protagonists a prodigious derivational inventiveness as part of their individual characterisation». Recordemos que Vázquez-Montalbán es uno de los autores incluidos en nuestro corpus. Por tanto, consideramos que la sufijación apreciativa puede ser considerada perfectamente una característica de los textos de nuestro corpus y, como tal, esperamos verla representada en los textos originales. La incógnita es si lo estará de la misma manera en las traducciones, puesto que, como hemos comentado, el inglés suele valerse de recursos léxicos distintos para expresar estos matices cuantitativos y afectivos.

¹⁹² Dice esta autora al respecto: «el hecho de que este tipo de sufijación aparezca habitualmente en la lengua oral, en un registro muy coloquial, presentando infinidad de posibilidades expresivas en función de diversos y complejos factores como la intención del hablante, etc., dificulta la realización de una descripción exhaustiva de las opciones expresivas y de matiz de los mismos, y lo que es aún más difícil, una descripción de su funcionamiento en la lengua real» (2006: 14). «Ante tan diferentes significados por tanto, es innegable el hecho de que el contexto y la entonación sean considerados elementos necesarios para la apropiada decodificación del mensaje y la comprensión de determinados enunciados, y esto mismo justifica el hecho de que la utilización de sufijos cuantificadores se dé en su gran mayoría en la expresión oral, o bien en textos escritos que reflejan la lengua hablada» (2006: 30).

¹⁹³ La Real Academia (2009: 656) comenta que «las variantes no lexicalizadas de los sufijos afectivos aparecen raramente en los discursos de contenido más objetivo, como la prosa científica, la didáctica, la jurídica o la administrativa. No son infrecuentes, sin embargo, en la prosa ensayística, en el lenguaje periodístico y en el literario, puesto que los significados afectivos mencionados son compatibles en gran medida con los contenidos que se transmiten en todos esos géneros».

La versión del corpus que utilizaremos será exclusivamente el de la fase II, al igual que en el estudio de los demostrativos; es decir, los subcorpus comparables de español original y traducido en su versión más reciente y el subcorpus paralelo de originales en inglés. No utilizaremos la versión etiquetada en esta ocasión, dado que TreeTagger no dispone de etiquetas para sufijos específicos. Fijarnos en el etiquetado podría servir, de todos modos, para descartar categorías gramaticales, en el caso, por ejemplo, en el que decidiéramos no incluir los derivados de verbos con cambio de categoría gramatical. Sin embargo, debido a los errores de etiquetado apreciados en los capítulos anteriores y a la menor cantidad de concordancias que los sufijos generan, hemos decidido revisar todas las concordancias también en esta ocasión.

Nuestro punto de partida al enfrentarnos a este estudio consistió en la confección de un catálogo de sufijos apreciativos, de donde poder seleccionar los más idóneos para el análisis. La clasificación de los sufijos españoles no es tarea sencilla, dado su elevado número y la variedad de valores que pueden aportar a la base léxica. Por este motivo, los inventarios difieren bastante de una fuente a otra. La sufijación apreciativa no es la excepción¹⁹⁴. Varios de los sufijos que hemos encontrado clasificados en ciertas fuentes como apreciativos pueden agruparse también con los no apreciativos puesto que son capaces de adoptar también valores objetivos (compárese *chavalín* con *bailarín*). Dentro de los apreciativos, un mismo sufijo es categorizado como diminutivo o como despectivo dependiendo del autor, o como ambos, dado que la realidad es que en muchas ocasiones pueden aportar tanto un matiz como otro. A veces dependerá de la combinación del significado del sufijo con el de la base (pensemos en los sentidos y connotaciones que adscribimos a palabras como *taconazo*, *madraza* o *cabronazo*, por ejemplo) y otras, como ya hemos mencionado anteriormente, será el contexto el que nos indique qué intención ha tenido el emisor en cada caso:

Es que el sufijo es afectivo pero no dice él de por sí qué signo lleva ese afecto. La situación, las actitudes varias de los hablantes, las relaciones coloquiales, las condiciones rítmico-melódicas y el modo de ser llevado el tema, orientan en cada caso sobre la cualidad del afecto (Alonso 1982: 167).

¹⁹⁴ Santos Muñoz (2006: 19) señala: «(...) tal y como apuntó Beinhauer (1978), la sufijación es el sector de la lengua en el que existe en mayor medida una desconcertante cantidad de excepciones, irregularidades y generalizaciones que están en competencia, siendo precisamente el capítulo de los sufijos diminutivos y aumentativos el mayor quebradero de cabeza de todas las gramáticas y métodos de español, algo que no resulta sorprendente si tenemos en cuenta la dificultad de sistematizar los matices de un diminutivo/aumentativo que depende del contexto semántico.»

En las tablas siguientes puede observarse la variedad de sufijos que es posible hallar en los repertorios de los distintos autores que hemos consultado. Para no hacer las tablas demasiado largas, solo incluimos aquellos sufijos que son mencionados al menos por dos autores, aunque sea en diferentes categorías (por ejemplo, en diminutivos y peyorativos)¹⁹⁵.

Cuando un autor considera que determinado sufijo presenta matices peyorativos importantes, lo indicamos entre paréntesis («pey»). Si no se trata de matices, sino que el sufijo en cuestión se considera miembro de otra categoría, se indica sin paréntesis, como «aum» o «pey». A partir de su inclusión en seis repertorios de autores distintos, incluimos el sufijo en nuestro estudio. Por tanto, excluirémos los sufijos *-aco/-aca*, *-ajol/-aja*, *-ato/-ata*, *-ón/-ona*, *-orrio/-orria*, *-orro/-orra*, *-ucho/-ucha*, *-ujo/-uja*, *-uzco/-uzca*, *-usco/-usca*, *-uzo/-uza* y *-zuelo/-zuela*; al menos del listado de diminutivos, pues incluiremos alguno en otro de los listados (p. ej. *-orro/a* y *-orrio/a*).

DIM	Almela Pérez (1999)	Alonso (1982)	Erichsen (2007a)	Gómez Torrego (1997)	Gooch (1970)	Kaplanová (2010)	Lang (1990)	Lázaro Mora (1999)	Náñez Fernández (1973)	RAE (2009)	Rajo (2009)	Santibáñez Sáenz (1999)	Santos Muñoz (2006)	Seco (1989)	Total
-aco / -aca	aum (+pey)	pey	pey (+aum)		pey	pey (+aum)	pey (+aum)	pey		pey	aum (+pey)			x (+pey)	1
-ajo / -aja	pey		pey	pey	pey	pey	pey	pey		pey	pey		pey	x (+pey)	1
-ato / -ata		x		pey						pey				x	2
-ejo / -eja	pey		pey	pey	x	pey	pey	x	x	x	pey	x (+pey)		x (+pey)	6
-ete / -eta	x (+pey)	x	x	x	x	x	x (+pey)	x (+pey)	x	x	x (+pey)	x	x	x	14
-ico / -ica	x	x	x	x	x	x	x	x	x	x	x (+pey)	x	x	x	14
-illo / -illa	x	x	x	x	x	x	x	x	x	x	x (+pey)	x (+pey)	x	x	14
-in / -ina	x		x	x	x	x	x	x	x	x	x	x		x	12
-ino / -ina	x	x					x		x	x			x		6
-iño / -iña		x	x	x			x		x			x	x		7
-ito / -ita	x	x	x	x	x	x	x	x	x	x	x (+pey)	x	x	x	14
-ón / -ona	aum	x	aum	aum	aum	aum	x	aum		aum	aum		aum	aum	2
-orrio / -orria	pey	pey		pey	pey			pey		pey				x	1
-orro / -orra	pey	x		pey	pey	pey (+aum)	pey (+aum)	pey		pey	pey		aum (+pey)	x (+pey)	2
-ucho / -ucha	pey		pey	pey	pey	pey	pey	pey		pey	pey		pey	x	1
-uco / -uca	pey	x	pey	x	pey	pey	pey	pey	x	x			x	x (+pey)	6
-ujo / -uja	pey				pey	pey	pey	pey						x	1
-uelo / -uela	x		x	x	x	x (+pey)	x (+pey)	x	x	x	pey	x (+pey)	x	x	12
-uzco / -uzca, -usco / -usca	pey					pey	pey				pey			x	1
-uzo / -uza	pey	pey			pey	pey	pey	pey		pey	pey			x	1
-zuelo / -zuela			x	pey											1

Tabla 111 - Sufijos diminutivos en la bibliografía

En los casos en los que no hay acuerdo sobre la clasificación del sufijo, lo hemos incluido en la categoría en la que más autores lo clasifican. Como vemos, *-ejo / -eja* y

¹⁹⁵ Quedan fuera los siguientes sufijos, que solo hallamos en el repertorio de un único autor: diminutivos > *-ezno* (Seco); aumentativos > *-zón* (Almela Pérez); peyorativos > *-aino*, *-anas*, *-anco*, *-ancho*, *-apo*, *-enco*, *-inche*, *-itis*, *-men*, *-oncho* y *-ucio* (Almela Pérez), *-ardo* y *-arria* (Kaplanová), *-eo*, *-ería*, *-izante* e *-izo* (Gooch).

uco / *-uca* son considerados diminutivos por la mitad de los autores que los mencionan, mientras que la otra mitad los considera despectivos o despectivos con matiz diminutivo. Hemos optado por incluirlos en este grupo, porque el grupo de los peyorativos es más numeroso y porque la RAE los considera diminutivos a ambos.

En el caso de los aumentativos, hay dos autores menos en la tabla, pues Santibáñez Sáenz y Nández Fernández no mencionan este grupo de sufijos apreciativos en sus respectivos trabajos. Parece haber bastante consenso con respecto a los sufijos *-azo/-aza*, *-ón/-ona* y *-ote/-ota*. Del resto, el único que llega al umbral de seis que habíamos impuesto es *-udo/-uda*. Por lo tanto, excluimos de nuestra búsqueda de aumentativos los sufijos *-al*, *-acho/-acha*, *-aco/-aca*, *-ada* y *-orro/-orra*.

AUM	Almela Pérez (1999)	Alonso (1982)	Erichsen (2007b)	Gómez Torrego (1997)	Gooch (1970)	Kaplanová (2010)	Lang (1990)	Lázaro Mora (1999)	RAE (2009)	Rojo (2009)	Santos Muñoz (2006)	Seco (1989)	Total
-al					x			x					2
-acho / -acha	x (+pey)	pey	pey (+ aum)	pey	pey	pey (+ aum)	pey (+ aum)	pey		x (+pey)			2
-aco / -aca	x (+pey)	pey	pey (+ aum)		pey	pey (+ aum)	pey (+ aum)	pey	pey	x (+pey)		dim (+pey)	2
-azo / -aza	x	x	x	x	x	x (+pey)	x	x (+pey)	x	x	x	x	12
-ada					pey			x (+pey)					1
-ón / -ona	x	x	x	x	x	x	x	x (+pey)	x	x (+pey)	x	x	12
-orro / -orra	pey	dim		pey	pey	pey (+ aum)	pey (+ aum)	pey	pey	pey	x (+ pey)	dim (+pey)	1
-ote / -ota	x	x	x (+pey)	x	x	x (+pey)	x (+pey)	x	x	x (+pey)	x	x	12
-udo / -uda	x		x		x	x (+pey)	x (+pey)	x					6

Tabla 112 - Sufijos aumentativos en la bibliografía

Los peyorativos son el grupo más abundante y en el que más discrepancias hay entre autores. En la tabla se indica, entre paréntesis, un nuevo matiz que algunos autores atribuyen a ciertos peyorativos, además de los matices diminutivos o aumentativos que ya habíamos visto en los otros dos grupos: el humorístico («hum»). Como en los casos anteriores, decidimos excluir del análisis aquellos sufijos que no se encuentran catalogados en seis fuentes o más. Una de las razones para hacerlo es que, al no estar recogidos en estos listados, es probable que su productividad sea baja y no encontremos demasiados ejemplos en el corpus. Otra es que ciertos sufijos tradicionalmente considerados no apreciativos se incluyen a veces en los repertorios debido al hecho de que la base transmite, en ocasiones y, sobre todo, en el caso de los peyorativos, su significado o connotaciones al sufijo (*rencoroso*, *grasiento*), en lugar de ser el mismo sufijo el que posea este valor y lo aporte a la base —como sí es el caso de *naricita* y *narizota*, por ejemplo— (Almela Pérez 1999: 100-101; Lang 1990: 118), que es lo que buscamos nosotros.

PEY	Almela Pérez (1999)	Alonso (1982)	Erichsen (2007a, b)	Gómez Torrego (1997)	Gooch (1970)	Kaplanová (2010)	Lang (1990)	Lázaro Mora (1999)	RAE (2009)	Rojo (2009)	Santos Muñoz (2006)	Seco (1989)	Soler Espiauba (1996)	Total
-acho / -acha	x (+aum)	x	x (+aum)	x	x	x (+aum)	x (+aum)	x	x	aum (+pey)				9
-aco / -aca	x (+aum)	x	x (+aum)		x	x (+aum)	x (+aum)	x	x	aum (+pey)		dim (+pey)	x*	9*
-ada					x			aum (+pey)						1
-ajo / -aja	x		x	x	x	x (+aum)	x (+aum)	x	x	x	x	dim (+pey)	x	11
-ales				x*	x	x (+hum)	x (+hum)	x	x	x				7*
-allo / -alla	x				x	x		x						4
-ángano / -ángana	x				x			x						3
-ango / -anga	x				x	x (+hum)	x (+hum)	x	x	x				7
-arro / -arra	x					x (+aum)	x (+aum)			x			x (+aum)	5
-astre	x				x			x						3
-astro /astra	x		x	x	x	x (+aum)	x (+aum)	x	x	x				9
-ato / -ata		dim		x					x			dim	x*	3*
-azos / -azas	x				aum		x*	x						3*
-ejo / -eja	x		x	x	dim	x (+dim)	x (+dim)	dim	dim	x		dim (+pey)		6
-engo, -enga						x (+hum)	x (+hum)		x					3
-engue	x				x	x (+hum)		x	x					5
-eras	x			x*									x*	3*
-ero / -era					x								x*	2*
-esco / -esca	x				x			x						3
-eta(s) / -eto	x						x*						x*	3*
-ica	x			x		x	x*							4*
-ijo / -ija	x					x							x	3
-il	x (+dim)				x			x						3
-ingo / -inga	x				x	x (+hum)	x (+hum)	x	x	x				7
-ingue					x			x						2
-oca													x*	1*
-oide	x (+dim)				x		x (+hum)	x						4
-olis				x*										1*
-ongo / -onga	x					x (+hum)	x (+hum)							3
-orio / -oria					x			x						2
-orrio / -orria	x	x		x	x			x	x			dim (+pey)	x (+aum)	7
-orro / -orra	x	dim		x	x	x (+aum)	x (+aum)	x	x	x	aum (+pey)	dim (+pey)	x (+aum)	9
-ote / -ota	aum	aum	aum (+pey)	aum	aum	aum (+pey)	aum (+pey)	aum (+pey)	aum	aum (+pey)	aum	aum	x*	1*
-ucho / -ucha	x		x	x	x	x (+dim)	x (+dim)	x	x	x	x	dim (+pey)	x	11
-uco / -uca	x	dim	x	dim	x	x	x	x	dim		dim	dim (+pey)		6
-uelo / -uela	dim		dim	dim	dim	dim (+pey)	dim (+pey)	dim	dim	x	dim	dim		1
-ujo / -uja	x				x	x	x	x				dim (+pey)		5
-uno / -una	x									x				2
-ungo / -unga	x					x (+hum)	x (+hum)							3
-urrio / -urria	x	x												2
-urro / -urra						x (+aum)	x (+aum)							2
-uta / -uto	x												x	2
-ute	x				x		x	x	x				x	6
-uzco / -uzca, -usco / -usca	x					x	x			x		dim (+pey)		4
-uzo / -uza	x	x			x	x	x	x	x	x		dim (+pey)		8
-zuelo / -zuela			dim	x										1

Tabla 113 - Sufijos peyorativos en la bibliografía

La razón última de que nos encontremos con esta variedad de repertorios es el dinamismo del sistema de sufijación apreciativa, que, en palabras de Lang (1990: 100),

«is not closed, calls where necessary on suffixes which normally would be classified as non-emotive and is subject to expansion and renewal through the use of the neologistic morphemes of modern speech which lend colour and vivacity to the language»¹⁹⁶. De hecho, hemos añadido una autora a esta tabla, Soler Espiauba, dado que en su artículo de 1996 habla de ciertos sufijos irónico-afectivo-despectivos típicos de áreas urbanas españolas —especialmente Madrid—, como *-ata*, *-eta*, *-aca*, *-oca*, *-eras*, *-ero*, *-aco* (*bocata*, *jubileta*, *masoca*, *guaperas*, etc.) que, aunque no se encuentren en los repertorios tradicionales, nos parecen un objeto de estudio muy interesante y coinciden en parte con los que Gómez Torrego denominaba sufijos familiares. Por este motivo, no haremos siempre caso a la Real Academia, y sí tendremos en cuenta ciertos sufijos o variantes como *-ota* o *-ica*, a pesar de que provocan un cambio de categoría gramatical de la base (generalmente marcaremos estos derivados con un asterisco, tanto en la tabla 3 como en las que iremos incluyendo a lo largo del análisis). Como explicaremos en el apartado correspondiente, esta es la razón por la cual en el listado final de despectivos incluimos *-ata/o*, a pesar de que este sufijo no alcanza el umbral de los cinco autores.

Los elementos que analizaremos en este capítulo (en sus variantes masculinas, femeninas, neutras y plurales) serán, entonces, los siguientes:

➤ DIMINUTIVOS (10)

1. -EJO / -EJA
2. -ETE / -ETA
3. -ICO / -ICA
4. -ILLO / -ILLA
5. -ÍN / -INA
6. -INO / -INA
7. -IÑO / -IÑA
8. -ITO / -ITA
9. -UCO / -UCA
10. -UELO / -UELA

¹⁹⁶ Gooch (1970: 25) explica que hay sufijos que gozaron de mayor popularidad en el pasado y que hoy son apenas productivos. Lang (1990: 120), habla de aquellos sufijos que surgen constantemente en el habla popular y son de naturaleza más bien efímera: «Particular suffixes are used more often humoristically than not. They are of a neologistic character, used spontaneously and form derivatives which are ad hoc or ephemeral. The stylistic register in which they occur is popular speech, the language of the young, and the satirical press. Here only a few are considered, as representative of a type which flourishes and declines according to the dictates of linguistic fashion» (se refiere a *-ales*, *-oide*, *-ango*, *-ingo*, *-ongo* y *-ungo*).

◆ AUMENTATIVOS (4)

11. -AZO/-AZA
12. -ÓN/-ONA
13. -OTE/-OTA
14. -UDO/-UDA

◆ PEYORATIVOS (13)

15. -ACHO / -ACHA
16. -ACO / -ACA
17. -AJO / -AJA
18. -ALES
19. -ANGO / -ANGA
20. -ASTRO /ASTRA
21. -ATO / -ATA
22. -INGO / -INGA
23. -ORRIO / -ORRIA
24. -ORRO / -ORRA
25. -UCHO / -UCHA
26. -UTE
27. -UZO / -UZA

El hecho de que muchos de estos sufijos cuenten con variantes alomorfas, es decir, de idéntico significado (p. ej.: **bichito** / **pastorcito** / **solecito**), es algo que preocupa a algunos estudiosos; sobre todo por el interés de establecer determinadas reglas que expliquen o predigan con exactitud qué variante se seleccionará para cada base y por qué. Sin embargo, a nosotros no nos interesa esta cuestión para el análisis del corpus, puesto que buscaremos siempre la forma básica del sufijo que aparece en todas las variantes (en este caso, *-ito*), tanto en masculino como en femenino y el plural de ambos.

En la búsqueda y clasificación de las palabras formadas por estos sufijos, habrá que tener en cuenta distintos factores. En primer lugar, lo que Gooch (1970: 22-23)

denomina «mirage suffixes», es decir, palabras que, aparentemente terminan en un sufijo español, pero en realidad no es así, puesto que provienen de otras lenguas, como el italiano (*cuneta, estafeta*), el francés (*billete, bayeta*), el árabe (*mamarracho, azulejo*) o, en el caso de muchos términos acabados en *-illo* e *-illa*, de diminutivos del latín que han perdido dicho sentido con el tiempo (*castillo, astilla*).

En segundo lugar, el grado de transparencia de las palabras derivadas, para identificar el estado de lexicalización de ciertos vocablos que, a pesar de tener su base en un término español, ya no pueden ser considerados el producto de la adición de un sufijo apreciativo a una base de manera facultativa por parte de un hablante. Estas palabras han pasado a designar un concepto diferente al de la palabra a la que se han añadido y, dado que el significado resultante de la unión de una base y un sufijo ha dejado de ser «transparente» para el hablante (en el sentido de que pueda interpretarse directamente a partir de la combinación de los significados de ambos, como es el caso de *vasito*, «vaso pequeño»), muchas veces ese nuevo significado aparece recogido en los diccionarios (por ejemplo, *camisa* > *camisón* o *central* > *centralita*). La Real Academia explica en la página 28 de su *Nueva Gramática* que el diccionario de la Real Academia no «da cabida a la mayor parte de las voces formadas con sufijos apreciativos (*arbolito, fiebrón, etc.*), salvo cuando se han fosilizado o están lexicalizadas...». Por tanto, en principio, si nos atenemos a esta afirmación, no deberíamos recoger en nuestro listado de palabras sufijadas las que se recojan en el DRAE. Sin embargo, como la misma Academia aclara (2009: 635), los conceptos de *transparencia* y *opacidad* son «naciones graduales, puesto que se basan en la conciencia lingüística del hablante, siempre variable, a menudo dependiente de su cultura particular y, por tanto, difícil de objetivar». A veces será necesario utilizar nuestro criterio para decidir si incluir una palabra o no en nuestro listado, primero porque el diccionario puede recoger ciertas palabras cuyo significado, desde nuestro punto de vista, sigue siendo transparente (p. ej. *gustazo* o *casucha*); segundo, porque también pueden incluirse usos particulares de palabras que, sin embargo, tienen un significado composicional transparente, como *cachorrillo*, que tiene el significado lexicalizado de «pistola pequeña», pero también se interpreta de manera transparente como «cachorro pequeño»; y, por último, porque a veces un término puede no aparecer en el diccionario a pesar de haber perdido transparencia su significado, como veremos en determinados casos.

6.3. Resultados y discusión

Para la descripción de los sufijos nos basaremos principalmente en los autores recogidos en las tablas 1, 2 y 3. El resto de nuestras fuentes aparecen en las notas a pie o recogidas en la bibliografía. Debajo de cada descripción incluimos la expresión utilizada para la búsqueda en el corpus, una tabla con los resultados de la búsqueda de palabras sufijadas, otra con la distribución de formas por autores en cada subcorpus y otra con la categorización de los elementos activadores del corpus de originales en inglés.

6.3.1. Diminutivos

6.3.1.1. -EJO / -EJA

Como comprobamos durante el proceso de selección de sufijos, no hay acuerdo en cuanto al valor fundamental de *-ejo/a*, pues algunos autores lo clasifican dentro de los diminutivos y otros, dentro de los peyorativos. Gooch (1970: 10) lo incluye dentro del grupo de los diminutivos, aun comentando su tendencia a expresar connotaciones peyorativas, que lo asemejan a *-illo*. Prieto (2005: 57), sin embargo, es tajante: «Of these, *-ejo* is not really a diminutive; it serves more pejorative and attenuating functions».

Nosotros hemos decidido, por las cuestiones prácticas que ya hemos comentado, incluirlo en este primer grupo. De todas formas, prácticamente todos los autores le asignan, en mayor o menor medida, ambos matices. Nández Fernández (1973: 44) afirma: «El sufijo *-ejo* tan abundante como un simple apreciativo con valor despectivo, es también diminutivo, y en alguna región española, como Guadalajara, muy usado bajo cualquier signo valorativo, abundan las formas positivas o neutras».

Se utiliza en La Mancha y en otras partes del área centromeridional de España, en algunas zonas del Río de la Plata y del español andino (RAE 2009: 631).

Ejemplos: lugarejo, bichejo, librejo

Ejemplos de términos lexicalizados: candileja, marmolejo, molleja

De origen latino: abeja, oreja, lenteja, vencejo

[word=".*ej(o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
calleja(s)	6	1,6			
palabreja	2	0,5			
tipeja/o	2	0,5	tipejo(s)	8	1,6
Total	10	2,7		8	1,6

Tabla 114 - Sufijo *-ejo* en ES y TRAD

Lexicalizadas ES: azulejo, mollejas

Lexicalizadas EN: bosquejo, candilejas

En este caso, las tres palabras halladas en el corpus podrían considerarse lexicalizadas, puesto que aparecen en el diccionario de la Real Academia Española. Sin embargo, Gooch no las incluye dentro de sus grupos de términos de significado cambiado o especializado, y creemos que el significado de las tres puede deducirse de sus componentes, aunque teniendo en cuenta dos valores diferentes del sufijo *-ejo/a*. Por un lado, *calleja*, que es claramente una calle más pequeña de lo normal, y, por otro, *palabreja* y *tipejo/a*, que son términos despectivos para referirse a personas o palabras ridículas o sin importancia. Por estos motivos y, dado que no hay más casos en el corpus, hemos decidido tenerlas en cuenta.

Encontramos más términos sufijados en ES (10) que en TRAD (8), a pesar de ser el primero el subcorpus con menos cantidad de palabras. Las frecuencias normalizadas nos indican una presencia de una única palabra sufijada más en ES que en TRAD cada 100 000 palabras del subcorpus correspondiente. Dicha diferencia, según el test de log-likelihood, no es suficiente para pasar el umbral de significación estadística.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	3	3,5	Ellroy	7	6,4
Martín	3	5,3	Leonard	1	1,3
Vázquez	2	2,3	Cornwell	0	0
Martínez	1	1,8	Grafton	0	0
Núñez	1	2,3	Highsmith	0	0
Madrid	0	0			

Tabla 115 - Distribución del sufijo *-ejo* por texto

En la tabla 5 podemos ver la frecuencia de uso del sufijo en los distintos textos. Como en los dos capítulos anteriores, las columnas de frecuencias normalizadas (tercera

y sexta) nos muestran dichas frecuencias multiplicadas por 100 000 y divididas por el número de palabras de la obra en la que aparece la palabra sufijada. Gracias a esta normalización, sabemos que Martín hace un uso relativamente más abundante de este sufijo que Giménez-Bartlett, a pesar de que esta autora presente el mismo número de concordancias.

En ES, *calleja* coloca con el adjetivo *estrecha* en tres ocasiones, lo que confirma la utilización diminutiva del sufijo con esta base. Las dos veces que aparece *tipejo* la connotación es, como habíamos adelantado, claramente despectiva, pues en un caso se refiere a la responsable de la creación de un «engendro» y en el otro, a un detective «de pocas luces». *Palabreja*, sin embargo, es más complicada de interpretar, pues en los textos se refieren con este vocablo sufijado a los términos *neurastenia* y *Ouroboros*, que no parecen ser dignas de desprecio ni poco importantes, sino más bien extrañas o complicadas. El personaje de Núñez, por ejemplo, «da con ella» después de pensar un rato. Quizás podríamos aventurar que en estos contextos, el sufijo expresa fastidio por no hallar la palabra adecuada de inmediato. El sentido despectivo de este sufijo es, por lo que vemos, bastante fuerte.

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo ¹⁹⁷	6
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	1
Información contextual ¹⁹⁸	1
Tono general del fragmento (compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
Ø	
Total	8

Tabla 116 - Elementos activadores del sufijo *-ejo* en traducciones

Todos los términos utilizados en los textos originales en inglés —que, recordemos, se traducen por *tipejo*— son coloquiales y prácticamente todos, a excepción de *dude* (el único perteneciente al texto de Leonard) son despectivos:

¹⁹⁷ Como lo define Schneider (2003: 19): «a third type of diminutives, in addition to the synthetic and analytic types (...). Examples of inherent diminutives include *town* in relation to *city*, and *hill* in relation to *mountain*».

¹⁹⁸ En este caso, la traducción de *dude* por el más despectivo *tipejo* se entiende si leemos el fragmento entero, en el que el personaje contrapone su vida a la de otra persona, cualquiera, que tuviera que sacrificar para no morir o ir a la cárcel, restándole valor continuamente:

<Leonard>: ¿Te crees que si algún <tipejo> se interpone en mi camino no me lo voy a cargar ?
 -->par_en: If it's him or you , or if it's him doing time or you doing time ? There's nothing to think about , man , you take him out. " Ordeil said , " Once I pick up the goods and make one more delivery ? I won't ever have to work again till I've spent something like a million bucks . You think some **dude** gets in my way I won't remove him ? "

*spook*¹⁹⁹ (un término peyorativo para referirse a personas negras), *jig rig* (a pesar de su traducción, se refiere a un coche clásico, *tuneado* y decorado con mal gusto), *queers* (sinónimo ofensivo de homosexual), *punks* (vándalo, gamberro), *shit* (en este caso, un vocativo ofensivo) precedido por el adjetivo *little*, que probablemente sea un incentivo bastante fuerte para la selección en español de un diminutivo), y *hump* (otro término despectivo para referirse a una persona) y *geek* (bicho raro) precedidos por los adjetivos *skinny* (delgado) y *bland* (insulso), respectivamente, que también apuntan a que el individuo en cuestión es poca cosa. Sin embargo, toda esta variedad de sustantivos tiene como equivalente único en español a *tipejo*, por lo que quizás sea este el equivalente lexicalizado o estándar para este tipo de denominaciones despectivas para personas en español —o una preferencia del traductor, puesto que siete de ocho ejemplos provienen del texto traducido por Sabaté Vargas—.

6.3.1.2. -ETE / -ETA

Gooch lo describe como un sufijo de carácter marcadamente humorístico (1970: 17), la RAE le asigna valores de ironía, complicidad o atenuación (2009: 653) y Lázaro Mora afirma de él que colorea los objetos nombrados con «una especie de aprecio burlador, capaz igualmente de resultar positivo o de descalificar y despreciar», como es el caso de la palabra *amiguetes* para referirse a «compinches de bar» (1999: 4677).

Lang (1990: 103-106) le atribuye un uso bastante restrictivo —dadas sus fuertes tendencias lexicalizadoras, como en *glorieta* o *burlete*—, más proclive a ser peyorativo que otros diminutivos como *-ito* o *-illo*, aunque no tanto, en su opinión, como *-uelo* (el autor no habla de *-ejo* en este caso, dado que lo incluye en el grupo de los peyorativos) y con tendencia a aparecer combinado con otros sufijos apreciativos, como en *pobretón*. Santos (2006: 34-37), sin embargo, le atribuye en muchos casos un marcado carácter afectivo cariñoso, como en la frase «era muy majete», aunque reconoce que puede ir unido a valores eufemísticos o atenuadores.

Es un sufijo más utilizado en el español europeo que en el americano y, aunque se utiliza de manera general en toda España, «destaca proporcionalmente en Aragón, Levante y Cataluña» (RAE 2009: 631).

¹⁹⁹ Las definiciones de los términos han sido extraídas de *Urban Dictionary* y *WordReference English Spanish Dictionary* [fecha de consulta: 11 de agosto de 2015].

Lang (1990: 105) habla de la variante *-etas* (muy parecida a *-ica* u *-ota*, que veremos más adelante), utilizada sobre todo en la jerga escolar americana para designar a personas de manera despectiva o con algo de censura (por ejemplo, *lloretas*, pero también menciona *acuse* y *charleta*). El problema con estas construcciones es que cambian la clase gramatical de la base que les da origen (*acusar*, *llorar*, *charlar*) de verbo a sustantivo, por lo que, en principio, *-etas* (y sus variantes deverbalizadoras *-ete* y *-eta*) no se ajustarían a la definición de sufijo apreciativo.

Sin embargo, dado que son términos que se utilizan para calificar de manera apreciativa a una persona, consideramos que estaría justificada su inclusión en este estudio, en el que, como hemos dicho en otros capítulos, no nos interesan tanto las categorías gramaticales estancas como los usos particulares del lenguaje que pueden tener interés a la hora de comparar el español original y el traducido. Por ello, sea cual sea nuestra decisión, se mantendrá para ambos subcorpus, y aclararemos qué palabras han quedado fuera de la selección por considerarse lexicalizadas. Por otro lado, no tendremos en cuenta los nombres propios creados a partir de diminutivos (*Enriqueta*), a excepción de que no se utilicen de manera sistemática, sino con una clara intencionalidad afectiva (Náñez Fernández 1973: 213).

Ejemplos: dinerete, amiguete, pobrete, guapete

Ejemplos de términos lexicalizados: palacete, vejete, careta, boquete, aleta

De origen árabe: alcahuete

De origen francés: florete, furgoneta, servilleta

De origen italiano: cuneta, estafeta

[word=".*et(e|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
amiguete	2	0,5	amiguetes	1	0,2
chorrete	1	0,3	regordete/a(s)	4	0,8
chuleta	2	0,5			
esbirrete	1	0,3			
majete	3	0,8			
negociete	1	0,3			
sicariete	5	1,4			
solete	1	0,3			
tieta	1	0,3			
vejete	1	0,3			
regordete/a	2	0,5			
Total	20	5,4		5	1

Tabla 117 - Sufijo *-ete* en ES y TRAD

Lexicalizadas ES: avioneta, banqueta, bragueta, caballete, camiseta, caldereta, carrete, caseta, carreta, cebolleta, chinchetas, chupete, claqueta, coleta, cubiletes, jugarreta, libreta, meseta, papeleta, pedorretas, peseta, piquete, plaqueta, rabieta, salmonetes, veletas, vinagreta, violeta
Lexicalizadas EN: agujetas, aletas, arete, banqueta, bragueta, caballete, camiseta, camioneta, carrete, careta, caseta, casquete, cazoleta, chancletas, chinchetas, coleta, colorete, cornetas, estiletes, falsete, grilletes, jugarreta, libreta, maceta, machete, palanqueta, paleta, panceta, papeletas, piquete, puñetas, sopletes, trompeta, vinagreta, violetas

No hemos hallado ejemplos de palabras derivadas de verbos ni mediante la variante *-etas*. Hemos tenido dudas con la inclusión de *regordete*, pues se encuentra en el DRAE y es una palabra especial, ya que precisa también un prefijo para su formación (no existe la palabra *regordo*), pero, dado que Gooch no la considera lexicalizada y que puede encontrarse en ambos subcorpus, la incluimos sin temor a sesgar los resultados en uno u otro sentido. El total de 20 palabras para ES y 5 para TRAD (o 5,4 por cada cien mil palabras frente a una) representa, esta vez, una diferencia estadísticamente significativa (LL +14,43 $p < 0,001$).

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	8	14,2	Grafton	2	2,1
Madrid	5	12,1	Cornwell	1	0,9
Vázquez	4	4,5	Ellroy	1	0,9
Giménez	1	1,2	Leonard	1	1,3
Martínez	1	1,8	Highsmith	0	0
Núñez	1	2,3			

Tabla 118 - Distribución del sufijo *-ete* por texto

Mientras que palabras como *amiguete*, *majete*, *solete*, *tieta*, *vejete* o *regordete* se encuentran en contextos amables, amistosos o humorísticos, *chuleta* es claramente una palabra despectiva que se utiliza para designar a alguien que no cae bien. *Esbirrete* disminuye la importancia de la persona a la que se atribuye el adjetivo y también denota antipatía, pues es alguien que está forzando a quien habla a hacer algo que no quiere. Lo mismo sucede con *sicariete*, acompañado además por adjetivos que refuerzan el desprecio (*torpe*, *canijo*, *loco*) y con el eufemístico *negociete*: el sufijo resta gravedad a la base porque se sabe que se está hablando de algo turbio.

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	
Sufijo apreciativo	3
Adjetivo o adverbio con connotación apreciativa + sustantivo	1
Información contextual	
Tono general del fragmento (razones estilísticas o compensación) ²⁰⁰	1
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
Ø	
Total	5

Tabla 119 - Elementos activadores del sufijo *-ete* en traducciones

El sustantivo que da pie al uso de *amiguete* en TRAD es *buddy*, un término informal para referirse a un amigo (compañero, colega). La marca de diminutivo inglés *-y* puede haber activado el uso del diminutivo en español, que no era imprescindible en este caso pero aporta un matiz irónico que va bien a la situación:

<Grafton>: - Ah . Ahora somos socios . El truco de los <amiguetes> - dije .

-->par_en: The buddy system , " I said

En dos apariciones de *regordete*, la descripción de la persona se hace en términos negativos (*desagradablemente, desaliñado*). En uno de los originales vemos que se utiliza simplemente el adjetivo *plump*, pero en el otro sí que existe una marca diminutiva en el adjetivo *little* antepuesto a *fat*. En los contextos en los que no existe esa connotación peyorativa, los términos utilizados en el original son *pudgy* y *chubby*, que contienen esa marca final de diminutivo que las hace más simpáticas.

6.3.1.3. -ICO / -ICA

Gooch lo considera más afín a *-ito* por sus matices favorables y cariñosos. Lo asocia especialmente con Aragón y Navarra, y pone como ejemplo el nombre de la patrona de Zaragoza, La Pilarica, y los gentilicios *navarrica* y *pamplonica* (1970: 26),

²⁰⁰ Distinguimos estos casos de los de la categoría «información contextual», dado que no se nos proporciona información en el contexto acerca de por qué el traductor puede haber llegado a la conclusión racional de que el uso del diminutivo era la mejor opción, sino que vemos que el tono general del fragmento ha propiciado su uso, como compensación de otros elementos apreciativos que no han podido traducirse, o simplemente por razones estilísticas (en el apartado de *-ito/a* se proporcionan más ejemplos para distinguir estas dos categorías):

<SueGrafton>: Tenía los brazos <regordetes> y unos muslos que tenían que frotarse con fuerza al andar .

-->par_en: Her hair was dark and straight , anchored unbecomingly behind her ears . She was wearing flip-flops , tight black pedal pushers , and a black sleeveless cotton top , stretched over drooping breasts . Her upper arms were **plump** and her thighs looked like they would chafe against one another when she walked . Everything about her looked unpleasantly damp .

aunque también se refiere a algunos usos del sufijo en Granada. Lang amplía su esfera de uso al este de Andalucía y a América Central y comenta que puede encontrarse en autores que en principio no presentan marcas regionales (1990: 108). Otros autores lo ubican también en Murcia y la Comunidad Valenciana. Al parecer, fue especialmente productivo entre los siglos XV y XVII (Prieto 2005: 57) y su uso en aquella época estaba más extendido en el centro y norte de España (RAE 2009: 630).

Como sucedía con *-ete*, cuenta con una variante epicena peyorativa, *-ica*, esta vez propia del español europeo, que Lang considera característica también del lenguaje infantil y puede generar cambio de categoría gramatical. Algunos ejemplos son los sustantivos o adjetivos *llorica*, *abusica*, *acusica*, *miedica*, *cobardica* o *quejica* (1990: 109), muchos de los cuales provienen de verbos. Evidentemente, la Real Academia (2009: 630) no los considera casos de derivación apreciativa, pero, una vez más, nosotros los incluiremos en nuestros listados por el interés que pueden tener para caracterizar y comparar el lenguaje utilizado en las novelas del corpus.

Ejemplos: fotico, cinturica, pajarico, pequeñoico

Ejemplos de términos lexicalizados: abanico, perico, vainica

[word=".*(i|í)c(o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
quejicas	1	0,3	quejicas	2	0,4
Total	1	0,3		2	0,4

Tabla 120 - Sufijo *-ico* en ES y TRAD

Lexicalizadas ES: abanico, perico

Lexicalizadas EN: abanico, marica

A pesar de encontrar numerosísimos casos del sufijo no apreciativo *-ica* («relativo a / origen / condición» [Almela Pérez 1999: 115]: *bíblico*, *céntrico*, *cinematográfico*), apenas hay casos apreciativos. Quizás sea este un sufijo demasiado marcado regionalmente, a pesar de lo que decía Lang. En el texto de Madrid hallamos el hipocorístico *Albertico*, pero no lo incluimos porque, dado que nadie se refiere al personaje de ninguna otra manera, consideramos que ha pasado a ser la manera común de denominarlo y no un uso afectivo del lenguaje.

Evidentemente, la diferencia entre subcorpus, a favor de TRAD esta vez, no puede ser tenida en cuenta desde el punto de vista estadístico.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	1	1,8	Ellroy	2	1,8
Giménez	0	0	Cornwell	0	0
Madrid	0	0	Grafton	0	0
Martínez	0	0	Highsmith	0	0
Núñez	0	0	Leonard	0	0
Vázquez	0	0			

Tabla 121 - Distribución del sufijo *-ico* por texto

Quejicas se utiliza claramente de manera despectiva: en el texto de Madrid va acompañado de otros adjetivos de connotaciones negativas (*lloronas, pusilánimes*). En cuanto a las dos apariciones en TRAD, hay que mencionar que, en realidad, son el mismo fragmento repetido, pues el protagonista está escuchando la misma grabación dos veces. La expresión original es «These places are filled with losers and lonesome creeps»; todos los elementos son negativos. La traducción de Sabaté Vargas es «Estos lugares están llenos de perdedores y de quejicas solitarios». *Creep* es una persona indeseable o repelente, pero la traducción es adecuada en este contexto porque lo que motiva esa afirmación es que en aquel lugar no paran de oírse sollozos.

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	
Información contextual	2
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
∅	
Total	2

Tabla 122 - Elementos activadores del sufijo *-ico* en traducciones

6.3.1.4. -ILLO / -ILLA

Según la Academia, fue uno de los sufijos predominantes en la lengua medieval y clásica (2009: 630) y en la época actual, presenta un uso proporcionalmente mayor en Andalucía, aunque su uso está muy extendido en todas las regiones. Aunque puede utilizarse en ocasiones de manera afectuosa, tiene facilidad para expresar matices

peyorativos, como ejemplifica Gooch (1970: 6) comparando los diferentes contextos en los que utilizaríamos *animalito* («es un animalito precioso») y *animalillo* («¡qué animalillo más repugnante!»). Por otro lado, puede aportar un matiz atenuador en los casos en los que el hablante busca restar importancia al significado de la palabra a la que se aplica el sufijo, como cuando decimos que alguien es *feíllo* o *envidiosillo* (Santos Muñoz 2006: 49).

Su tendencia a la especialización de significados es también muy fuerte (Lang 1990:102). Lázaro Mora observa que las lexicalizaciones con *-illo* son mucho más numerosas que con otros sufijos y lo explica de la siguiente manera:

En efecto, *-illo/a*, el más antiguo de los tres sufijos, a partir del siglo XIV había experimentado un desgaste evidente en la expresión del afecto, y eso es, justamente, lo que a partir de entonces la permitió desprenderse con mayor facilidad de su componente afectivo. *-Ito/a* e *-ico/a*, en cambio, son mucho más firmes en mantener su filiación semántica (*casita* sigue siendo una casa, aunque pequeña) y su orientación positiva del afecto.

Por el contrario, *-illo/a* (desgastado y sin un componente afectivo intenso) frena mucho menos la posibilidad de que el vocablo se lexicalice con acepciones propias, y de que signifique peyorativamente o con desafecto» (1999: 4676).

Ejemplos: bichillo, durillo, guapilla, nubecilla

Ejemplos de términos lexicalizados: cigarrillo, estribillo, manecilla, boquilla

De origen latino: tobillo, cuchillo, astilla, rodilla

[word=".*(i|í)ll(o|a)(s)?"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
abogadillo	1	0,3	asuntillo	1	0,2
hatillo	3	0,8	bosquecillo	1	0,2
bigotillo	1	0,3	cachorrillos	1	0,2
calorcillo	2	0,5	chavalillos	1	0,2
cancioncillas	1	0,3	cosilla	2	0,4
carcajadilla	1	0,3	duendecilla/o	2	0,4
carpetilla	1	0,3	estatuilla	1	0,2
carrerilla	1	0,3	hilillo	2	0,4
chalequillo	1	0,3	huesecillos	1	0,2
chavalilla	1	0,3	listillo	2	0,4
choricillos	2	0,5	mentirijilla	1	0,2
cohecillo	1	0,3	pececillos	1	0,2
cocinilla	2	0,5	pedrecillas	1	0,2
cosillas	2	0,5	problemilla	1	0,2
desvergonzadillos	1	0,3	risilla	1	0,2
enteradilla	1	0,3	rojillo(s)	7	1,4
figurillas	1	0,3	ruedecilla(s)	3	0,6

golfillo(s)	2	0,5	torrecillas	1	0,2
hilillo	1	0,3	trabajillo	1	0,2
hombrecillo	2	0,5			
honorcillo	1	0,3			
jardincillo	1	0,3			
jersecillo	1	0,3			
jovencillo	1	0,3			
listillo	2	0,5			
mentirijillas	1	0,3			
mentulilla	1	0,3			
mierdecilla	1	0,3			
montajillo	2	0,5			
moquillo	1	0,3			
musiquilla	2	0,5			
nubecilla	1	0,3			
ojillos	7	1,9			
olorcillo	2	0,5			
pajarillos	2	0,5			
pastorcilla	1	0,3			
pelillos	1	0,3			
pesetillas	3	0,8			
pielecilla	1	0,3			
pobrecillo	2	0,5			
politiquillo	1	0,3			
poquillo	1	0,3			
puestecillo	1	0,3			
putillas	1	0,3			
sonrisilla	2	0,5			
tijerillas	1	0,3			
tonelillo	1	0,3			
trabajillo	1	0,3			
trotecillo	1	0,3			
venillas	1	0,3			
vientecillo	1	0,3			
Total	74	20,1		31	6,3

Tabla 123 - Sufijo *-illo* en ES y TRAD

Lexicalizadas ES: alcantarilla, altillo, avanzadilla, barandilla, barbilla, bocadillo, bolsillo, bombilla, bordillo, buhardilla, cajetilla, calderilla, calzoncillos, camilla, campanilla, carillas, casillas, cepillo, chiquillos, cigarrillo, colilla, conejillo, comidilla, coronilla, cuartillas, cucharilla, cursillo, descansillo, dinerillo, dobladillo, escalerillas, espinillas, flequillo, frenillo, gatillo, gentecilla, gravilla, guerrilla, gusanillo, horquilla, lamparillas, lazarillo, manecillas, maquinilla, marisabidilla, mesilla, mirilla, muletillas, mundillo, nudillos, panecillos, pardillo, pasillo, patillas, plantilla, platillo, plumilla, presillas, puntillas, rabillo, rejilla, rosquillas, solomillo, (de) tapadillo, taquilla, tetilla, tonillo, tortilla, vaquillas, ventanilla, visillos, zancadillas, zapatillas

Lexicalizadas EN: alcantarilla, alfombrilla, almohadilla, ardilla, banquillo, barandilla, barbilla, bocadillo, barquillo, bolsillo, bombilla, boquilla, bordillo, buhardilla, cajetilla, calderilla, calzoncillos, camilla, campanillas, carboncillo, cardenillo, carretilla, cartilla, casilla, casquillo, cepillo, cerilla, chiquilla/o(s), cigarrillo, coletillas, colilla, comillas, coronilla, corrillos, cuadernillo, cucharilla, cursillo, descansillo, dobladillo, ensaladilla, escalerilla, escuadrilla, escobilla, espinilla, estampilla, esterilla, gatillo, gravilla, hornillo, horquilla, manecilla,

maquinilla, marisabidilla, mascarilla, mercadillo, mesilla, mezclilla, mirilla, nudillos, palillos, panecillos, palomilla, pandilla, pasillo, pastelillo, patillas, pepinillos, picadillo, platillo, puntillas, rabadilla, rabillo, redecilla, rejilla, solomillo, tablilla, taquilla, trabillas, trampilla, vainilla, ventanilla, visillos, zapatilla(s)

En este caso, la cantidad de palabras es muy grande, por lo que en bastantes ocasiones nos han surgido dudas acerca de si determinados términos debían considerarse lexicalizados o podían incluirse en los listados. Nuestro criterio, como ya hemos adelantado, es no incluirlas si aparecen en el diccionario con una acepción que sea diferente a la apreciativa y no pueda deducirse directamente de los significados aportados por la base y el sufijo (RAE 2009: 657). El criterio último si el diccionario no aclara la cuestión es, por supuesto, nuestra intuición lingüística como nativos de que dicha palabra no se emplea ya con intención afectiva, sino que el sufijo ha pasado a formar parte inherente de ella, como en *campanilla* o *calzoncillos*.

Hay una diferencia bastante importante entre las apariciones de este sufijo en ES y en TRAD (43 casos; casi 14 más en ES por cada 100 000 palabras si normalizamos las cifras), muy significativa según el test estadístico (LL +32,57 $p < 0,0001$).

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	18	21,2	Ellroy	13	11,8
Martín	17	30,2	Grafton	9	9,7
Madrid	16	38,6	Highsmith	4	4,1
Martínez	8	14,7	Leonard	3	3,8
Núñez	8	18,8	Cornwell	2	1,8
Vázquez	7	8			

Tabla 124 - Distribución del sufijo *-illo* por texto

Entre las concordancias del subcorpus de originales en español encontramos que los términos incluidos en el listado cumplen las siguientes funciones:

- Restar importancia o gravedad (atenuar): choricillos, cosillas, golfillo, honorcillo, mentirijillas, montajillo, carcajadilla, desvergonzadillos, olorcillo, pesetillas, puestecillo, trabajillo, trotecillo;
- Simplemente disminuir: hatillo, carpetilla, chalequillo, chavalilla, cochecillo, cocinilla, figurillas, hilillo, jardincillo, mentulilla, moquillo, ojillos, pajarillos, pelillos, pielecilla, poquillo, tijerillas, venillas;
- Aportar un tono humorístico, ridiculizar o insultar: bigotillo, cancioncillas, carrerilla, enteradilla, hombrecillo, listillo, mierdecilla, politiquillo, abogadillo,

putillas, sonrisilla, tonelillo;

- Referirse a algo agradable o a alguien de manera afectuosa: calorcillo, musiquilla, nubecilla, pastorcilla, pobrecillo, vientecillo.

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	12
Sufijo apreciativo	6
Adjetivo o adverbio con connotación apreciativa + sustantivo	7
Información contextual	3
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional ²⁰¹	2
Información distribuida en más de un elemento lingüístico ²⁰²	1
Ø	
Total	31

Tabla 125 - Elementos activadores del sufijo *-illo* en traducciones

En cuanto al subcorpus de traducciones, el elemento activador más frecuente es simplemente una palabra con significado diminutivo o peyorativo inherente: *puppy* (traducido como *cachorrillo*), *pixie* (*duendecillo*), *trickle* (*hilillo*), *fib* (*mentirijilla*), *giggle* (*risilla*) o *hotshot* (en este caso, se refiere a alguien que tiene una opinión de sí mismo exageradamente buena y se traduce por *listillo*). También podemos encontrar bastantes casos de adjetivos de connotación diminutiva añadidos al sustantivo: *small* (*forest, fish*), *little* (*kids, thing, bones*) y *short* (*tower*). Tenemos algunos casos de diminutivos o atenuación cuya utilización es comprensible si atendemos al contexto:

<Ellroy>: no vaya a jorobarnos ese <asuntillo> personal que tenemos entre manos .

-->par_en: We do n't want him screwing up this personal thing of ours . "

<Grafton>: oftalmoscopio - dijo - . Los utilizan los oculistas para mirar dentro de los ojos . Permite ver las

<ruedecillas> de la combinación y saber cómo se mueven .

-->par_en: This shines light on the combination wheels so I can see what we got going . "

Finalmente, hallamos un diminutivo atenuador formado por información procedente de varios elementos lingüísticos (*kind of a problem: problemilla*), un sufijo

²⁰¹ Rojo (2009: 239-240): «Sometimes translators may opt for a functional equivalent that gives priority to stylistic naturalness over semantic accuracy».

²⁰² Como hemos explicado más arriba, es la estrategia explicada por Rojo (2009: 239) pero en la dirección de traducción inversa a la que estamos analizando (traducción de diminutivos españoles al inglés): «Sometimes the meaning of the stem is provided in a separate adjective and the information of the suffix is conveyed in an agentive noun».

repetido seis veces (*pinko: rojillo*) y dos equivalentes funcionales (*estatuilla de la Academia* por *Academy Award* y *pedrecillas* por *jewels*).

6.3.1.5. -ÍN / -INA, -INO / -INA, -IÑO / -IÑA

Como *-ico* e *-ito*, *-ín* presenta matices favorables (Gooch 1970: 10), aunque hay autores que le atribuyen un valor peyorativo moderado (Lang 1990: 107; Kaplanová 2010: 33). Es la variante regional típica de la zona noroccidental de España (RAE 2009: 631). Nández Fernández (1973: 89) cita un fragmento de la obra de Martínez de Moretín *Estudios Filológicos*, de 1857, para explicar el uso de este sufijo, bastante proclive a la lexicalización: «Los diminutivos terminados en in (sic) no son muchos, y con frecuencia expresan lo contrario de lo que indica la palabra *primitiva*. Espada y espadín representan objetos que el arte distingue bien; y lo mismo sucede con peluca, y peluquín».

Lang (1990: 101) opina que *-iño/-a* e *-ino/-a*, que a veces se consideran sufijos independientes, son variantes alomorfas de *-ín*, que vienen determinadas más por usos regionales que por la forma de la base. Según Kaplanová (34) «*-Ín/-ina* tiene una mayor representación en León (Coruña, Armental, el Bierzo) mientras que su forma *-ino/-ina* se extiende por Extremadura y Andalucía (Huelva, Sevilla, Córdoba)». Nández Fernández (1973: 69) localiza este último también en Asturias y Salamanca. *-Iño/a*, por su parte, es típico del español hablado en Galicia y proviene de la influencia del gallego (RAE 2009: 631).

Ejemplos: chiquitín, tontín / chavalina, pequeñino / besinho, dineriño

Ejemplos de términos lexicalizados: botín, bombín, plumín / neblina, padrino

[word=".*ín|.*(i|í)n(o|a)(s?)"]; [word=".*(i|í)ñ(o|a)(s?)"];²⁰³

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
mentulina	1	0,3	poquitín	2	0,4
poquitín	1	0,3			
Total	2	0,5		2	0,4

Tabla 126 - Sufijos *-ín*, *-ino* e *-iño* en ES y TRAD

²⁰³ Estas dos expresiones regulares podrían haberse agrupado en una ([word=".*ín|.*(i|í)(n|ñ)(o|a)(s?)"];), pero la realidad es que en un principio realizamos las búsquedas por separado.

Lexicalizadas ES: banderín, batín, boletín, borrachín, botellín, botiquín, botín, cajetín, camarín, corbatín, figurín, folletín, fortín, fútbolín, langostino, maletín, neblina, orquestina, papelina, peluquín, pizarrín, polvorín

Lexicalizadas EN: batín, boletín, botiquín, botín, cajetín, calcetín, chalinas, figurín, listín, maletín, marquesina, neblina, papelinas, chocolatina, hornacina

Evidentemente, las cifras recogidas en la tabla no son suficientes para extraer conclusiones, más allá del hecho de que estos tres sufijos, o variantes, son escasamente utilizados tanto en español original como traducido, al menos en su calidad de apreciativos. Sin embargo, los términos lexicalizados son numerosos, lo que indica que antaño fueron bastante más productivos pero, como hemos visto en la introducción de este apartado, su uso se circunscribe ahora a determinados ámbitos regionales.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	2	3,5	Cornwell	2	1,8
Giménez	0	0	Elroy	0	0
Madrid	0	0	Grafton	0	0
Martínez	0	0	Highsmith	0	0
Núñez	0	0	Leonard	0	0
Vázquez	0	0			

Tabla 127 - Distribución de los sufijos *ín*, *-ino* e *-iño* por texto

Las dos palabras que encontramos son *mentulina*, diminutivo de *méntula* (palabra no recogida en el DRAE que en latín significa *pene*), y el cuantificador *un poquitín*, que en TRAD viene de la traducción de *a little (bit)* y *a tad*, dos expresiones que sirven para indicar «cantidad pequeña». Del sufijo *-iño/a* no hemos hallado ningún ejemplo.

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	1
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	1
Información contextual	
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
∅	
Total	2

Tabla 128 - Elementos activadores de los sufijos *-ín*, *-ino* e *-iño* en traducciones

6.3.1.6. -ITO / -ITA

Según la Real Academia, es el sufijo diminutivo más extendido actualmente en todo el mundo hispánico, tras haber desbancado a *-illo* después del siglo XV (Prieto: 57), y algunos autores afirman que en Latinoamérica ha prácticamente desplazado a los demás. Se asocia con Castilla pero es el diminutivo menos marcado regionalmente (Lang 1990: 103). Es muy productivo pero también se presta a abundantes lexicalizaciones.

Como adelantamos en la descripción de *-illo*, este sufijo suele utilizarse para transmitir matices favorables y afectuosos (Gooch 1970: 6; Santos Muñoz 2006: 37). Añade Gooch un uso enfático o aumentativo en expresiones como «¡qué añito!» o «¡vaya tiempesito!». Santos Muñoz menciona también el matiz superlativo o de insistencia, hallado en frases como «hazlo rapidito» o «la carne me gusta bien fritita», en las que *rapidito* equivaldría a muy rápido y fritita a *muy frita*. Comenta esta autora también el matiz halagador, presente en la exclamación «¡qué zapatitos más monos!», que busca resaltar las cualidades del sustantivo, pero que fácilmente puede dar lugar a una interpretación irónica o despectiva si se modifican o eliminan ciertos elementos del contexto, como el adjetivo *monos*, o simplemente la entonación. Alonso, por su parte, comenta que este sufijo puede llegar a ser despectivo y rebajador cuando se aplica «al enemigo», como en «¡sermoncitos a mí!», un uso que pretende rebajar al contrario o al obstáculo que se nos presenta y que era frecuente antiguamente con nombres propios, pues expresa familiaridad impertinente (1982: 167).

Ejemplos: pueblecito, calentita, poquito, tacita

Ejemplos de términos lexicalizados: centralita, cabrito, pajarita, pepita

[word=".*(iî)t(a)o(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
agujeritos	1	0,3	amiguita/o	3	0,6
amiguito/a(s)	6	1,6	amorcito	1	0,2
angelitos	1	0,3	angelito(s)	5	1
animalito	1	0,3	armarito	4	0,8
arbolitos	1	0,3	bandejita	1	0,2
Arielito	1	0,3	barquito	1	0,2
arito	1	0,3	blanquita/o(s)	6	1,2
armarito(s)	6	1,6	bocadito	1	0,2
arruguitas	1	0,3	bolita	1	0,2
bajito/a(s)	5	1,4	bolsita	3	0,6

barquitos	2	0,5	bragueta	1	0,2
besito	1	0,3	burrito	5	1
bigotito	2	0,5	cabecita	2	0,4
bocaditos	1	0,3	cabezadita	1	0,2
bolsita	5	1,4	cadenita	1	0,2
bolsito	3	0,8	cajita	4	0,8
bombillitas	1	0,3	cajoncito	1	0,2
boquita	3	0,8	calladitos	1	0,2
borreguito	1	0,3	cariñito	1	0,2
botecito	1	0,3	carrerita	2	0,4
botellita(s)	4	1,1	carrito(s)	32	6,5
burguesitos	1	0,3	carterita	1	0,2
cachetitos	1	0,3	casita(s)	8	1,6
cafecito	1	0,3	chorritos	1	0,2
cafetito	1	0,3	cochecito(s)	2	0,4
cajita	8	2,2	conejo/as	2	0,4
cajoncito	1	0,3	copita	3	0,6
calentito	1	0,3	copitos	1	0,2
calladito/a(s)	3	0,8	corderito	1	0,2
calorcito	1	0,3	cortecitos	1	0,2
carita	1	0,3	cosita	3	0,6
carrito	7	1,9	credito	1	0,2
cartoncito	1	0,3	cuartito	2	0,4
casita(s)	8	2,2	cubanito	21	4,3
cervecita	2	0,5	despacito	2	0,4
chalecito	4	1,1	espejito	1	0,2
charquito	1	0,3	estrellitas	1	0,2
chispitas	2	0,5	frasquito	1	0,2
chuletitas	1	0,3	galletitas	2	0,4
chulito	1	0,3	gatita/o(s)	13	2,7
churrillos	1	0,3	golpecito(s)	9	1,8
cochecito(s)	2	0,5	gordito	1	0,2
collarcito	1	0,3	gorrita	2	0,4
copita(s)	6	1,6	gotitas	2	0,4
corazoncito(s)	2	0,5	guapito	2	0,4
corderitos	1	0,3	heridita	1	0,2
cosita(s)	2	0,5	hermanita/o(s)	3	0,6
costillitas	1	0,3	hijito	1	0,2
coñitos	2	0,5	hojita	1	0,2
cristalitos	1	0,3	jovencita/o(s)	5	1
cuadernito	1	0,3	jueguecitos	2	0,4
cuadritos	8	2,2	jugueteito	1	0,2
cuidadito	3	0,8	lamparita	1	0,2
culitos	1	0,3	librito	1	0,2
cunita	2	0,5	lucecita(s)	7	1,4
curritos	1	0,3	manita	1	0,2
dedito	1	0,3	mantelitos	1	0,2
delgadita	1	0,3	maridito	1	0,2
despacito	1	0,3	mayorcito	1	0,2
duritos	1	0,3	mesita*	31	6,3
empresita	1	0,3	montoncito	1	0,2
escenitas	1	0,3	morenita	1	0,2
figuritas	1	0,3	morritos	1	0,2
finito	2	0,5	muchachito	2	0,4
golpecitos	3	0,8	mujercita	1	0,2
gordito	3	0,8	navajita	2	0,4
gorrita	3	0,8	negrita/os	3	0,6
grititos	2	0,5	neverita	1	0,2
grupito	4	1,1	niñita/os	2	0,4
gusanito	1	0,3	numerito(s)	7	1,4
hermanita(s)	2	0,5	palmadita(s)	7	1,4
hilito	2	0,5	papaíto	1	0,2
hombrecito	1	0,3	pantaloncitos	1	0,2
huequecito	1	0,3	papelito	1	0,2
igualito	1	0,3	paseíto	1	0,2
invasioncita	1	0,3	paquetitos	1	0,2
jardincito	1	0,3	pataditas	1	0,2

Javierito	1	0,3	pedacito(s)	3	0,6
jovencita/o(s)	6	1,6	pelito	1	0,2
librito	1	0,3	perita	1	0,2
lonchitas	1	0,3	perrita/o	2	0,4
Lucecita	4	1,1	piececitos	1	0,2
Luisito	1	0,3	pistolita	1	0,2
machito	1	0,3	platito	3	0,6
mamadita	1	0,3	poquito	2	0,4
mamaíta	1	0,3	pueblecito(s)	4	0,8
manita(s)	4	1,1	ramita	6	1,2
maquinita	1	0,3	ramito	2	0,4
masajito	1	0,3	ratita	1	0,2
mayorcita/os	2	0,5	ratito	3	0,6
medallita	1	0,3	regalitos	1	0,2
mejorcito	1	0,3	rellenita	1	0,2
melenita	1	0,3	risita(s)	15	3,1
mesita(s)*	14	3,8	rodajita	1	0,2
minutitos	5	1,4	rollitos	2	0,4
montoncito(s)	2	0,5	rubita/o	5	1
muchachito	1	0,3	ruidito	1	0,2
mujercitas	1	0,3	salita	4	0,8
mulatita	1	0,3	saloncito	1	0,2
negrita/os	2	0,5	sobrecito(s)	3	0,6
niñito	1	0,3	soldaditos	1	0,2
numerito(s)	9	2,4	sonrisita	2	0,4
obrita	1	0,3	tacita(s)	3	0,6
oscurita	1	0,3	tiendecitas	1	0,2
osito	1	0,3	toquecito	1	0,2
ovejita	1	0,3	trabajito (s)	8	1,6
pajarito(s)	2	0,5	traguito	1	0,2
palillitos	1	0,3	trocito	2	0,4
palmadita(s)	4	1,1	tubito	1	0,2
pantaloncitos	2	0,5	varita	1	0,2
papelito	4	1,1	vasitos	2	0,4
paquetito(s)	3	0,8	velitas	1	0,2
paradita	1	0,3	venitas	1	0,2
partidita	1	0,3	viejecita	1	0,2
pasitos	1	0,3	vocecita	2	0,4
pastelito	1	0,3			
pastillita(s)	3	0,8			
pataditas	1	0,3			* Sin contar «mesita de noche»
pechitos	1	0,3			
pedacito(s)	4	1,1			
pelito	1	0,3			
pequeño/a(s)	4	1,1			
pobrecito	3	0,8			
poquita/o	8	2,2			
pueblecito	1	0,3			
puentecitos	1	0,3			
puntita	1	0,3			
puñadito	1	0,3			
quietecito	1	0,3			
rarita	1	0,3			
ratito	4	1,1			
rayitas	3	0,8			
rayito	1	0,3			
regalito	1	0,3			
risita(s)	4	1,1			
ruiditos	1	0,3			
sabanitas	1	0,3			
salita	3	0,8			
saloncito	1	0,3			
saltitos	1	0,3			
saquito(s)	7	1,9			
semanita	1	0,3			
sentadito	1	0,3			
silloncitos	1	0,3			

Silvita	2	0,5		
solita/o	7	1,9		
soniditos	1	0,3		
sonrisita(s)	3	0,8		
sorbitos	2	0,5		
sueldecito	1	0,3		
tijeritas	1	0,3		
tirito	1	0,3		
trabajito(s)	4	1,1		
traguito	1	0,3		
trocito(s)	3	0,8		
vasito	2	0,5		
verduritas	1	0,3		
vueltecita	1	0,3		
Total	342	93,2	344	70,4

Tabla 129 - Sufijo *-ito* en ES y TRAD

Lexicalizadas ES: bonito, cabrito, chiringuito, chorlito, cubito, gomitas, mosquita (muerta), nevadito (tipo de droga), pajarita, palomitas, pepitas, señorita/o, varita, mesita (de noche)

Lexicalizadas EN: bonita/o, burrito (comida), centralita, cordita, gallinita (ciega), gallito, manguitos, mariquita, mesita (de noche), mosquita (muerta), negrita (letra), pajarita, palomitas, perrito (caliente), señorita/(s), tiritita, tortitas, tortolitos

Lo primero que llama la atención de la tabla es que la suma de todas las apariciones de este sufijo son prácticamente las mismas en ambos subcorpus (342 en ES y 344 en TRAD) y, sin embargo, la columna de ES (158 celdas) es bastante más larga que la de TRAD (115 celdas). En el subcorpus de originales hay, por tanto, mayor variedad de bases que dan lugar a diminutivos (43 más). Las cifras totales normalizadas nos dicen que hay también una diferencia de frecuencias si analizamos los subcorpus teniendo en cuenta el número de palabras de cada uno, y el test estadístico nos confirma que esta diferencia es significativa (LL +13,10 $p < 0,001$).

En los originales españoles hemos encontrado multitud de nombres propios entre los diminutivos con *-ito* (en total, 119 casos de Alvarito, Amparito, Antoñito, Arielito, Arturito, Carmencita, Currito, Javierito, Laurita, Luisito, Luquitas, Merceditas, Mudarrito, Paquito, Raquelita, Ricardito, Silvita y Terecita). Sin embargo, solo hemos dejado en el listado unos pocos; aquellos que se utilizaban como apreciativos en contextos específicos, y no los que constituyen simplemente la manera habitual de llamar a un determinado personaje. En TRAD hay muchos menos casos de hipocorísticos (11 en total: Anita y Juanita). Sin embargo, hemos tenido dudas con «el cubanita», puesto que es un apelativo que se repite muchas veces, pero no llega a ser un nombre propio: siempre aparece en la narración en tercera persona y es traducción de

«the Cuban bus boy», una manera algo despectiva de referirse sistemáticamente a alguien, aunque esos sean su trabajo y su procedencia, especialmente si se conoce su verdadero nombre. Finalmente hemos decidido contarlo (aparece 21 veces en TRAD), aunque, de haberlo eliminado, la diferencia entre subcorpus habría sido más significativa. No hallamos demasiados casos de términos lexicalizados; este sufijo es el más productivo de los diminutivos que hemos analizado.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Madrid	90	217,2	Highsmith	91	92,6
Núñez	70	164,8	Leonard	72	92,5
Martín	56	99,6	Ellroy	69	62,9
Giménez	49	57,7	Cornwell	56	51
Vázquez	46	52,4	Grafton	56	60,3
Martínez	31	57			

Tabla 130 - Distribución del sufijo *-ito* por texto

En ES hallamos representadas todas las funciones descritas por las gramáticas: cariño, humor, atenuación, intensificación, ironía, desprecio o simple disminución (los ejemplos están en ese orden):

<Martín>: Y mi padre aguantando , aguantando . Siempre sonriente . Siempre recuerdo a mi padre sonriente , con <chispitas> en los ojos , « ¡eh , machote , vamos a jugar ; eh , machote , a ver si me pillas , chuta !

<Martínez>: Entonces , dio una orden : - ¡Rosario , ven ! Como un <borreguito> , obediente y de trote alegre , la aludida compareció en el escenario y se quedó muda.

<Madrid>: Inchausti llegó al bar pasadas las dos de la madrugada , cuando yo me había bebido dos <botellitas> de vino de Colmenar de Oreja .

<Núñez>: - Hola , Víctor - saludé . Levantó la cabeza y me miró de arriba abajo . Y ése fue su saludo . <Igualito> que un asno .

<Núñez>: . Si no conoces a ninguno . hazte cuenta de que te ha tocado la lotería . - Ya me imagino que no son <angelitos> .

<Martín>: todas sus compañeras dirán « se fue con un tío así y asá , <bajito> , mal afeitado , muy feo , con cara de mono » .

<Madrid>: había una maqueta que reproducía una ciudad en escala , con bloques de viviendas , calles , plazas , <arbolitos> , coches y gente paseando.

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	117
Sufijo apreciativo	24
Adjetivo o adverbio con connotación apreciativa + sustantivo	81
Información contextual	34
Tono general del fragmento (razones estilísticas o compensación)	11
Equivalente funcional	30
Información distribuida en más de un elemento lingüístico	29
Ø	18
Total	344

Tabla 131 - Elementos activadores del sufijo *-ito* en traducciones

A continuación, resumimos lo que hemos encontrado con respecto a los elementos que activan la utilización de este sufijo en las traducciones.

La categoría que mejor representada está es la de palabras que presentan connotaciones apreciativas sin necesidad de afijos o modificadores, porque representan objetos muy pequeños (*cottage, sprig, scrap*) o acciones muy breves (*pat, dash, tap*), o son inherentemente despectivas (*ofay, snicker*) o cariñosas (*sweetheart*). Estas palabras pueden activar o no el uso de diminutivos; recordemos que el traductor puede tener a su disposición equivalentes en su lengua que no lo requieran.

En segundo lugar se encuentra la categoría que, en principio pensábamos que sería la que en más ocasiones diera lugar al uso de diminutivos: la adición de adjetivos o adverbios del campo semántico de lo pequeño o breve, como *little* (*tray, wife, while, plump*), *small* (*boat, stars, injury, voice*), *tiny* (*ranch, head, tomboy, town*), *young* (*woman, man, boy, kid*), *low* (*dresser, table*), *kid* (*brother*), *baby* (*donkey*) y también de lo afectuoso o despectivo (*sweet boy, silly laughter*). Encontramos además, dos ejemplos de *real* para enfatizar los adjetivos *slow* y *short*.

A continuación, proporcionamos ejemplos de las categorías restantes:

- Información distribuida en más de un elemento lingüístico: en el original encontramos la información separada en dos o más palabras, generalmente, un adjetivo y un sustantivo. El diminutivo une la información de ambas en una única palabra. Ejemplos: *white ho'* (blanquita), *white kid* (blanquito), *cuban busboy* (cubanito), *colored toddlers* (negritas), *black kids* (negritos), *pussy gun* (pistolita) o *mousy little girl* (ratita).

- Utilización de un equivalente funcional: *count my blessings* (soñar con los angelitos), *reside in Snooze City* (echar una cabezadita), *easy now* (despacito), *fresh napkins* (mantelitos limpios), *striptease* (numerito).
- El contexto nos proporciona la información para entender por qué el traductor utiliza el diminutivo: en este caso, Moya utiliza *cajita de costura* para *mending kit* por continuar con la denominación que se ha elegido la primera vez, en la que se denominaba a la cajita *little mending kit*, aunque ahora en el original ya no aparezca el adjetivo *little*.

<Grafton>: Saqué del bolso la <cajita> de costura .

-->par_en_barbara: I took the complimentary **mending kit** from my handbag .

Aquí el traductor entiende que se está hablando de un acto irregular aunque en el original se usen los más neutrales *job* o *work*.

<Grafton>: Todavía tengo en la cabeza un chichón del tamaño de un puño . Y todo por un <trabajito> por el que ni siquiera me pagaron .

-->par_en_barbara: And all this for a **job** for which I did n't even get paid !

En este caso no es necesario el diminutivo porque no se utiliza vocabulario especialmente afectuoso, infantil o humorístico en el fragmento, pero, dado que se está hablando de una señora mayor que elige tomar una copa de helado de características bastante infantiles, su uso es una elección estilística justificada que aporta un toque humorístico a la descripción.

<Highsmith>: Una gran bola de helado de chocolate simulando un <conejito> con dos orejas de barquillo , y un generoso brochazo de crema batida entre las mismas

-->par_en: A great globe or scoop of chocolate ice cream with two ladyfingers perched like **rabbit's** ears upon it , a generous daub of whipped cream between the ears

- También, como adelantaban Rojo y Schneider, existen algunos diminutivos en inglés que pueden ser un elemento activador: *baggie* (*bolsita*), *bunnies* (*conejitas*), *kitten* (*gatita*), *fats* (*gordito*), *daddy* (*papaíto*) o *hubby* (*maridito*).

- A veces puede aparecer un diminutivo donde no hay elemento activador en inglés para compensar si no se ha conseguido utilizar donde correspondía, o incluso si en ciertos fragmentos el tono general del original no se ha logrado transmitir por falta de otros recursos en español:

<Cornwell>: Una nevisca mezclada con lluvia cubría la hierba y las paredes con <gotitas> de hielo .

-->par_en Sleet mixed with rain beaded grass and walls with ice .

<Leonard>: Se le ocurre disparar a alguien , a otro poli , se fuga y nosotros nos quedamos con el <papelito> .

-->par_en: " You're talking about a guy , " Max said , " who's the highest kind of risk that he'll take off , who's a threat to the community.. . He shoots somebody , another cop , he runs and we're holding his paper .

- La última categoría corresponde a aquellos casos en los que no hemos logrado encontrar ningún elemento lingüístico, estilístico o contextual concreto que justifique el uso del diminutivo, aunque no por ello queremos decir que su uso no sea correcto ni adecuado, puesto que puede formar parte del estilo del traductor o del concepto o imagen mental que posee o se ha formado acerca del objeto o la idea representados.

<Leonard>: Por eso tenían agentes de seguridad que iban dando vueltas con sus <cochecitos> blancos para ayudar .

-->par_en: It was why they had security guys driving around in those white utility cars , GMC Jimmys , to help you out .

<Cornwell>: Se sentó en el borde de la chimenea y sacó su <navajita> del ejército suizo .

-->par_en: He sat on the edge of the hearth and got out his Swiss Army knife .

6.3.1.7. -UCO / -UCA

Comenta Gooch (1970: 28) que este sufijo se asocia a Santander (a toda Cantabria, según la RAE), «a province known indeed to her sons as *la tierra*». En otras regiones puede considerarse a veces despectivo.

Ejemplos: mesuca, casuca, beatuco

[word=".*(u|ú)c(o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
mujeruca	2	0,5	ventanuco	5	1
Total	2	0,5		5	1

Tabla 132 - Sufijo *-uco* en ES y TRAD

Lexicalizadas ES: —

Lexicalizadas EN: —

La palabra hallada en TRAD, *ventanuco*, aparece definida en el DRAE como «ventana pequeña y estrecha». En principio es un término lexicalizado, pero el significado es justamente el de disminuir la base, por lo que lo hemos incluido en el cómputo. De todas formas, lo consideremos válido o no, las frecuencias son muy bajas y las diferencias entre subcorpus no son significativas de ninguna de las dos maneras.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	2	3,5	Grafton	5	5,4
Giménez	0	0	Cornwell	0	0
Madrid	0	0	Ellroy	0	0
Martínez	0	0	Highsmith	0	0
Núñez	0	0	Leonard	0	0
Vázquez	0	0			

Tabla 133 - Distribución del sufijo *-uco* por texto

El único ejemplo de este sufijo que hemos encontrado en ES es el de Andreu Martín, *mujeruca*. El contexto nos dice que este término está utilizado de manera despectiva («una mujeruca patética, una de esas ancianas amargadas, petrificadas ante el televisor»), lo que tiene sentido si consideramos que el autor es catalán, no cántabro.

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	
Información contextual	1
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	2
Información distribuida en más de un elemento lingüístico	
∅	2
Total	5

Tabla 134 - Elementos activadores del sufijo *-uco* en traducciones

A pesar de corresponder todas a un único término, *ventanuco*, hemos clasificado los elementos activadores del texto original en tres categorías diferentes. Las dos primeras veces que aparece, su uso no parece estar activado por nada en el original:

<Grafton>: Ray llamó a la puerta , cuyo <ventanuco> estaba cruzado por barrotes de adorno . Por el **ventanuco** vimos la luz que salía de una habitación

-->par_en: Ray tapped on the door , which had ornamental burglar bars across the **glass** . Through the **window** I could see bright light pouring from a room on the left

Las dos siguientes apariciones podrían venir dadas por la descripción que se da en el texto de dichas ventanas y el equivalente que Moya posee en su diccionario mental para las características que se mencionan:

<Grafton>: mal iluminado y de un metro ochenta de altura , bordeado por la parte superior por una serie de <ventanucos> que se abrían con manivela ; algunos estaban abiertos y por ellos entraba luz artificial

-->par_en: Before me was a dimly lighted , musty passageway about six feet high , rimmed along the top with a continuous series of hand-cranked **window panels** , some of which were standing open , admitting artificial light .

<Grafton>: y apenas daban luz . En la circunferencia que trazaban los bordes del techo , la cenefa continua de <ventanucos> con cristal de espejo ocultaba el pasadizo en que me encontraba --

>par_en_barbara: Around the periphery , up near the ceiling where I was , the continuous rim of mirrored-glass **windows** concealed the space where I hid .

El último caso es simplemente una utilización de la información contextual aparecida previamente, dado que estamos hablando de las mismas ventanas para las que ya se ha elegido denominación:

<Grafton>: Abrí el bolso a la luz que entraba por los <ventanucos> y saqué la billetera .

-->par_en: By the light coming through the **windows** , I hunkered down and opened my bag , taking out my wallet .

6.3.1.8. -UELO / -UELA

Como *-ico*, era bastante más utilizado antiguamente (Gooch 1970: 25), pero es más productivo que *-ete* o *ín*, pues no está restringido por regionalismos (Lang 1990: 106-107). Posee un matiz despectivo bastante marcado; de hecho, Lang (1990: 101)

comenta que a veces suele agrupárselo con los sufijos peyorativos. Sin embargo, aclara este autor que se trata de un matiz leve que puede dar lugar a un tono amistoso o humorístico.

Náñez Fernández (1973: 89), citando una vez más a Martínez de Morentín, comenta sobre dicho valor peyorativo: «Los diminutivos acabados en uelo, uela, zuelo, zuela, ezuelo, ezuela, indican también el *poco valor* de las cosas, y el poco caso que se hace de las personas; y con frecuencia el *soberano desprecio* con que las miramos».

La Academia habla de los usos expresivos que pueden tener en palabras como *ojuelos* (ojos vivos o risueños) y de los matices atenuadores o de indulgencia que pueden añadir a significados como el de *ladrón* en *ladronzuelo* (ladrón de poca monta) (RAE 2009: 653).

Ejemplos: escritorzuelo, chicuela, mozuelo

Ejemplos de términos lexicalizados: pañuelo, cajuela, castañuela

De origen latino: abuela

[word=".*uel(o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
callejuela	1	0,3	callejuela(s)	3	0,6
gordezuelo	1	0,3	plazuela	1	0,2
jovenzuelo/as	3	0,8	polluelos	1	0,2
mujerzuelas	1	0,3	portezuela	4	0,8
polluelo	1	0,3	riachuelo	2	0,4
Total	7	1,9		11	2,2

Tabla 135 - Sufijo *-uelo* en ES y TRAD

Lexicalizadas ES: castañuelas, lentejuelas, morteruelo, pañuelo, portezuela (de coche)

Lexicalizadas EN: cazuela, hojuelas, hoyuelo, lentejuelas, pañuelo, portezuela (de coche), tachuelas

Las diferencias entre subcorpus no llegan a ser significativas. La duda más grande que hemos tenido en este grupo se refiere a *portezuela*, dada su elevada frecuencia de aparición en TRAD. Aparece recogida en el DRAE con el significado específico de «puerta de carruaje». En la mayor parte de los casos de ambos subcorpus se utiliza *portezuela* como puerta de vehículo pero, evidentemente, de automóviles modernos. Dado que no se está valorando de ninguna manera dicha puerta cuando se la

denomina *portezuela*, sino que, al parecer, ha adquirido el significado específico de «puerta de coche», no la contaremos cuando esté utilizada de esta manera, pero sí cuando se refiera a otro tipo de *puerta* o se la califique de manera diminutiva o despectiva, por ejemplo, lo que elimina prácticamente todos los casos (4 en ES y 34 en TRAD), a excepción de cuatro en TRAD:

<Grafton>: había empotrada una caja fuerte de pequeño tamaño . La <portezuela> de la caja tendría cuarenta centímetros de anchura por treinta y cinco de altura .

-->par_en: The safe door was approximately sixteen inches wide and fourteen inches tall .

<Grafton>: Abrió las <portezuelas> que había debajo del fregadero y vi que arrugaba el entrecejo .

-->par_en: She opened the cabinet under the sink , and I saw a frown cross her face .

Mujerzuela aparece en el DRAE con el sentido de «mujer de poca estima, perdida o de mala vida». Dado que el sentido recogido es apreciativo (despectivo) no consideramos que dicha acepción tenga un sentido diferente al que aporta el sufijo y, por tanto, incluimos la palabra en el cómputo. Por los mismos motivos incluimos también *riachuelo*.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Núñez	3	7,1	Grafton	5	5,4
Vázquez	2	2,3	Ellroy	3	2,7
Giménez	1	1,2	Highsmith	2	2
Martín	1	1,8	Cornwell	1	0,9
Martínez	0	0	Leonard	0	0
Madrid	0	0			

Tabla 136 - Distribución del sufijo *-uelo* por texto

El sentido que este sufijo aporta es diminutivo en *jovenzuelo*, *polluelo*, *plazuela*, *portezuela* y *riachuelo*, y humorístico o despectivo en *gordezuelo* y *mujerzuela* (ambos en ES). Encontramos ambas connotaciones en la única aparición de *callejuela* en ES:

<Núñez>: Me interné en una <callejuela> . Es estrecha y casi siempre está en la oscuridad . Los drogas y las parejas se encargan de romper las farolas a pedradas .

En TRAD, *callejuela* se utiliza cuando hay en el original una indicación de estrechez, pequeñez o importancia secundaria (*little street*, *narrow street*, *side street*).

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	3
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	3
Información contextual	2
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	2
∅	1
Total	11

Tabla 137 - Elementos activadores del sufijo *-uelo* en traducciones

6.3.2. Aumentativos

6.3.2.1. -AZO / -AZA

Este sufijo es el más ambivalente de todos los aumentativos, pues puede connotar positivamente o negativamente en la misma medida, dependiendo del contexto o del contenido semántico de la base a la que se aplique (RAE 2009: 658; Gooch 1970: 5, 12): compárese un *plomazo* con un *cuerpazo*.

Lang (1990: 112-113) lo considera el más explícito e hiperbólico de los aumentativos, y afirma que es muy productivo y en Sudamérica está ganando terreno a los otros sufijos de este grupo. En algunas regiones denota accidentes (*trenazo*, *avionazo*) o, aplicado a nombres propios y de lugar, acciones políticas (*Cordobazo*, *Fujimorazo*). Más recientemente, se han acuñado términos con el significado de protesta popular (*cacerolazo*) o medidas políticas drásticas (*tarifazo*), un uso que se ha extendido también en Europa (Rainer 2010: 29-34).

Estos valores provienen probablemente del sentido de acción o resultado de una acción que posee el sufijo (Gooch 1970: 12). Transmite especialmente la idea de golpe, real o metafórico (p. ej. *hachazo*, *telefonazo*, *frenazo*) y los derivados pueden crearse a partir del objeto con el que se da el golpe o del que lo recibe (*botellazo* o *cogotazo*) (Lázaro Mora 1999: 4673). Sin embargo, la Real Academia no acepta como apreciativos los derivados en *-azo* que se refieren a «nombres de golpe, sonido o movimiento brusco» (2009: 659); solo lo son si designan cosas de gran tamaño (por lo tanto, palabras como *cañonazo* pueden ser apreciativas o no, dependiendo de si nos referimos al objeto o al disparo). Prieto (2005: 59) opina que esta diferencia de significados podría

constituir un fenómeno de homofonía si lo analizamos de manera sincrónica, pero no tanto si lo hacemos de manera diacrónica. Como ya sabemos, en opinión de la Real Academia, ningún apreciativo debería cambiar el género de su base. Sin embargo, al parecer, existen ciertos casos polémicos, como *multazo*, *broncazo*, *gripazo*, *latazo*, *pestazo* o *trancazo*, que pueden interpretarse como construcciones asimiladas a los nombres de golpe o como excepciones a la regla general (RAE 2009: 659).

Por otra parte, existen algunos derivados en *-azas* o *-azos* que se utilizan para referirse a alguien de manera bastante peyorativa. Lang (1990: 113) los considera variantes alomorfas de *-azo* utilizadas en expresiones más actuales, como por ejemplo, *bocazas*, *manazas* o *calzonazos*, que, según Lázaro Mora (1999: 4673), se traducen respectivamente por «el que habla demasiado y fanfarronea», «persona poco hábil» y «hombre muy condescendiente». Nosotros incluiremos estos sustantivos por el interés que pueden tener en el estudio de las diferencias entre español original y traducido, marcándolos con un asterisco (*).

Ejemplos: golazo, acentazo, manaza, ojazos

Ejemplos de términos lexicalizados: barcaza, carnaza y melaza.

[word=".*az(o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
bigotazo(s)	2	0,5	bocazas*	2	0,4
billetazo	1	0,3	cabronazo(s)	4	0,8
bocazas*	1	0,3	gustazo	1	0,2
bolerazo	1	0,3	jefazo	2	0,4
chotaza	1	0,3	madraza	1	0,2
detallazo	1	0,3	manazas	3	0,6
gustazo	1	0,3			
jesuitazo	1	0,3			
locaza	2	0,5			
manaza(s)	3	0,8			
padraza	2	0,5			
propinaza	1	0,3			
taconazos	1	0,3			
talentazo	1	0,3			
topicazo	1	0,3			
Total	20	5,4		13	2,7

Tabla 138 - Sufijo *-azo* en ES y TRAD

Lexicalizadas ES: carnaza, coñazo

Lexicalizadas EN: barcaza, espinazo

Como vemos, hay más del doble de términos diferentes en ES que en TRAD, aunque la cantidad de apariciones en total no se diferencie tanto entre los dos subcorpus. Sin embargo, esta moderada diferencia es significativa en términos estadísticos (LL +4,17 $p < 0,05$).

Han quedado fuera de nuestro listado, sin embargo, la mayor parte de los derivados de *-azo*, puesto que su significado es el de golpe. También en este grupo hay mayor variedad de términos en ES (47 frente a 27), pero, en este caso, más cantidad de apariciones en TRAD (118 frente a 196), aunque 77 de estas últimas corresponden a una única palabra, *vistazo*. Si los sumáramos a los casos anteriores, la diferencia entre subcorpus pasaría a no ser significativa.

ES	Frec	TRAD	Frec.
alfilerazos	3	aguijonazos	1
arañazos	1	arañazos	9
bajonazo	1	balazo(s)	5
balazo	1	bocinazos	1
bandazo(s)	2	brochazo(s)	2
batacazo	2	cachiporrazo	2
bocinazo(s)	2	cambiazo	1
cabezazos	2	carpetazo	1
carpetazo	1	chivatazo(s)	7
cañonazo	1	codazo(s)	9
chispazo	1	escopetazo	1
chivatazo	1	espadazos	1
claxonazos	1	fogonazo	1
codazo(s)	5	manotazo(s)	5
cogotazo	1	esquinazo	1
derechazo	1	navajazo(s)	2
escobazo	1	patinazo(s)	3
fogonazo	1	pinchazo(s)	8
frenazos	1	porrazo(s)	2
golpetazos	1	portazo(s)	11
guantazo(s)	6	puñetazo(s)	20
jeringazos	1	rodillazo	3
latigazo(s)	2	salivazo(s)	2
manotazo(s)	8	tanganazos	1
martillazo	2	timbrazo(s)	18
mazazo	1	tortazo	1
navajazo(s)	8	vistazo(s)	77
palmetazo	1	zarpazos	1
papirotazos	1		
patinazo(s)	2		
pelotazo	4		
pinchazo	1		
pistoletazo	1		
plumazo	2		
porrazo(s)	2		
portazo	7		
puntillazo	1		
puñetazo(s)	15		
ramalazo	1		
rodillazo	1		
salivazo	1		
telefonazos	1		
timbrazo(s)	7		
topetazo	1		
tortazo	2		
trompazos	1		
vistazo(s)	5		
zarpazo	3		
Total	118		196

Tabla 139 - Derivados de *-azo* no incluidos

Hemos comprobado que la afirmación que realizábamos en la introducción de este sufijo respecto a que los valores pueden ser igualmente negativos o positivos, es acertada, puesto que hallamos prácticamente la misma cantidad de términos con connotación positiva que negativa y que simples aumentativos.

Negativos o despectivos: locaza, chotaza, bocazas, bolezazo, topicazo, cabronazo

Positivos: detallazo, gustazo, madraza, padrazo, propinaza, talentazo

Aumentativo neutro: bigotazo, billetazo, jesuitazo, manaza, taconazos, jefazo

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	7	8,2	Leonard	5	6,4
Martín	5	8,9	Ellroy	3	2,7
Núñez	3	7,1	Highsmith	3	3
Martínez	2	3,7	Cornwell	2	1,8
Vázquez	2	2,3	Grafton	0	0
Madrid	1	2,4			

Tabla 140 - Distribución del sufijo *-azo* por texto

En cuanto a los elementos utilizados en inglés, encontramos, como en los diminutivos, varios equivalentes funcionales, como *mother hen* (madraza) o *full of shit* (bocazas), adjetivos de sentido aumentativo, como *big* y *heavy* para *hands*, y palabras cuyo significado incluye la connotación de tamaño, aprecio o desprecio, como *motherfucker* (cabronazo) o *brass* (jefazo). Este último lo hemos clasificado en la categoría anterior porque está modificado por los adjetivos *heavy* y *high*.

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	4
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	4
Información contextual	1
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	4
Información distribuida en más de un elemento lingüístico	
∅	
Total	13

Tabla 141 - Elementos activadores del sufijo *-azo* en traducciones

6.3.2.2. -ÓN / -ONA

Gooch (1970: 10) lo considera el sufijo aumentativo más versátil, con tendencia a formar sustantivos y adjetivos de connotación peyorativa (*matón*, *respondón*), aunque también puede reforzar cualidades positivas, como en *simpaticón* o *bonachón*. Sin embargo, comenta el autor que en estas últimas lo normal es percibir cierto matiz peyorativo denotado por el exceso de esa cualidad. A veces puede incluso atenuar el significado de ciertos términos negativos, aportándoles matices de familiaridad y restando severidad a la base (p.ej. *egoistón*, *torpón*). Santos Muñoz (2006: 54), sin embargo, afirma que «se trata de un sufijo cuyo principal valor es el de halagador, en muchas ocasiones mezclado con una importante carga de de cariño».

En opinión de Lang (1990: 111-114) *-ón* es el sufijo menos emotivo y con más lexicalizaciones; su productividad no es tan grande como la de *-azo* pero, como este, muestra un sentido meliorativo o peyorativo dependiendo de la base a la que se adjunte. También se le asemeja en su función de indicar acción repentina o resultado de la acción denotada por la base, como en *apagón* o *empujón*. Puede aportar sentido diminutivo a algunas bases, como *ratón*, *tapón* o *escalón*. Por último, mediante este sufijo se derivan adjetivos de bases verbales, como *mirón* o *preguntón* (o los dos ejemplos de Gooch que hemos mencionado al principio). Sin embargo, estos tres últimos grupos de derivados no son aceptados por la Real Academia: «aunque añaden matices expresivos diversos, modifican la categoría léxica de su base» (2009: 630).

Algunos autores se limitan a describir estas funciones sin posicionarse sobre si deberían considerarse apreciativos o no. Sin embargo, en Prieto (2005: 58) encontramos la siguiente afirmación:

The AUG *-ón* comes from Latin *-ō*, *-ōnis*, which in Latin attaches to nouns and verbs to designate people that are particularly characterized by some action, feature or habit, undesired generally. In Spanish, we still see this type of use in words such as *llorón*, *chillón* (cry+*ón* = person that cries too much) and *frentón* (forehead+*ón* = person with a big forehead). Thus, this type of function is still an AUG function.

Otras funciones que puede cumplir este sufijo según la Real Academia (2009: 657) son la de designar a personas de cierta edad (p. ej. *cuarentón*, *cincuentona*) o estatus marital (*solterón*) de manera jocosamente familiar o, a veces, despectiva; la de

expresar crítica, ironía o censura amables, como en *torpón*; y la de formar palabras a partir de partes del cuerpo, que pueden interpretarse como derivados despectivos (*narigón* sería una nariz grande) o como sustantivos de persona, no aumentativos, que designan al que posee esa parte del cuerpo anormalmente grande (*cabezón*). Nosotros incluiremos ambos sentidos, en sus formas sustantivas y adjetivas. Finalmente, hay que mencionar que las lexicalizaciones de este sufijo son dudosas en algunos casos. Por ejemplo, las palabras *dulzón* y *nubarrón* están en el DRAE, por lo que habría que considerarlas lexicalizadas, pero también figuran en la gramática de la RAE como voces claramente transparentes.

Ejemplos: patadón, novelón, fiestón,

Ejemplos de términos lexicalizados: salón, sillón, cinturón, portón

[word=".*ón|. *on(a|e)(s?)"];</p>
</div>
<div data-bbox="160 442 826 892" data-label="Table">
<table border="1">
<thead>
<tr>
<th>ES</th>
<th>Frec.</th>
<th>Frec. norm.</th>
<th>TRAD</th>
<th>Frec.</th>
<th>Frec. norm.</th>
</tr>
</thead>
<tbody>
<tr>
 azulones | 1 | 0,3 | caserón | 2 | 0,4 |
</tr>
<tr>
 barrigón/a(s) | 5 | 1,4 | cincuentón | 1 | 0,2 |
</tr>
<tr>
 bigotón | 1 | 0,3 | comodona | 1 | 0,2 |
</tr>
<tr>
 bobalicona | 1 | 0,3 | cuarentón(es) | 2 | 0,4 |
</tr>
<tr>
 cabezón | 2 | 0,5 | desgarrón | 1 | 0,2 |
</tr>
<tr>
 caderona | 2 | 0,5 | dulzón | 1 | 0,2 |
</tr>
<tr>
 cincuentón/a(s) | 5 | 1,4 | fortachón | 1 | 0,2 |
</tr>
<tr>
 coquetón | 1 | 0,3 | gigantón | 1 | 0,2 |
</tr>
<tr>
 cuarentón/a(s) | 8 | 2,2 | lanchones | 1 | 0,2 |
</tr>
<tr>
 culona | 1 | 0,3 | macetones | 1 | 0,2 |
</tr>
<tr>
 dulzón/a | 3 | 0,8 | notición | 1 | 0,2 |
</tr>
<tr>
 facilón | 1 | 0,3 | novelones | 1 | 0,2 |
</tr>
<tr>
 frivolona | 1 | 0,3 | nubarrones | 1 | 0,2 |
</tr>
<tr>
 hombión(es) | 2 | 0,5 | ricachonas/ón | 2 | 0,4 |
</tr>
<tr>
 lagrimones | 2 | 0,5 | sesentona | 1 | 0,2 |
</tr>
<tr>
 narizón | 1 | 0,3 | tristones | 1 | 0,2 |
</tr>
<tr>
 negrona | 1 | 0,3 | ventarrón | 1 | 0,2 |
</tr>
<tr>
 notición | 2 | 0,5 | zapatones | 1 | 0,2 |
</tr>
<tr>
 nubarrones | 2 | 0,5 | | | |
</tr>
<tr>
 panzón | 1 | 0,3 | | | |
</tr>
<tr>
 pastón | 2 | 0,5 | | | |
</tr>
<tr>
 patadón | 1 | 0,3 | | | |
</tr>
<tr>
 picarón | 1 | 0,3 | | | |
</tr>
<tr>
 pistolón | 3 | 0,8 | | | |
</tr>
<tr>
 reservón | 1 | 0,3 | | | |
</tr>
<tr>
 ricachón | 1 | 0,3 | | | |
</tr>
<tr>
 santón | 1 | 0,3 | | | |
</tr>
<tr>
 sesentón | 2 | 0,5 | | | |
</tr>
<tr>
 sofocón | 1 | 0,3 | | | |
</tr>
<tr>
 solterón/a(s) | 2 | 0,5 | | | |
</tr>
<tr>
 tiarrón(es) | 3 | 0,8 | | | |
</tr>
<tr>
 tripón | 1 | 0,3 | | | |
</tr>
<tr>
 vomitona | 1 | 0,3 | | | |
</tr>
<tr>
 vozarrón | 1 | 0,3 | | | |
</tr>
<tr>
 Total | **64** | **17,4** | | **21** | **4,3** |
</tr>
</tbody>
</table>
</div>
<div data-bbox="387 893 609 907" data-label="Caption">
<p>Tabla 142 - Sufijo -ón en ES y TRAD</p>
</div>
<div data-bbox="480 924 516 940" data-label="Page-Footer">
<p>443</p>
</div>

Lexicalizadas ES: almohadón, batallón, bodegón, boquerones, cabrón, cajón, callejón, camisón, cartón, cañón, cinturón, colchón, copón, faldones, fondona, frontón, garrafón, goterones, hampón, hormigón, jarrón, jergón, judiones, litrona, listón, mamón, mantón(es), maricón(a), matón, montón(es), morrón, moscón, mosquetón, piñón, portón, ratón, resopón, roscón, rosetón, salón, santurrón, sillón, tablón, tacón, tazón, terrón, tumbona

Lexicalizadas EN: arcón, balón, batallón, bonachón, buscona(s), cabrón(es), cagones, cajón(es), callejón, camisón, cartón, cascarón, cañón, chaquetón(es), cebones, cinturón(es), colchón(es), culebrón, escuadrón(es), espigón, espolones, faldón(es), fregona, frontón, galeones, hampón(es), hormigón, jarrón, jergones, listón(es), mamón(es), mantón, maricón, mariposón, matón(es), mechón(es), montón(es), mosquetones, piñón, plumón, portalón, ratón, respingona, resultón, salón(es), saltón, sillón(es), tablón(es), tacón(es), tazón(es), telón, torreón, tumbona

Como vemos, las lexicalizaciones son mucho más numerosas que en el caso de -azo/a. Aparecen dos apodos en TRAD, *Dulzón* (7) y *Grandullón* (21), que no hemos contado porque no se utilizan de manera apreciativa; son simplemente los moteles mediante los cuales los demás personajes se refieren a estas personas durante toda la historia. Aunque los contáramos, la diferencia entre subcorpus, que es estadísticamente significativa a favor del subcorpus de originales (LL +36,81 p<0,0001), seguiría siéndolo, aunque un poco menos.

Los términos que han quedado fuera de nuestro listado por ser sustantivos derivados de verbos y, por tanto, cambiar la categoría gramatical de la palabra — recordemos que la RAE no los acepta como apreciativos— son también numerosos. Los recogemos en la siguiente tabla:

ES	Frec	TRAD	Frec.
apretón(es)	4	borrones	7
bajón	1	burlona/es	7
bofetón	1	chillón/a(s)	13
borbotón(es)	3	dormilona	1
borrón	5	empujón(es)	7
burlón/a(s)	11	encerrona	5
chillona	1	figón	2
críticón	1	gruñona	1
empellón(es)	4	hinchazón(es)	3
empujón(es)	10	juguetona	1
encerrona	3	ligón	1
gritonas	1	llorones	2
llorón/as/es	4	madrugón	1
madrugón(es)	2	mirón(es)	101
mirón(es)	9	peleón	1
respondón(es)	2	pisotón(es)	3
retortijones	1	remendón	1

reventón	1	remolón	1
revolcón	1	resbalón	2
sacudión	1	restregones	1
saltones	4	retortijón(es)	2
sobón	1	retozona	1
soplón(es)	3	soplón/a(s)	20
temblonas	1	temblona	1
tirón(es)	8	tirón(es)	21
tragón(es)	2	tragón	1
trompicones	2		
zumbón	4		
Total	91		207

Tabla 143 - Derivados de -ón no incluidos

Si los contáramos, las diferencias dejarían de ser significativas (LL -0,93), pero estos resultados se verían sesgados por la gran cantidad de veces que aparece la palabra *mirón* en la traducción del texto de Ellroy (99 de las 101 apariciones y casi siempre como traducción de *peeper* o *peeping*).

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	14	24,9	Grafton	9	9,7
Vázquez	12	13,7	Ellroy	5	4,5
Giménez	10	11,8	Leonard	3	3,8
Núñez	11	25,9	Cornwell	2	1,8
Madrid	9	21,7	Highsmith	2	2
Martínez	8	14,7			

Tabla 144 - Distribución del sufijo -ón por texto

Los términos hallados parecen, en su mayor parte, intensificar matices negativos (*panzón*, *bobalicona*), aunque también hay ejemplos de lo contrario (*notición*, *pastón*). Sin embargo, no queda muy claro a partir de las concordancias qué origina el uso de este aumentativo en traducciones, dado que los elementos activadores están distribuidos de manera equilibrada en todas las categorías.

Encontramos palabras con significado aumentativo inherente (por ejemplo, *hulk* se traduce por *gigantón* y *gale* por *ventarrón*), casos de adición del adjetivo *big* (*big news* por *notición* y *big house* por *caserón*), información separada en varios elementos que el aumentativo unifica (tres de los cuatro casos corresponden a edades: *fifty-some*, *about forty*, *in his mid-forties*) o nada en concreto que motive la adición del sufijo:

<Grafton>: El aire olía al perfume acre de la pólvora y percibí al instante un sabor <dulzón> al final de la lengua .

-->par_en: The air smelled acrid , and I could suddenly taste something sweet at the back of my throat .

Bastantes casos son motivados probablemente por el contexto en el que aparece el objeto aludido y la imagen mental del mismo que se hace el traductor, o por el tono general del fragmento (en el segundo ejemplo, algo despectivo):

<Grafton>: Vi dos hileras de <macetones> con palmas a ambos lados de una pisoteada alfombra roja que conducía a recepción .

-->par_en: Two rows of potted fan palms were arranged on either side of a length of trampled-looking red carpet that heralded the path to the front desk

<Leonard>: Si no , los domingos , no ves más que <ricachonas> que sacan a sus chuchos para echar una meada .

-->par_en: Otherwise , Sunday , what you have mostly are rich ladies come out with their little doggies to make wee-wee .

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	3
Sufijo apreciativo	1
Adjetivo o adverbio con connotación apreciativa + sustantivo	3
Información contextual	4
Tono general del fragmento (razones estilísticas o compensación)	2
Equivalente funcional	1
Información distribuida en más de un elemento lingüístico	4
∅	3
Total	21

Tabla 145 - Elementos activadores del sufijo -ón en traducciones

6.3.2.3. -OTE / -OTA

Gooch considera que este sufijo presenta muchas similitudes con respecto a -ón, dada su connotación peyorativa y capacidad atenuadora (1970: 12). La Real Academia (2009: 660) está de acuerdo con ambas apreciaciones: «Este sufijo aúna a menudo el significado aumentativo al despectivo y aparece en los textos junto a otros calificativos de sentido marcadamente despreciativo (...) No obstante, recibe a menudo connotaciones atenuadoras en sustantivos y adjetivos» (p. ej. *brutota*, *feote*). Se diferencia de -azo en los matices de condescendencia o simpatía, a veces burlesca, que puede aportar a ciertos adjetivos como *sencillote* o *tranquilote*. También afirma la Real Academia que se prefiere este sufijo antes que -azo en México y Centroamérica.

Lang (1990: 111, 115) también comenta que -ote aúna el matiz aumentativo con rasgos de simpatía y humor —un ejemplo sería saludar a un niño con un «¡Hola,

muchachote!»—, pero la mayor parte de las veces se utiliza con sentido peyorativo. Según el autor, no suele producir lexicalizaciones (aunque la RAE sí que menciona algunas en la página 660 de su gramática, como *cabezota* en el sentido de *testarudo*, *machote* y *monigote*).

Como en el caso de *-etas*, *-azas* o *-ica*, existe la variante epicena *-otas*, que tiene valor despectivo al designar a alguien con algún defecto físico (p. ej. *un narizotas* o *un gafotas*). La RAE (2009: 487) comenta que en español europeo se registran ciertas formaciones con sentido despectivo a partir del acortamiento de ciertas bases y la adición del sufijo *-ota*, como *pasota* (de *pasar*) o *drogota* (de *drogadicto*). Observábamos el mismo fenómeno en la descripción de *-ete* (*fumeta*), aunque no halláramos ejemplos en el corpus. Una vez más, incluiremos estos términos coloquiales en nuestro listado, indicándolos con un asterisco (*).

Ejemplos: serrote, brutota, casota, muchachote

Ejemplos de términos lexicalizados: camarote, palabrota, cascote, islote

[word=".*ot(a)e(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
chavalote	11	3	drogota*	1	0,2
drogotas*	1	0,3	grandota/es	3	0,6
generalote	1	0,3			
grandote	2	0,5			
pasota*	2	0,5			
Total	17	4,6		4	0,8

Tabla 146 - Sufijo *-ote* en ES y TRAD

Lexicalizadas ES: cascotes, Palotes (Perico de los), machote, chicote

Lexicalizadas EN: cabezota, cascotes, cipote, islotes, palabrota

La diferencia entre subcorpus es significativa a favor de ES (LL +12,81 p < 0,001) aunque los ejemplos hallados no son muchos. Juan Madrid utiliza *chavalote* como vocativo y también para hablar en tercera persona del personaje al que se dirige ese vocativo. En estos casos su función parece ser la de buscar complicidad o transmitir cierto grado de simpatía por parte del travesti hacia su interlocutor:

<Madrid>: Usted , **chavalote** , puede dar color a nuestro espectáculo . ¿Lo has visto alguna vez , <**chavalote**> ? - No . - Recuérdalo , puedes entrar gratis . Entrada y copa .

Machote también se utiliza como apelativo afectuoso pero, como hemos mencionado en la presentación de este sufijo, la RAE lo considera lexicalizado. *Grandote* es aumentativo en Núñez (se refiera a un cenicero), pero aporta un toque humorístico algo peyorativo a la descripción de Martín:

<Martín>: no los robaran los gitanos o el Papa Ogro , que menudo pinta era . Salía cada mañana con su coche , <grandote> que era , con sus bigotazos , la barriga enorme , los ojos chispeantes , la sonrisa cínica , depravada.

En cambio, *drogotas* es claramente despectivo (se contraponen los «drogotas, marginales y demás» a los «ciudadanos de a pie»); *generalote* también lo es, pues aparece en medio de un discurso en el que un personaje se queja del daño que hizo a España la dictadura de Franco, y *pasota* se utiliza para describir «la lengua cheli de los años setenta»:

<Vázquez>: Dictadura del proletariado en <pasota> : Los rojeras gustan pasar por el aro a los tragones hasta arrascar el raje con el fregao de los colores . La curranda ha de antoligar el cotarro . O bien una de las tesis de abril de Lenin : Hay que esparrarar el bando gambeante endiñando el cotarro a los rojeras , también llamados rogelios . El pasota-leninismo capaz de traducir ¿Qué hacer ? de Lenin , por ¿Cómo montárselo ?

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Madrid	11	26,5	Leonard	3	3,8
Vázquez	3	3,4	Ellroy	1	0,9
Martín	2	3,5	Cornwell	0	0
Núñez	1	2,3	Grafton	0	0
Giménez	0	0	Highsmith	0	0
Martínez	0	0			

Tabla 147 - Distribución del sufijo *-ote* por texto

En cuanto a los términos hallados en TRAD, contamos *monkey* en la categoría de los términos inherentemente apreciativos, pues el elemento que da pie a la utilización de *drogota* («El drogota, un tipo duro»), *tough monkey*, puede utilizarse para referirse a alguien con una fuerza extraordinaria, o ser un término despectivo para personas negras (la persona en cuestión lo es), como se relata en un artículo del *Los Angeles Times* referido a un polémico incidente en el que un comentarista deportivo llamó de esa manera al jugador de baloncesto Allen Iverson (Stewart 1996). En cuanto a *grandota*,

parece cumplir una función atenuadora, por ejemplo, cuando se refiere a una mujer atractiva pero entrada en carnes. Las tres veces que aparece esta última palabra encontramos el adjetivo *big* en los originales:

<Leonard>: En esos trece años , la chica fina aunque <grandota> se había vuelto más grande aún , sus pechos parecían de circo pero no dejaban de ser buenos pechos

-->par_en: The fine big girl had in thirteen years become bigger , show tits grown to circus tits but still okay

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	1
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	3
Información contextual	
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
Ø	
Total	4

Tabla 148 - Elementos activadores del sufijo *-ote* en traducciones

6.3.2.4. -UDO / -UDA

Este sufijo, que forma derivados únicamente adjetivales y se utiliza especialmente en sentido peyorativo (Gooch 1970: 12), no es considerado por la Real Academia como aumentativo ni como despectivo, puesto que sus derivados «no se interpretan agregando la connotación de menosprecio al significado de su base» (2009: 661). La RAE describe sus derivados como palabras que caracterizan a personas o animales que destacan por el tamaño, el exceso, la desproporción o la malformación de alguna parte de su cuerpo, como *hocicudo*. También denotan la abundancia de pelo (*greñudo*) o de alguna característica que se posee en exceso, como *confianzudo* (2009: 530-531).

Lang (1990: 116) afirma que, a pesar de mostrar una mayor tendencia a formar palabras peyorativas, la abundancia de determinadas características puede ser positiva si se consideran atributos físicos atractivos, como en el caso de *tetuda*. Kaplanová (2010: 41), por su parte, comenta que a veces simplemente significa que algo tiene forma de otra cosa, como sucede en *ganchudo* o *picudo*.

La reticencia de la Academia a incluirlo en el listado de apreciativos se debe probablemente a que en la mayor parte de los casos (si no todos), cambia la categoría

gramatical de la palabra, de sustantivo a adjetivo (*vello* > *velludo*). Sin embargo, hemos decidido incluirlo en este inventario por haber sido descrito como aumentativo por seis autores.

[word=".*ud(o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
barbudo	1	0,3	huesudo/a(s)	6	1,2
huesudo/a	2	0,5	narigudo	1	0,2
hievuda	1	0,3	peludos	2	0,4
melenuda	1	0,3	tetuda	1	0,2
panzudos	1	0,3	velludo/a(s)	3	0,6
peludo/a(s)	3	0,8			
sañudo	1	0,3			
talentudo	1	0,3			
tetuda	1	0,3			
velludo/a(s)	3	0,8			
Total	15	4,1		13	2,7

Tabla 149 - Sufijo *-udo* en ES y TRAD

Lexicalizadas ES: campanudas, ceñudo/a, cojonudo/a, concienzudo, cornudo(s)

Lexicalizadas EN: (cuero) cabelludo, ceñudo/a, concienzudo, cornudo/a(s), puntiaguda

Nuestro criterio de inclusión de palabras es, como hemos hecho en casi todos los casos, que el significado se pueda extraer de aplicar a la base el sentido del sufijo (en este caso, «posesión en abundancia»). De esta manera, *barbudo*, aunque figure en el DRAE, entra dentro de nuestra clasificación, pero no lo hace *campanudas*, pues no significa tener muchas campanas, sino que tiene el significado lexicalizado de «personas que se expresan con un estilo hinchado o retumbante». Siguiendo este criterio, no encontramos diferencias estadísticamente significativas entre textos originales y traducidos en cuanto al uso de este sufijo aumentativo.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	5	5,9	Grafton	5	5,4
Madrid	4	9,6	Leonard	4	5,1
Martín	3	5,3	Ellroy	2	1,8
Núñez	2	4,7	Cornwell	1	0,9
Vázquez	1	1,1	Highsmith	1	1
Martínez	0	0			

Tabla 150 - Distribución del sufijo *-udo* por texto

En ES vemos que estos adjetivos suelen utilizarse en contextos en donde lo que se busca es un matiz despectivo o de exageración, o incluso caricaturizar a un personaje:

<Núñez>: Alto , <huesudo> , de aspecto y cara macilenta , daba cortas y nerviosas chupadas a un cigarrillo sin filtro .

<Núñez>: Nalgas de mamut que bambolea con orgullo y sus brazos son cortos , rollizos y extraordinariamente <velludos> .

Los adjetivos formados con el sufijo *-udo* son en su mayoría equivalentes de adjetivos ingleses formados también con sufijos, especialmente con *-y*, como *knobby*, *bony* y *skinny* (los tres con connotaciones de demasiada delgadez) o *shaggy*, *furry*, *downy* y *fuzzy*, que significan, simplemente, «lleno de» pelo o vello. En una de las concordancias de *velludo* (*flocked*) encontramos también un caso del sufijo *-ed*, que sirve para formar adjetivos de bases nominales o verbales y significa «possessing or having the characteristics of»²⁰⁴. El último sufijo es *-eer*, que suele formar sustantivos «significantly associated with the referent of the base word (auctioneer; engineer; mountaineer; pamphleteer); now frequently pejorative (profiteer; racketeer)»²⁰⁵. Con este sufijo se ha formado, a partir de *brisket* (un término coloquial para referirse a los pechos), el sustantivo *brisketeer* para referirse a una mujer en un contexto informal de tono humorístico.

El único término que hemos incluido en otra categoría es *hooter*, un término coloquial para referirse a la nariz que se ha traducido por un aumentativo pues se refiere a la nariz de Lincoln que, como todos sabemos, era bastante grande.

<Grafton>: y pinté a Lincoln un bigote que le redujo la nariz . Hasta entonces no me había dado cuenta de lo <narigudo> que era .

-->par_en: I 'd never paid attention to what a **hooter** he had .

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	
Sufijo apreciativo	12
Adjetivo o adverbio con connotación apreciativa + sustantivo	
Información contextual	1
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
∅	
Total	13

Tabla 151 - Elementos activadores del sufijo *-udo* en traducciones

²⁰⁴ ed. *Collins English Dictionary - Complete & Unabridged 10th Edition* [fecha de consulta: 11 de agosto de 2015].

²⁰⁵ eer. *Dictionary.com Unabridged* [fecha de consulta: 11 de agosto de 2015].

6.3.3. *Peyorativos*

Por la variedad de clasificaciones y la cantidad de formas existentes, los sufijos peyorativos no están descritos con tanto detalle como los miembros de los dos grupos anteriores en los inventarios de los autores consultados. En algunas ocasiones, solo podremos proporcionar una breve descripción o algún ejemplo de uso. Las apariciones en el corpus son escasas en este grupo, por lo que tendremos en cuenta todas las palabras sufijadas aunque aparezcan en el diccionario de la Real Academia, a excepción, por supuesto, de que se haya perdido por completo la transparencia de significado y el sufijo no aporte ya connotaciones despectivas, como es el caso de *aguilucho*.

6.3.3.1. -ACHO / -ACHA

Gooch (1970: 13) y Lang (1990: 118) consideran que se trata de un sufijo peyorativo con matices aumentativos. Lang afirma que suele formarse sobre bases adjetivales y los derivados atribuyen cualidades a seres animados. Cantera Ortiz de Urbina (2012: 104) le atribuye en ocasiones «un matiz más bien simpático».

Ejemplos: ricacho, vivaracho

[word=".*ach(o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
populacho	1	0,3	ricacho	1	0,2
ricacho	1	0,3			
Total	2	0,5		1	0,2

Tabla 152 - Sufijo *-acho* en ES y TRAD

Como vemos, apenas hay algún caso de este sufijo en el corpus, por lo que las diferencias no pueden ser significativas. Además de las palabras apuntadas en la tabla, en TRAD hallamos un ejemplo de *gabacho*, que es un término fuertemente despectivo para referirse a los franceses, aunque en el contexto de la novela equivale a *gringo*. No obstante, no lo hemos incluido pues, según el DRAE, proviene de la palabra provenzal *gavach* («que habla mal»), con lo que la terminación no vendría dada por la adición del

sufijo apreciativo *-acho*, sino por la desinencia de género. En el original estadounidense se utiliza este mismo término, dado que es proferido por un hispanohablante.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	1	1,8	Ellroy	1	0,9
Vázquez	1	1,1	Cornwell	0	0
Giménez	0	0	Grafton	0	0
Madrid	0	0	Highsmith	0	0
Martínez	0	0	Leonard	0	0
Núñez	0	0			

Tabla 153 - Distribución del sufijo *-acho* por texto

Tanto *populacho* como *ricacho* son términos claramente despectivos, como puede observarse a partir del contexto de las concordancias:

<Martín>: El contador de historias podrá ser un bufón que entretiene al <populacho>, pero jamás será un auténtico literato

<Vázquez>: El policía había aprendido a sostenerle la mirada , a fingirse tan entero como aquel <ricacho> de mierda y alcanzó un tono de voz tranquilo cuando preguntó

<Ellroy>: Hurwitz era una especie de jugador arruinado y ese jodido <ricacho> de Exley le compró todas las pieles y le dijo a Johnny cómo simular el robo .

-->par_en: Hurwitz was some kind of gone-bust gambler , and fuckin ' rich-kid Exley , he bought all the furs and told Johnny how to stage the heist . "

Vemos que el término original en la concordancia de TRAD de *ricacho* es *rich-kid*, que ya de por sí conlleva cierta connotación despectiva, dado que suele referirse a alguien joven que no necesita trabajar —un *señorito*²⁰⁶—, pero, además, dicha connotación está reforzada por *fuckin*'.

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	1
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	
Información contextual	
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
Ø	
Total	1

Tabla 154 - Elementos activadores del sufijo *-acho* en traducciones

²⁰⁶ señorito. *WordReference English-Spanish Dictionary* [fecha de consulta: 11 de agosto de 2015]

6.3.3.2. -ACO / -ACA

Como el sufijo anterior, aporta matices aumentativos. Su productividad es limitada y tiene tendencia a unirse con el sufijo peyorativo *-arro* (Lang 1990: 118).

Ejemplos: libraco, tiparraca, bicharraco

[word=".*ac(o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
libraco(s)	3	0,8			
moraco	1	0,3			
pajarraco	1	0,3			
sudacas	3	0,8			
Total	8	2,2		0	0

Tabla 155 - Sufijo *-aco* en ES y TRAD

No hallamos ningún caso de este sufijo en TRAD, lo que hace que los ocho casos de ES cobren mayor importancia. Las diferencias son estadísticamente significativas (LL +13,54 $p < 0,001$)²⁰⁷.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	3	5,3	Cornwell	0	0
Núñez	3	7,1	Ellroy	0	0
Giménez	1	1,2	Grafton	0	0
Vázquez	1	1,1	Highsmith	0	0
Madrid	0	0	Leonard	0	0
Martínez	0	0			

Tabla 156 - Distribución del sufijo *-aco* por texto

Las tres apariciones de *libraco* corresponden a Núñez, y las utiliza como aumentativos, para designar libros muy extensos de disciplinas complejas, pero también con el matiz despectivo de libro pesado que, a pesar de decir mucho, no sirve para nada.

<Núñez>: Comentaron que la semana había sido muy fuerte . Que estaban intoxicadas de <libracos> . Y no exageraban . Esta ha sido la primera vez que las he visto estudiar en serio .

<Núñez>: No pude soportar por más tiempo aquella conversación . Y es que las frases lapidarias sacadas de un <libraco> de sicología me revientan .

²⁰⁷ Test exacto de Fisher: 0,001148 ($p < 0,01$)

Aparece también *moraco* y *sudacas*, ambos términos fuertemente despectivos derivados de *moro* y *sudamericanos*, respectivamente, y *pajarraco*, que denota antipatía o, en este caso, hartazgo:

<Giménez>: no hay descanso , al final acabas cabreándote con el padre prior o te molestan los graznidos de los <pajarracos> que hay en el cocotero bajo el que has decidido dormir

6.3.3.3. -AJO / -AJA

Lang (1990: 119) le atribuye, una vez más, sentido peyorativo aumentativo²⁰⁸ y lo describe como el más productivo del grupo, aunque no puede aplicarse sobre bases animadas. En cuanto a las inanimadas, menciona *hierbajo*, *espantajo* y *latinajo*, tres palabras que figuran en el diccionario, pero ya hemos dicho que ese no será nuestro criterio a la hora de incluir derivados peyorativos en la tabla, sino la transparencia de su significado. No sucede lo mismo con las bases adjetivales (*pequeñajo*, *chiquitajo*, *rubiajo*): ninguna aparece en el DRAE. No incluiremos los derivados nominales de verbos (*comistrajo*, *bebistrajo* o *escupitajo*) que, como ya sabemos, no son considerados apreciativos por la RAE pero, además, en este caso, su significado suele estar muy lexicalizado.

[(word=".*aj(o|a)(s?)") & (word!="abajo|bajo|baja|debajo|caja(s?)" %c)];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
latinajos	1	0,3	hierbajos	2	0,4
migajas	3	0,8	migajas	3	0,6
			hatajo	2	0,4
Total	4	1,1		7	1,4

Tabla 157 - Sufijo -ajo en ES y TRAD

Lexicalizadas ES: colgajo, comistrajo, escupitajo, gargajos, zarandajas

Lexicalizadas EN: colgajo, escupitajo, zarandajas

Esta vez, las apariciones de este sufijo son más numerosas en TRAD que en ES, pero la diferencia no llega a ser significativa.

²⁰⁸ En esto difiere Cantera Ortiz de Urbina (2012: 106), que le atribuye valores diminutivos y despectivos.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Giménez	2	2,3	Highsmith	5	5,1
Martínez	1	1,8	Ellroy	2	1,8
Vázquez	1	1,1	Cornwell	0	0
Madrid	0	0	Grafton	0	0
Martín	0	0	Leonard	0	0
Núñez	0	0			

Tabla 158 - Distribución del sufijo *-ajo* por texto

Latinajo, *hatajo* y *hierbajo* aparecen con una connotación despectiva mientras que *migajas*, en general, parece aportar un matiz simplemente diminutivo incluso cuando se utiliza en sentido metafórico:

<Vázquez>: Álvaro ya les había metido en la habitación la sacristana , por los muchos <latinajos> que Carvalho le había detectado

<Ellroy>: ¡El dinero para el vino no se repartirá hasta las seis en punto ! <Hatajo> de borrachos atontados ! ¡Esto es un plató para filmación de exteriores , no la misión de Cristo

-->par_en: Wine money not disbursed until six o'clock ! The tsuris you **stumblebums** inflict !

<Highsmith>: Estaba limpiando de <hierbajos> uno de los lechos de rosas que había cerca de la casa .

-->par_en: He was then **weeding** in one of the rose beds near the house .

<Highsmith>: Tom sacudió los granos de arena del suelo de la camioneta , <migajas> grises de tierra .

-->par_en: With a rough cloth he found in a corner of the garage , Tom flicked out some grains of sand from the back of the station wagon , gray **crumbs** of soil .

<Giménez>: Quizá sería más prudente ir ofreciéndoles algunas <migajas> de información , a lo mejor estaríamos más tranquilos

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	7
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	
Información contextual	
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
∅	
Total	7

Tabla 159 - Elementos activadores del sufijo *-ajo* en traducciones

Los elementos que dan pie a la utilización de estos derivados son todas palabras de significado inherentemente diminutivo o peyorativo: *crumb*, *weed* y, para *hatajo*, los insultos *shmuckface* («hatajo de estúpidos») y *tumblebum* («hatajo de borrachos

atontados»). La adición de *hatajo* contribuye a intensificar la idea de un grupo de personas indeseables.

6.3.3.4. -ALES

Este sufijo se usa, según la Real Academia (2009: 662), casi exclusivamente en España. Lang (1990: 121) lo considera peyorativo con tintes humorísticos y cita la siguiente descripción de Manuel Seco: «de carácter marcadamente popular, expresa siempre una actitud humorística, de burla o de atenuación». Comenta que hay quien lo considera una variante del sufijo no apreciativo *-al* y que no es muy productivo. Produce derivados nominales y adjetivales de bases adjetivales, aunque las bases nominales que suelen mencionarse, como fresco, rubio, viejo o vivo, son también adjetivos. Recordemos que Gómez Torrego (1997: 23) no lo clasificaba dentro del grupo de los despectivos, sino de los sufijos familiares, que presentan significados de ironía o desprecio pero con marcado carácter familiar.

Ejemplos: frescales, rubiales, viejales, vivales.

`[(word=".+al|.+ales") & (word!="Conesal|mal|igual|tal|cual")];209`

El único ejemplo que hemos encontrado del uso de este sufijo es el siguiente, en el subcorpus de originales españoles:

<Giménez>: Fue usted quien desde el principio creyó en la personalidad de un loco para realizar estos delitos . Que sea muy inteligente no quiere decir que esté menos **<mochales>** .

Sin embargo, no podemos incluirlo porque es una palabra completamente lexicalizada: aparece en el diccionario y su significado (chiflado) no puede deducirse del de la base (cabeza humana).

²⁰⁹ Esta expresión regular incluye las concordancias de *-al*, puesto que en un principio considerábamos la posibilidad de incluir este sufijo junto a los aumentativos.

6.3.3.5. -ANGO / -ANGA, -INGO / -INGA

Según Selva (1949: 195), nace probablemente como variante de *-anco/a*. Lang lo clasifica, como a *-ales*, dentro de los peyorativo-humorísticos, agrupándolo junto a *-engo/a*, *-ingo/a*, *-ongo/a* y *-ungo/a*, sufijos que, a excepción de *-ingo/a*, no han entrado en nuestra clasificación (aunque *-engue* se quedó a las puertas). No obstante, en esta búsqueda incluiremos todas las formas, dado que no esperamos obtener abundantes ejemplos. Estas variantes «are not allomorphs but haphazardly distributed according to individual whim» (1990: 121). Según Lang, son característicos del español de Sudamérica y reflejan influencias indias y africanas. La RAE comenta que en Bolivia se usa *-ango/a* como aumentativo, como en *casanga* o *puertanga* y, con respecto a *-ingo/a*, afirma lo siguiente: «Posee gran vitalidad en algunas regiones andinas —muy destacadamente en la parte oriental de Bolivia— tanto con adjetivos (*quietingo*, *friingo*) como con sustantivos (*casinga*, *mesinga*) y adverbios (*ahoringa*, *cerquina*)» (RAE 2009: 631). Sin embargo, añade que actualmente estos sufijos se han popularizado en el habla popular neológica de España, por lo que quizás hallemos ejemplos de su uso en nuestro corpus.

Ejemplos: fritanga, señoringa, frailengo, bailongo, chulengo, facilongo, blandengo

[word=".*ang(o|a)(s?)"];

[(word=".*eng(ue|o|a)(s?)") & (word!="(t|v)eng(a|o)" %c)];

[word=".*ing(o|a|ue)(s?)"];

[(word=".*ong(o|a|ue)(s?)") & (word!="(su)?pong(o|a)" %c)];

[word=".*ung(o|a|ue)(s?)"];²¹⁰

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
mandanga(s)	2	0,5	ringorrango	1	0,2
			blandengue	3	0,6
Total	2	0,5		4	0,8

Tabla 160 - Sufijos *-ango*, *-engo*, *-ingo*, *-ongo* y *-ungo* en ES y TRAD

²¹⁰ Una vez más, estas expresiones probablemente podrían haberse juntado en una única expresión regular, de haber sabido que posteriormente agruparíamos los cinco sufijos.

Lo cierto es que, al ser este un corpus de novelas escritas y traducidas a la variante de español europeo, no hemos encontrado prácticamente ningún caso de este sufijo. Como es de esperar, las diferencias entre subcorpus no son significativas. La palabra *mandanga* (indolencia o tonterías, en plural) parece llevarlo, pero su origen no es del todo claro²¹¹, por lo que su inclusión en la lista es dudosa. Ocurre lo mismo con *ringorrango* (adorno superfluo o rasgo de pluma exagerado)²¹², en TRAD.

<Vázquez>: Pero luego , con tantas libertades y tantas <mandangas> pues hay que tirar del botón para tapar y abrir bocas

<Grafton>: una complicada colección de nombres y números impresos con variada intensidad , dibujos con mucho <ringorrango> geométrico y sellos oficiales .

-->par_en: Paper currency has a baroque assortment of shaded names and numbers , lacy scrollwork , and official seals .

Scrollwork, en principio, no nos proporciona ese matiz negativo de exageración, pero quizás el adjetivo *baroque*, que puede tenerlo, haya propiciado el uso del término elegido en la traducción.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	1	1,8	Ellroy	3	2,7
Vázquez	1	1,1	Grafton	1	1,1
Giménez	0	0	Cornwell	0	0
Madrid	0	0	Highsmith	0	0
Martínez	0	0	Leonard	0	0
Núñez	0	0			

Tabla 161 - Distribución de los sufijos -ango, -engo, -ingo, -ongo y -ungo por texto

No encontramos ninguna palabra terminada en *-inga* o *-ingo*, *-ongo*, *-onga*, *-ungo*, *-unga* ni en *-enga* o *-engo*, pero sí en *-engue*, un sufijo que no habíamos incluido en la lista por aparecer solo en cinco repertorios de los consultados: *blandengue* aparece en tres ocasiones en TRAD, siempre como traducción de calificativos de intencionalidad despectiva.

²¹¹ Según la obra *Origen de las palabras estrafalarias*, de José Calles Vales (Editorial LIBSA, 2002) citada en Álvarez (2004).

²¹² Según Selva (1949: 196): «proviene, a mi ver, de *ringla* desde que expresa los trazos caprichosos que salen de la *regla* o *renglón*. Hay que contar la superposición de dos incrementos *-orra* y *-ango*, que hacen más despectiva la voz».

<Ellroy>: está colado , coladísimo , por el astro de la producción , Rock Rockwell , ese seductor

<blandengue>

-->par_en: is hot , hot , hotsville with the star of the movie : **limpwristed** lothario Rock Rockwell .

<Ellroy>: La mirada de rayos X ; no estaba mal para un <blandengue> .

-->par_en: This X-ray stare-not bad for a **punk** .

<Ellroy>: . - <Blandengue> , chivato , hijo de puta !

-->par_en: Leroy : " **Sissy** fink stool pigeon motherfucker . "

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	3
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	
Información contextual	
Tono general del fragmento (razones estilísticas o compensación)	1
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
∅	
Total	4

Tabla 162 - Elementos activadores de los sufijos *-anga*, *-engo*, *-ingo*, *-ongo* y *-ungo* en traducciones

6.3.3.6. -ASTRO / -ASTRA

Sufijo peyorativo-aumentativo en opinión de Lang (1990: 119), que suele denotar ineficacia en determinada profesión (*poetastro*, *criticastro*, *medicastro*). También se aplica a relaciones de parentesco, pero en estos casos, los derivados suelen estar fuertemente lexicalizados, como en *madrastro*, *hijastro* o *padastro*, que no tienen por qué transmitir connotaciones negativas. Incluiremos también los derivados en *-astre*.

[word=".*astr(e|o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
pillastre	1	0,3	politicastro	1	0,2
Total	1	0,3		1	0,2

Tabla 163 - Sufijo *-astro* en ES y TRAD

Lexicalizadas ES: hermanastras, madrastra

Lexicalizadas EN: padastro

No parece ser un sufijo muy utilizado. Los casos no son suficientes para poder hacer comparaciones entre subcorpus.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Vázquez	1	1,1	Ellroy	1	0,9
Giménez	0	0	Cornwell	0	0
Madrid	0	0	Grafton	0	0
Martín	0	0	Highsmith	0	0
Martínez	0	0	Leonard	0	0
Núñez	0	0			

Tabla 164 - Distribución del sufijo *-astro* por texto

Los siguientes son los únicos dos casos que hemos localizado, el segundo con connotaciones más negativas que el primero, que parece aportar simplemente una nota jocosa a la explicación:

<Vázquez>: El Jaguar no es mío . Usted y yo somos invitados de un <pillastre> riquísimo que se llama Lázaro Conesal . Un rico como hay pocos , de los que enseñan los aviones

<Ellroy>: « Welles Noonan es un <politicaastro> poco escrupuloso , con falsos antecedentes liberales .
-->par_en: " Welles Noonan is an unscrupulous **hack politician**²¹³ with spurious liberal credentials .

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	1
Información contextual	
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
∅	
Total	1

Tabla 165 - Elementos activadores del sufijo *-astro* en traducciones

6.3.3.7. -ATO / -ATA

Este sufijo —o su variante epicena *-ata*— en principio no supera la criba inicial, pues solo lo mencionan cinco autores; tres como despectivo (Gómez Torrego 1997: 22,

²¹³ «a professional who renounces or surrenders individual independence, integrity, belief, etc., in return for money or other reward in the performance of a task normally thought of as involving a strong personal commitment: a political hack». De: *hack*. *Dictionary.com Unabridged* [fecha de consulta: 12 de agosto de 2015].

RAE 2009: 662 y Soler Espiauba 1996), y dos como diminutivo (Alonso 1982: 189 y Seco 1989: 213). Sin embargo, hemos creído adecuado estudiarlo también en nuestro corpus, como sufijo despectivo, dada su afinidad formal y funcional con otros sufijos o variantes de las que hemos estado hablando en este trabajo, como *-etas*, *-azas*, *-ica*, *-otas* y, sobre todo, *-ota* y *-eta*, que dan lugar a términos propios del argot (*drogota*, *porreta*), comunes en los ámbitos en los que transcurre la acción de las novelas de nuestro corpus.

El sufijo despectivo *-ato/a* forma derivados despectivos como *cegato*, *niñato* o *novata* (RAE 2009: 662). Según la Real Academia, la variante *-ata*, propia del lenguaje coloquial —especialmente, del ámbito juvenil y jergal europeo—, se considera un sufijo con entidad propia, mediante el cual se crean derivados a partir de bases nominales, adjetivales y verbales (*bocata*, *sociata*, *fregata*). Marcaremos los derivados de esta variante con un asterisco (*).

`[(word=".*at(o)a(s?)") & (word!="asesinato(s?)|dato(s?)|plato(s?)|trat(o)a(s?)|rato|psicópata|Chato" %c)];`
`[(word=".*at(o)a(s?)") & (word!="asesinato(s?)|dato(s?)|plato(s?)|trat(o)a(s?)|rato|psicópata" %c)];214`

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
cubata(s)*	2	0,5	negrata(s)*	2	0,4
drogata(s)*	2	0,5			
niñato(a)	3	0,8			
sociata(s)*	2	0,5			
Total	9	2,4		2	0,4

Tabla 166 - Sufijo *-ato* en ES y TRAD

Lexicalizadas ES: chivato, culata, mulato, novato, perorata

Lexicalizadas EN: chivato, culata, mulato, novato

Algunas palabras del corpus que pueden parecer derivados de este sufijo, tienen su origen en otras lenguas, como *carromato* y *bravata* (del italiano), *columnata* y *literato* (del latín) o *garabato*, de origen prerromano.

Encontramos varios ejemplos, especialmente en el subcorpus de originales españoles y de la variante jergal del sufijo (*cubata* de *cubalibre*, *drogata* de *drogadicto*,

²¹⁴ En este caso las expresiones regulares difieren entre subcorpus porque las concordancias eran numerosas y decidimos excluir de la búsqueda las palabras que más se repetían en cada uno.

sociata de *socialista* y *negrata* de *negro*). La diferencia de uso a favor de dicho subcorpus es estadísticamente significativa (LL +7,04 p<0,01)²¹⁵.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Vázquez	3	3,4	Cornwell	1	0,9
Martínez	2	3,7	Leonard	1	1,3
Núñez	2	4,7	Ellroy	0	0
Giménez	1	1,2	Grafton	0	0
Martín	1	1,8	Highsmith	0	0
Madrid	0	0			

Tabla 167 - Distribución del sufijo *-ato* por texto

Ambas concordancias de *negrata* provienen de términos despectivos en el original:

<Leonard>: No , señor . Soy afroamericano . Antes era un negro , betún , un <negrata> , pero ahora soy afroamericano .

-->par_en: I used to be Neegro , I was cullud , I was black , but now I'm African-American .

<Cornwell>: En cuanto a los blancos , iban desde un guardagujas retirado que aún llamaba « <negratas> » a los negros hasta una rica ama de casa que sólo veía a los negros cuando salía uno de ellos por televisión

-->par_en: The professions of the white jurors ranged from a retired railroad switchman who still called blacks niggers ' to a rich housewife whose only exposure to blacks was when she watched the news

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	2
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	
Información contextual	
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	
∅	
Total	2

Tabla 168 - Elementos activadores del sufijo *-ato* en traducciones

6.3.3.8. -ORRIO / -ORRIA, -ORRO / -ORRA

Estos sufijos, especialmente *-orro/a*, han recibido algo más de atención por parte de los autores consultados. Recordemos que había disparidad de opiniones en cuanto a

²¹⁵ Test exacto de Fisher: 0,012349 (p < 0,05)

su clasificación en uno u otro grupo de apreciativos. Lang (1990: 120) lo considera peyorativo-aumentativo y afirma que los derivados surgen más comúnmente de adjetivos (*feorro*, *viejorro*, *calentorro*) que de sustantivos (*vidorra*, *tiorra*, *pitorro*). Comenta este autor que *-arro* y *-urro* se consideran a veces variantes de *-orro*, que no suelen darse solas sino en combinación con *-ón* (*mansurrón*, *vozarrón*, *mozarrón*). Las incluiremos en nuestra búsqueda, teniendo en cuenta que *-arro/a* no entró en nuestro listado de sufijos por muy poco (aparecía en cinco repertorios).

Santos (2006: 59), que lo incluye dentro del grupo de los aumentativos, señala, sin embargo, que *-orro* podría considerarse un sufijo más calificador que cuantificador y apunta asimismo que puede conferir matices tanto positivos (*vidorra* es una buena vida llena de lujos; si decimos que una persona está *buenorra* es que es muy atractiva) como despectivos (*ventorro*). *-Orrio/a*, sin embargo, parece dar lugar solo a términos despectivos (*villorrio*, *bodorrio*²¹⁶, *pastelorrio*). Incluiremos también en la búsqueda la variante *-orio*.

[word=".*orr(o|a)(s?)"];

[word=".*or(r?)i(o|a)(s?)"];²¹⁷

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
calentorros	1	0,3			
machorra	1	0,3			
tascorro	1	0,3			
papelorios	1	0,3			
vejestorio	1	0,3			
Total	5	1,4		0	0

Tabla 169 - Sufijos *-orro*, *-orrio* y *-orio* en ES y TRAD

Lexicalizadas ES: abejorro, ceporro, pedorra

Lexicalizadas EN: abejorro

Todos los términos recogidos son peyorativos. A pesar de la observación de Lang, vemos más presencia de sustantivos que de adjetivos en la tabla, aunque son muy pocas concordancias para extraer conclusiones al respecto. No hemos hallado ejemplos de palabras derivadas con *-orrio*. En cuanto a otras variantes, hallamos un caso en ES de

²¹⁶ Comenta de Miguel (2010) lo siguiente, a propósito del sentido que se atribuyen hoy en día en el habla coloquial a *bodorrio*: «Así, el "bodorrio" no es "un desastre de boda, mal organizada y populachera", como indica el sufijo <orrio>. Es el caso, por ejemplo, de "villorio". Realmente, "bodorrio" quiere decir ahora una boda de tronío, lujosa».

²¹⁷ Estas dos expresiones podrían haberse juntado en una: [word=".*or(r?)(i)?(o|a)(s?)"].

bandarra, pero no podemos considerarla una palabra derivada, pues, al aparecer su etimología es otra²¹⁸; cinco casos en ES y dos casos en TRAD de *macarra* (pero, según el DRAE, esta palabra tiene sus orígenes en el catalán y el francés); uno en TRAD de *chaparro* (del vasco) y también un caso en ES y cinco en TRAD de *cacharro*, un vocablo muy lexicalizado cuyo significado nada tiene que ver ya con *cacho*. Quizás podríamos incluir *cazurro*, pues existe la palabra *cazo*, que significa *torpe*, y *cotarro*, de *coto*, referido a determinada población o terreno (ambos en ES), pero estos significados son de dudosa transparencia. Sin incluirlos, la diferencia a favor de ES es significativa (LL +8,46 p<0,01)²¹⁹.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Martín	3	5,3	Grafton	0	
Giménez	1	1,2	Leonard	0	
Vázquez	1	1,1	Ellroy	0	
Madrid	0	0	Cornwell	0	
Martínez	0	0	Highsmith	0	
Núñez	0	0			

Tabla 170 - Distribución de los sufijos *-orrio*, *-orro* y *-orio* por texto

6.3.3.9. -UCHO / -UCHA

Casi todos los autores atribuyen a este sufijo peyorativo matices diminutivos. Se aplica a adjetivos, a los que dota de una nota despectiva (como *larguirucho*, *feúcho* o *blancucho*), por lo que no es aplicable a bases de contenido semántico positivo (Lang 1990: 120). También se aplica a sustantivos, como en *teatrucho* o *animalucho* (RAE 2009: 661). Cantera Ortiz de Urbina (2012: 115) le atribuye a veces un matiz afectivo o más bien compasivo (*debilucho*, *malucho*).

Ejemplos: *periodicucho*, *flacucho*, *casucha*, *cuartucho*

Ejemplos de términos lexicalizados: *aguilucho*, *serrucho*

`[(word=".*uch(o|a)(s?)") & (word!="much(o|a)(s?)" %c)];`

²¹⁸ Según De Deira (2008): «Si atendemos al Diccionario de la Real Academia, *bandarra* es sinónimo de *sinvergüenza*, aunque puede tener otras acepciones como *vago* o *gandul*. Pero además de su semántica, *Bandarra* es el nombre de una persona que, por cierto, influyó de algún modo en unos acontecimientos ocurridos en el siglo XVI», un poeta nacido en el año 1500 en la ciudad portuguesa de Trancoso.

²¹⁹ Test exacto de Fisher: 0,014534 (p < 0,05)

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
casucha	1	0,3	casuchas	1	0,2
hotelucho	1	0,3	delgaduchas	1	0,2
larguirucho	1	0,3	flacuchas	1	0,2
			larguirucho/as	2	0,4
			revistucha	1	0,2
Total	3	0,8		6	1,2

Tabla 171 - Sufijo *-ucho* en ES y TRAD

En ES hallamos dos sustantivos y un adjetivo. Podríamos haber contado *paparrucha* («tontería, estupidez», de la base *páparo*²²⁰), pero su significado no nos ha parecido lo suficientemente transparente. En TRAD este sufijo se utiliza más, siempre de manera peyorativa, pero, dado que este subcorpus cuenta con más cantidad de palabras, la diferencia no llega a ser significativa.

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Madrid	1	2,4	Ellroy	3	2,7
Martín	1	1,8	Leonard	2	2,6
Vázquez	1	1,1	Cornwell	1	0,9
Giménez	0	0	Grafton	0	0
Martínez	0	0	Highsmith	0	0
Núñez	0	0			

Tabla 172 - Distribución del sufijo *-ucho* por texto

La mayor parte de los elementos activadores de este sufijo son palabras cuyo significado lleva ya incorporado un matiz despectivo: los sustantivos *dump* para *casucha* y *rag* para *revistucha*, y los adjetivos *spindly* (alto y delgado pero generalmente también frágil) y *scrawny* (delgado pero en exceso) para *larguirucho*. En el caso de *skinny*, consideramos que uno de los ejemplos en los que se traduce mediante el sufijo puede venir dado por el concepto no demasiado positivo que se tiene, en general, de la delgadez de las modelos, o del mundo de la moda al que se hace referencia, puesto que el tono del fragmento no puede considerarse especialmente negativo:

<Leonard>: Tipos con las gafas de sol levantadas hacían posar a chicas <flacuchas> en la playa y las fotografiaban .

-->par_en: It was no more Louis's scene than the hotel where the Colombians hung out , but the show was better : all the tank tops and hundred-dollar pump-up basketball shoes . Louis had lived here ten years ago when old retired people from New York sat on the hotel porches wearing hats , their noses painted white , and boat-lift Cubans worked their hustles down the street . Five years ago when it was beginning to change

²²⁰ Según el DRAE: «Aldeano u hombre del campo, simple e ignorante, que de cualquier cosa que ve, para él extraordinaria, se queda admirado y pasmado» [fecha de consulta: 12 de agosto de 2015].

he had returned to rob a bank not ten blocks from here , up by Wolfie's Deli .Now it was the hip place to be in South Florida. Guys with sunglasses in their hair posed **skinny** girls on the beach and photographed them

Con respecto al otro caso, la connotación peyorativa viene dada más claramente por la palabra que sigue a *skinny*, *redneck* (que se refiere a alguien paleta, cerril o reaccionario):

<Leonard>: Estaba ocupado viendo pasar a las chicas nazis , aquellas jóvenes <delgaduchas> con el pelo corto como los chicos .

-->par_en: He was busy watching the Nazigirls , all of them **skinny rednecks** , their hair cut short as boys ' .

Elemento activador EN	Frec.
Palabra de significado inherente apreciativo	4
Sufijo apreciativo	
Adjetivo o adverbio con connotación apreciativa + sustantivo	
Información contextual	1
Tono general del fragmento (razones estilísticas o compensación)	
Equivalente funcional	
Información distribuida en más de un elemento lingüístico	1
Ø	
Total	6

Tabla 173 - Elementos activadores del sufijo *-ucho* en traducciones

6.3.3.10. -UTE

Se trata de un sufijo con carga humorística y despectiva. A pesar de aparecer en seis de las obras consultadas, no hemos encontrado otros ejemplos de uso más que el que nombran todos los autores, aplicado a un gentilicio: *franchute*. Ahora bien, si consultamos a Ulašín (2006: 65), veremos que también incluye la variante *-uta* en su clasificación de sufijos del argot español, lo que le permite proporcionar otros ejemplos, como *camaruta* (camarera de un bar de alterne) *chaluta* (chalé) o *manguta* (delincuente).

`[(word=".*ut(a|e|o)(s?)") & (word!="puta(s?)|minuto(s?)" %c)];`

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
camaruta	2	0,5		0	0
Total	2	0,5		0	0

Tabla 174 - Sufijo *-ute* en ES y TRAD

Evidentemente, no es este un sufijo de uso muy extendido en nuestro corpus. La sorpresa ha sido hallar dos ejemplos de *camaruta* y no del nombradísimo *franchute*, que era lo que esperábamos. El único término que se asemeja a estos y aparece tanto en ES como en TRAD es *canuto*, pero su origen, según el DRAE, es mozárabe. Las diferencias entre subcorpus son, como es lógico, no significativas.

<Madrid>: Un poli es un poli siempre , te lo digo yo , me cago en la leche . Mira , yo no me veo de **<camaruta>** , qué quieres que te diga . Aunque ahora mismo , dueño del bar como soy , pues bueno ,

<Martín>: Encontré a mis choricillos , como quien no quiere la cosa , en ; un bar de **<camarutas>** , casi a la hora de cerrar . Ellos no me vieron . Yo apenas me asomé , los localicé

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Madrid	1	2,4	Grafton	0	0
Martín	1	1,8	Leonard	0	0
Giménez	0	0	Ellroy	0	0
Martínez	0	0	Cornwell	0	0
Núñez	0	0	Highsmith	0	0
Vázquez	0	0			

Tabla 175 - Distribución del sufijo *-ute* por texto

6.3.3.11. -UZO / -UZA

No hemos encontrado más información referida a este sufijo despectivo que la proporcionada por Nández Fernández (1973: 385-386), quien comenta el desatino de algunos gramáticos al considerarlo, junto con otros grupos de palabras, diminutivo:

Al tomar como significación básica la pequeñez, ciertos gramáticos, desorientados por aferrarse a la definición etimológica, han incluido bajo la misma denominación otras palabras derivadas que significaban objetos pequeños o más pequeños que los positivos (ventorro), o crías de animales (cachorro, lobato, lobezno, palomino, etc.). Pero en realidad se refieren a seres perfectamente diferenciados objetivamente. No obstante dichos términos tenían en común con el diminutivo un indudable contenido afectivo, lo mismo que sucede con ciertos hipocorismos apreciativos (Paco, Pepe, etc.) por lo que en ambos casos dichas formas han sido puestas bajo el mismo denominador común: el diminutivo. (...) Por esta razón a veces se han incluido otros sufijos en la lista de los diminutivos, a causa de esa proximidad de significación (-istrajo, bebistrajo; -uza, gentuza etc.), con manifiesta desorientación.

Ejemplos: borrachuzo, pestuzo

[word=".*uz(o|a)(s?)"];

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
gentuza	4	1,1		0	0
Total	4	1,1		0	0

Tabla 176 - Sufijo *-uzo* en ES y TRAD

Solo hemos encontrado la palabra *gentuza*, utilizada cuatro veces en ES. También aparece *chapuza* (obra hecha sin esmero), tanto en ES como en TRAD, y *escaramuza*, solo en TRAD, pero el DRAE nos aclara que estos términos provienen del francés antiguo y del italiano, respectivamente. Al haber cuatro apariciones en el subcorpus de originales y ninguna en el de traducciones, que, como sabemos, contiene mayor número de palabras, el test estadístico da los resultados como significativos, aunque por poco margen (LL +6,77 $p < 0,01$)²²¹ —menos todavía si consultamos el resultado del test de Fisher—.

<Madrid>: Eso es lo que ha traído esta <gentuza> que nos gobierna , el despido libre y la falta de respeto por el trabajador - dijo - .

<Martínez>: Si no conseguía mi objetivo , quedaría sumido en el peor de los ridículos , a merced de la peor <gentuza> que había conocido en mis años de profesión .

ES	Frec.	Frec. norm.	TRAD	Frec.	Frec. norm.
Madrid	2	4,8	Grafton	0	0
Martínez	1	1,8	Leonard	0	0
Núñez	1	2,3	Ellroy	0	0
Giménez	0	0	Cornwell	0	0
Martín	0	0	Highsmith	0	0
Vázquez	0	0			

Tabla 177 - Distribución del sufijo *-uzo* por texto

²²¹ Test exacto de Fisher: 0,033877 ($p < 0,05$)

6.3.4. *Análisis de otras alternativas de traducción*

En este apartado analizaremos el corpus de originales ingleses para averiguar qué tienden a hacer los traductores al encontrarse palabras del campo semántico de lo grande y lo pequeño. También compararemos los corpus ES y TRAD en relación con ciertos elementos que pueden considerarse alternativas de traducción a los sufijos apreciativos en español para dichas palabras de EN, dado que expresan matices similares. Dada la imposibilidad de dar cuenta del universo completo de elementos activadores ingleses de los sufijos apreciativos, puesto que, a diferencia de lo que ocurría en el caso del sistema de demostrativos, no se trata de un conjunto cerrado, sino todo lo contrario —y lo mismo puede decirse acerca de las alternativas de traducción españolas—, nos vemos obligados a centrar nuestro análisis en algunos elementos concretos que consideramos más representativos.

No es factible intentar localizar todos los ejemplos de la primera categoría de elementos activadores por frecuencia de uso, las palabras de significado apreciativo inherente, pues, como hemos visto, estaríamos hablando de todo tipo de sustantivos, adjetivos y expresiones de naturaleza muy diversa. Nos fijaremos, entonces, en los elementos que pueden distinguirse morfológicamente, es decir, los sufijos ingleses que mencionan Rojo y Schneider, y en algunas de las palabras más representativas del campo semántico de la expresión del tamaño: *pequeño* y *grande* en español y *small*, *little*, *tiny*, *big* y *large* en inglés. No nos ocuparemos de los peyorativos debido a que no es tan fácil como en los otros dos casos hallar términos genéricos que nos permitan obtener cantidades suficientes de concordancias analizando solo unos pocos elementos.

Con respecto a la primera categoría, la de los sufijos diminutivos, lo cierto es que no son muy abundantes en inglés, al menos los que hemos analizado en este trabajo. Hay que mencionar, no obstante, que hemos elegido no estudiar los sufijos *-y*, *-o*, *-s* y *-er* porque, a pesar de que alguno sería probablemente bastante productivo como diminutivo —suponemos que sobre todo *-y*—, darían lugar a una cantidad inabarcable de concordancias, la mayor parte de las cuales resultarían irrelevantes.

Los sufijos *-poo*, *-pegs* e *-een* no arrojan resultados en nuestro corpus, al menos de palabras diminutivas. Con *-kin* solo hallamos *napkin* (*servilleta*), que fue diminutivo en sus orígenes pero ya no se percibe como tal. *-Ette* presenta en el corpus sobre todo

palabras de origen francés, como *brunette* (traducida como *morena* o *morenita* en nuestro corpus) o palabras lexicalizadas como *cigarette* (*cigarrillo*). Sin embargo, sí encontramos una palabra de origen estadounidense formada mediante este sufijo: *dinette* («a small space or alcove, often in or near the kitchen, serving as an informal dining area»²²²), traducido simplemente como *comedor* en el corpus. Con *-let* encontramos dos derivados diminutivos: *boardlet* (*tabla de planchar*), y *townlet* (*pueblo*). Como vemos, solo en un caso se activa la utilización de diminutivos con los sufijos analizados.

Los dos sufijos más productivos de los analizados son *-ey* e *-ie*. Casi todas las palabras a las que se añaden son nombres propios, como Davey y Eddie. Sin embargo, con el primer sufijo hallamos 2 casos del coloquial *gooey* (*pegajoso* o *empalagoso*, traducido como *jubiloso* y *dulce*), 4 de *homey* (*hogareño* como adjetivo o *amigo/colega* como sustantivo; traducido como *tipo* y *de aquí*) y 1 de *pricey* (equivalente informal de *caro*, traducido como *lujoso*). Con el segundo, aún más: 2 casos de *baggie* (*bolsita*), 2 de *beanie* (*casquete*), 5 de *blondie* (*rubio* y *rubito*), 6 de *bookie* (*apostador*, *apuestas*), 2 de *cheapie* (*película barata*, *de bajo presupuesto*), 10 de *Commie* (*comunista*, *rojillo*), 1 de *girlie* (*de chicas*), 2 de *hankie* (*de handkerchief*, traducido como *llorera* porque va seguido de la palabra *binge*), 1 de *indie* (*independiente*), 6 de *prostitie* (*prostituta*, *puta*), 5 de *juvie* (*juvenil*), 3 de *nightie* (*camisón*), 2 de *Okie* (*de Oklahoma*), 1 de *piggie* (*cerdo*), 4 de *sweetie* (*encanto*) y 1 de *weenie* (*minga*): 53 ejemplos, sin contar los ya quizás demasiado lexicalizados *junkie* (11 casos, traducidos por *yonqui*, *drogadicto* o *enganchado*) 8 *walkie-talkie* (tal cual o traducido como *emisor-receptor*) y 87 *movie* (*cine*, *película*, *filmación*).

En las traducciones de estos dos últimos sufijos observamos mayor presencia de diminutivos, aunque la relación tampoco es sistemática, ni mucho menos (8 diminutivos de 60 concordancias).

Sufijo EN	Suf. apr. TRAD	No suf. apr. TRAD
-een	0	0
-ette	0	1
-ey	0	7
-ie	8	45
-kin	0	0
-let	0	2
-pegs	0	0
-poo	0	0
Total	8	55

Tabla 178 - Uso de sufijos apreciativos españoles en la traducción de sufijos ingleses

²²² *dinette*. *Dictionary.com Unabridged* [fecha de consulta: 15 de agosto de 2015].

Ya hemos dicho que en inglés no existen sufijos aumentativos. Sí hay prefijos de este tipo, como *-arch(i)*, *-macro*, *-maxi-*, *-mega*, *-out*, *-super-* y *-ultra* (Huddleston y Pullum 2002: 1678-1680), pero no los hemos encontrado representados en nuestro corpus, a excepción de algunos derivados lexicalizados de *-ultra*, como *ultraviolet*, de *-super*, como *supermarket* (*supermercado* o *grandes almacenes*, en TRAD) o *superintendent* (*conserje* o *portero*), y de *-out*, como *outfox* (*pegársela*), *outmaneuvered* (*despistarlo*), *outdid* (*se excedió*), *outnumber* (*eran más que*), *outsize* (*inmensa*), *outweigh* (*le saco más de veinte kilos*) o *outrun* (*eludir*).

Con respecto a la estrategia analítica, es decir, la formación de diminutivos, aumentativos y peyorativos a través de adjetivos con significado apreciativo, queremos saber si se utiliza de la misma manera en textos originales españoles y textos traducidos al español. Lo primero que haremos será analizar las concordancias de la expresión regular «pequeñ.*», exceptuando los casos en los que equivale a *niño* o se utiliza como vocativo. Con respecto a este último caso, solo encontramos un ejemplo en ES, pero durante el análisis hemos comprobado que este uso es bastante común en TRAD con otras palabras como *nena* o *chato*, sobre todo en la traducción de *baby*. Aquí, no obstante, solo hemos encontrado 7 casos (siempre en femenino).

Hallamos 123 casos de «pequeñ.*» en ES y 212 en TRAD (las cifras normalizadas serían 33,5 y 43). La diferencia no es muy grande pero alcanza para ser considerada estadísticamente significativa (LL -5,31 $p < 0,05$). Por tanto, al menos en nuestro corpus, *pequeño/a* se utiliza más en traducciones que en originales españoles para formar diminutivos.

Ahora nos interesa saber qué hacen los traductores al encontrarse con estos apreciativos analíticos en inglés; es decir, si siguen la misma estrategia (y por eso hallamos más estructuras analíticas en traducciones) o se decantan por otras opciones, entre las que pueden encontrarse nuestros sufijos apreciativos. Buscamos todas las concordancias de *small* y *little*. Los resultados son los siguientes.

Como traducciones de *small* (184 casos en EN) hallamos 102 *pequeño/a(s)* (un 55,4 %), 27 diminutivos (14,7 %), utilizados especialmente por la traductora de Highsmith, Núñez (13 casos), y el de Cornwell, Mustieles, (con 12, aunque 2 son lexicalizados: *mesita* y *esterilla*). En 31 casos (16,8 %), *small* no se traduce al español y en 24 (13 %), encontramos sinónimos aproximados con el significado de pequeño o

breve, como *escaso, reducido, menudencia, bajo, de poca monta, estrecho, leve, fácil, tontería, cierto, (un/muy) poco, algo* y otros equivalentes como *charlar, conversación y palique* para *small talk*; *riñones, espalda* o *región lumbar* para *small of one's back*; *gatera* para *small opening* o *no ser de extrañar* para *no small wonder*.

Como traducciones de *little* (317 casos en EN), dado que no funciona solo como adjetivo, sino también, por ejemplo, como sustantivo y adverbio, encontramos mucha variedad de elementos. En 46 ocasiones se traduce por un diminutivo (un 14,5 %, aunque 2 son lexicalizados: *coleta* y *mariquita*). No se traduce por nada en 76 ocasiones (24 %, aunque 9 son nombres propios que se transfieren tal cual, como *Little Rock* o *Little Venice*). Esta vez, la utilización de *pequeño/a* no es tan abundante (solo 45 casos, un 14,2 %). El resto (150 casos, 47,3 %) son expresiones bastante variadas; sobre todo *(un/muy) poco*, pero también encontramos casos de *un tanto, algo, algún(a), un/a cierto/a, apenas, leve(mente), ligero/a(mente), escaso, lo mínimo, casi, prácticamente, un rato, estrecho, diminuto, insignificante, breve, canijo, relativo, ridículo, en seguida*, y también palabras de sentido diminutivo inherente (algunas con matices claramente peyorativos) como *cuchitril, chucho, meñique, infantil* o *adolescente*.

Si juntamos ambos términos, tenemos 501 concordancias, de las que 147, un 29,3 % corresponden a *pequeño*, 73 (14,6 %) se traducen por un diminutivo, 107 (21,3 %) no se trasladan a la traducción, y un 34,7 % (174 casos) corresponden a diversas expresiones con sentido diminutivo o equivalentes funcionales.

Existen otros elementos que podríamos buscar en EN, como *short, young, baby* o *kid*, pero son muy variados y no proporcionan tantas concordancias. Por ejemplo, *tiny* cuenta con 30 concordancias; de ellas, 6 corresponden a *pequeño/a* (20 %), 5 a diminutivos (16,7 %), 1 no se traduce (3,3 %), y las 18 restantes (60 %), corresponde a diversos equivalentes, de los cuales el más numeroso es *minúsculo*, con 9 casos, pero también están *diminuto, menudo, insignificante, leve, poco* y *nada*.

Con respecto a los aumentativos, hemos decidido buscar en ES y TRAD todas las apariciones de «grande/gran/grandísimo». En ES hallamos 200 (quitando 8 nombres propios conocidos por todos y que, por tanto, no proporcionan libertad de traducción, como *la Gran Vía* o *el Gran Gatsby*), y en TRAD, 336 sin contar *el Gran Zoco* ni expresiones referidas a dinero, como «cincuenta de los grandes» (*fifty grand*: cincuenta mil dólares). Las frecuencias normalizadas son 54 y 69, respectivamente. La diferencia

a favor de TRAD es significativa, algo más que en el caso de *pequeño/a* (LL -6,92 $p < 0,01$).

Si analizamos las concordancias de *big* en EN (288 casos), vemos que se traduce por *grande* 136 veces (47,2%), 37 veces no se traduce por nada (12,8%) y solo 10 veces por un sufijo —3,5% o 10,8% o si contamos también el mote *Grandullón*, que aparece en 21 ocasiones—, aunque de las 10 deberíamos eliminar también las lexicalizadas *barrote*, *boquete* y *matón*. Las 84 restantes (un 29,2%) se traducen por sinónimos o expresiones diversas que dan a entender gran cantidad, importancia o tamaño: *mucho/a*, *buen/a*, *tanto/a*, *todo/a*, *importante*, *un montón*, *corpulento*, *gordo*, *fuerte*, *ancho*, *mole*, *largo*, *mayor*, *amplio*, *pesado*, *grueso*, *robusto*, *voluminoso*, *enorme*, *aparatoso*, *inmenso*, *absoluto*, *enorme*, *de primera*, *en cantidad*, *bastante*, *colosal*, *genial*, *en tropel*, *general*, *de aquí te espero*, *sonado*, *multitudinario*, *destacado*, *sagrado*, *peor*, sustantivos como *saca* o *bidón* (para *big bag* y *big can*) y equivalentes de expresiones varias como *limpiar* (*to rob big*) *entusiasmar* (*to be a big fan*), *no pasar nada* (*to be no big deal*), *pasta* (*big money*) o *adulto* (*big boy*).

Existen muchos otros términos que podríamos haber estudiado, como *giant* o *huge*, pero no proporcionan tantas concordancias. De todas formas, hemos querido analizar uno más, *large*, con 37 casos (no contamos los cuatro de *at large*). 21 se traducen por *gran/grande(s)* (56,7%), 3 se dejan sin traducir (8,1%), no existen casos de aumentativos, y el resto de concordancias (35,2%) corresponden a equivalentes como *amplio*, *extenso*, *espacioso*, *considerable*, *importante*, *elevado*, *suficiente*, *concurrido* y (tamaño) *familiar*. Estos porcentajes no se alejan demasiado de las de los equivalentes de *big*.

Palabra EN	Suf. Apr.	%	Pequeño/grande	%	Otros equivalentes	%	Nada	%	Total
Small	27	14,7	102	55,4	24	13	31	16,8	184
Little	46	14,5	45	14,2	150	47,3	76	24	317
Tiny	5	16,7	6	20	18	60	1	3,3	30
Big	31	10,8	136	47,2	84	29,2	37	12,8	288
Large	0	0	21	56,7	13	35,2	3	8,1	37
Total	109	12,7	310	36,2	289	33,8	148	17,3	856

Tabla 179 - Adjetivos de tamaño en EN y sus equivalentes en TRAD

Cuando agrupamos todos los elementos analizados, vemos que los equivalentes más comunes son los adjetivos *pequeño* y *grande* y, en segundo lugar, la categoría de otros elementos equivalentes diversos. Los sufijos apreciativos, como era de esperar, constituyen la opción menos habitual de estas cuatro, aunque esta tendencia es más

marcada en *big* y *large*, puesto que en *small*, *little* y *tiny* ocupan el tercer lugar, y en *little*, la categoría «pequeño/grande» baja al cuarto puesto, debido al abundante uso de expresiones como *un poco* para traducir *a little*.

Finalmente, hemos querido comparar el uso de *pequeño* y *grande* en ambos subcorpus, en relación con la posición del adjetivo. Varios autores advierten de la tendencia, ya sea por influencia del francés o del inglés, de anteponer el adjetivo al sustantivo en situaciones en las que no estaría justificado en español:

Una advertencia importante procede hacer al hablar del diminutivo español. La de evitar el galicismo que se puede cometer en el empleo indebido de *pequeño* como marca de diminutivo. Por ejemplo, algunas veces al decir *pequeña mujer* que no suele ser correcto en español, a diferencia de *mujer pequeña* que sí suele serlo (Cantera Ortiz de Urbina 2012: 99)

Y llegamos a los “novísimos”, los jóvenes iconoclastas de los noventa, también creadores y destructores de lengua. Ray Loriga en *Caídos del cielo* (1995) nos cuenta que el hermano del *enano* era guapísimo, o sea “más guapo que la hostia” y que “estaba la hostia de orgulloso con aquella chica que tenía aún sus *tetitas* al aire”. Abandona la sufijación del castellano para adoptar una sintaxis anglosajona: “sería una preciosa, honesta chica muerta”, “tu *pequeña puta*”, “no hablaban ni del hombre malo ni de su *pequeña mujer*”. Cogió una *papeleta* de coca porque los vigilantes le daban el *coñazo*, ya ves qué *putada*. (Soler Espiauba 1996)

Estos autores hablan de novelas originales escritas en español. Sin embargo, hemos querido comprobar si los textos de partida ingleses ejercen algún tipo de influencia sobre la posición del adjetivo en las traducciones al español. Para ello, hemos contado todos los casos en los que el adjetivo *pequeño* se encuentra antepuesto al sustantivo tanto en ES como en TRAD y hemos calculado el porcentaje con respecto al total de concordancias válidas de *pequeño* de cada subcorpus:

ES: 71 de 123 - 58%
TRAD: 124 de 212 - 58%

Asombrosamente, los porcentajes coinciden en ambos subcorpus, por lo que si esta anteposición del adjetivo es un fenómeno que va en aumento en el español contemporáneo, no podemos deducir de este estudio que sea un fenómeno propiciado o agravado por las traducciones del inglés. Hacemos lo mismo con *gran/grande*, sin embargo, y los porcentajes se distancian un poco más, haciendo las diferencias significativas (LL +3,98 $p < 0,05$), pero en la dirección opuesta a la que esperábamos, pues la estructura es más frecuente en español no traducido:

ES: 160 de 200 – 80 %
TRAD: 217 de 336 – 65 %

6.3.5. *Datos agrupados*

En la tabla 180 agruparemos los resultados por grupos y, posteriormente, realizaremos los cálculos para todos los elementos juntos. Como vemos, en todos los grupos hay sufijos que no presentan diferencias significativas de uso entre subcorpus (recordemos que las casillas amarillas indican diferencia a favor de ES no significativa y las fucsia, diferencia a favor de TRAD no significativa). Sin embargo, las cifras finales de cada grupo y de todos los grupos en conjunto muestran una diferencia estadística muy significativa. Al parecer, si el sufijo en cuestión se utiliza habitualmente en español y cuenta con suficientes concordancias en nuestro corpus, estará mejor representado en los textos escritos originalmente en español (por ejemplo, *-illo* e *-ito* en los diminutivos u *-ón* en los aumentativos). Hay que tener en cuenta que los diminutivos cuentan con casi el triple de casos que los de los otros dos grupos sumados, por lo que su resultado como grupo influye indudablemente en el conjunto. No obstante, en ningún caso hallamos la tendencia contraria (que aparecería en la tabla como una casilla roja); es decir, que haya una diferencia significativa a favor de los textos traducidos, como sí sucedía con los adverbios demostrativos o algunas perífrasis. Por consiguiente, si no tuviéramos en cuenta los elementos sin diferencias significativas entre subcorpus, el resultado sería una clara infrarrepresentación de los restantes.

DIM	ES	Frec. norm.	TRAD	Frec. norm.	LL	p	Verif. Hipótesis
ejo/a	10	2,7	8	1,6	+1,16	n. s.	NO
ete/a	20	5,4	5	1	+14,43	< 0,001	SÍ
ico/a	1	0,3	2	0,4	-0,12	n. s.	NO
illo/a	74	20,1	31	6,3	+32,57	< 0,0001	SÍ
ín, ino/a, iño/a	2	0,5	2	1,8	+0,4	n. s.	NO
ito/a	342	93,2	344	70,4	+13,10	< 0,001	SÍ
uco/a	2	0,5	5	1	-0,61	n. s.	NO
uelo/a	7	1,9	11	2,2	-0,12	n. s.	NO
Total DIM	458	124,8	408	83,5	+34,81	< 0,0001	SÍ
AUM	ES	Frec. norm.	TRAD	Frec. norm.	LL	p	Verif. Hipótesis
azo/a	20	5,4	13	2,7	+4,17	< 0,05	SÍ
ón/a	64	17,4	21	4,3	+36,81	< 0,0001	SÍ
ote/a	17	4,6	4	0,8	+12,81	< 0,001	SÍ
udo/a	15	4,1	13	2,7	+1,28	n. s.	NO
Total AUM	116	31,6	51	10,4	47,96	< 0,0001	SÍ
PEY	ES	Frec. norm.	TRAD	Frec. norm.	LL	p	Verif. Hipótesis
acho/a	2	0,5	1	0,2	+0,69	n. s.	NO
aco/a	8	2,2	0	0	+13,54	< 0,001 ²²³	SÍ
ajo/a	4	1,1	7	1,4	-0,2	n. s.	NO
ales	0	0	0	0	0	n. s.	NO
ango/a, etc.	2	0,5	4	0,8	-0,23	n. s.	NO
astro/a/e	1	0,3	1	0,2	+0,04	n. s.	NO
ata/o	9	2,4	2	0,4	+7,04	< 0,01 ²²⁴	SÍ
orro/a, etc.	5	1,4	0	0	+8,46	< 0,01 ²²⁵	SÍ
ucho/a	3	0,8	6	1,2	-0,35	n. s.	NO
ute/a	2	0,5	0	0	+3,39	n. s.	NO
uzo/a	4	1,1	0	0	+6,77	< 0,01 ²²⁶	SÍ
Total PEY	40	10,9	21	4,3	12,69	< 0,001	SÍ
TOTAL	614	167,3	480	98,2	+77,03	< 0,0001	SÍ

Tabla 180 - Frecuencias de uso de sufijos apreciativos en ES y TRAD

Aun así, el valor p del grupo de los peyorativos es inferior al de los otros dos grupos. Si contamos cada sufijo peyorativo por separado, los casos en los que las diferencias no son significativas (7) superan a los otros (4). Sin embargo, no tiene sentido contarlos por separado, dadas las frecuencias de uso tan bajas que presenta cada uno. Como grupo, están claramente infrarrepresentados en traducciones.

Si hiciéramos el ejercicio propuesto por Chesterman —que aplicamos al grupo de perífrasis en el primer capítulo— para identificar elementos únicos *a posteriori*, estableciendo el umbral en un valor p de $< 0,001$, tendríamos que quedarnos únicamente con los diminutivos *-ete/a*, *-illo/a* e *-ito/a* y los aumentativos *-ón/a* y *-ote/a*. Ningún peyorativo presenta diferencias de uso significativas según el umbral establecido si tenemos en cuenta el resultado del test de Fisher.

En la tabla 181 sumamos la cantidad de veces que utiliza cada autor los sufijos apreciativos, primero por grupos y, posteriormente, en conjunto. Las frecuencias están

²²³ Test exacto de Fisher: 0,001148 ($p < 0,01$)

²²⁴ Test exacto de Fisher: 0,012349 ($p < 0,05$)

²²⁵ Test exacto de Fisher: 0,014534 ($p < 0,05$)

²²⁶ Test exacto de Fisher: 0,033877 ($p < 0,05$)

normalizadas con respecto a la cantidad de palabras de cada texto. Con verde marcamos el autor que más utiliza cada sufijo por separado y con rojo, los autores que menos lo hacen.

DIM	Giménez	Madrid	Martín	Martínez	Núñez	Vázquez
ejo/a	3,5	0	5,3	1,8	2,3	2,3
ete/a	1,2	12,1	14,2	1,8	2,3	4,5
ico/a	0	0	1,8	0	0	0
illo/a	21,2	38,6	30,2	14,7	18,8	8
ín, ino/a, iño/a	0	0	3,5	0	0	0
ito/a	57,7	217,2	99,6	57	164,8	52,4
uco/a	0	0	3,5	0	0	0
uelo/a	1,2	0	1,8	0	7,1	2,3
Total DIM	84,8	267,9	159,9	75,3	195,3	69,5
AUM	Giménez	Madrid	Martín	Martínez	Núñez	Vázquez
azo/a	8,2	2,4	8,9	3,7	7,1	2,3
ón/a	11,8	21,7	24,9	14,7	25,9	13,7
ote/a	0	26,5	3,5	0	2,3	3,4
udo/a	5,9	9,6	5,3	0	4,7	1,1
Total AUM	25,9	60,2	42,6	18,4	40	20,5
PEY	Giménez	Madrid	Martín	Martínez	Núñez	Vázquez
acho/a	0	0	1,8	0	0	1,1
aco/a	1,2	0	5,3	0	7,1	1,1
ajo/a	2,3	0	0	1,8	0	1,1
ales	0	0	0	0	0	0
ango/a, etc	0	0	1,8	0	0	1,1
astro/a/e	0	0	0	0	0	1,1
ata/o	1,2	0	1,8	3,7	4,7	3,4
orro/a, etc.	1,2	0	5,3	0	0	1,1
ucho/a	0	2,4	1,8	0	0	1,1
ute/a	0	2,4	1,8	0	0	0
uzo/a	0	4,8	0	1,8	2,3	0
Total PEY	5,9	9,6	19,6	7,3	14,1	11,1
TOTAL	116,6	337,7	222,1	101	249,4	101,1

Tabla 181 - Frecuencias de uso normalizadas de sufijos apreciativos en ES por texto

Si contamos las casillas verdes, los autores que más parecen utilizar sufijos apreciativos son Martín (9), Madrid (7) y Núñez (4), y las cifras lo confirman, pero el orden cambia: Madrid (337,7), Núñez (249,4) y Martín (222,1). Recordemos, de todas formas, que hablamos de cifras normalizadas.

Martín es quien presenta frecuencias más altas en más cantidad de sufijos diminutivos (5), es decir, que utiliza mayor variedad de diminutivos (de hecho, los utiliza todos), pero Madrid, que solo gana en dos ocasiones (y, en realidad, solo utiliza tres sufijos diferentes), es quien cuenta con más casos de diminutivos, puesto que esos dos (-illo e -ito) son los más abundantes en el subcorpus. Núñez ocupa el segundo puesto en frecuencia de uso de diminutivos, antes que Martín, y cuenta con casos en cinco de las ocho categorías de diminutivos.

En los aumentativos vuelve a ganar Madrid, tanto por número de sufijos como por frecuencia de uso de dichos sufijos. En los peyorativos, hay un empate entre Madrid

y Martín en cuanto a casillas verdes, pero Martín gana en frecuencia de uso, y Núñez le sigue. Sin embargo, Martín es quien, una vez más, utiliza mayor variedad de sufijos (siete de las once categorías de peyorativos), mientras que los otros dos se quedan en tres cada uno.

Es curioso constatar que cada uno de los autores mencionados proviene de una comunidad lingüística diferente: Madrid de Andalucía, Martín de Cataluña y Núñez de País Vasco.

Los que menos sufijos parecen utilizar son Giménez-Bartlett, Martínez Reverte y Vázquez Montalbán, aunque este último gana a Madrid tanto en variedad como en frecuencia de uso de peyorativos (utiliza pocos pero muy variados).

En cuanto a las traducciones, Moya (Grafton) y Sabaté Vargas (Ellroy) son los que más variedad de sufijos utilizan en mayor medida que el resto; sin embargo, no son los que mayor frecuencia de uso presentan: en este sentido ganan los traductores Núñez (Highsmith) y de Hériz Ramón (Leonard). Mustieles (Cornwell) utiliza poca variedad de sufijos y menos veces que los demás.

DIM	Cornwell	Ellroy	Grafton	Highsmith	Leonard
ejo/a	0	6,4	0	0	1,3
ete/a	0,9	0,9	2,1	0	1,3
ico/a	0	1,8	0	0	0
illo/a	1,8	11,8	9,7	4,1	3,8
ín, ino/a, iño/a	1,8	0	0	0	0
ito/a	51	62,9	60,3	92,6	92,5
uco/a	0	0	5,4	0	0
uelo/a	0,9	2,7	5,4	2	0
Total DIM	56,4	86,5	82,9	98,7	98,9
AUM	Cornwell	Ellroy	Grafton	Highsmith	Leonard
azo/a	1,8	2,7	0	3	6,4
ón/a	1,8	4,5	9,7	2	3,8
ote/a	0	0,9	0	0	3,8
udo/a	0,9	1,8	5,4	1	5,1
Total AUM	4,5	9,9	15,1	6	19,1
PEY	Cornwell	Ellroy	Grafton	Highsmith	Leonard
acho/a	0	0,9	0	0	0
aco/a	0	0	0	0	0
ajo/a	0	1,8	0	5,1	0
ales	0	0	0	0	0
ango/a, etc	0	2,7	1,1	0	0
astro/a/e	0	0,9	0	0	0
ata/o	0,9	0	0	0	1,3
orro/a, etc.	0	0	0	0	0
ucho/a	0,9	2,7	0	0	2,6
ute/a	0	0	0	0	0
uzo/a	0	0	0	0	0
Total PEY	1,8	9	1,1	5,1	3,9
TOTAL	62,7	105,4	99,1	109,8	121,9

Tabla 182 - Frecuencias de uso normalizadas de sufijos apreciativos en TRAD por texto

Si hablamos de los diminutivos por separado, el panorama es exactamente el que hemos pintado para todos los sufijos en conjunto. Sin embargo, la cosa cambia en los aumentativos, pues de Hériz Ramón cuenta con mayor variedad y mayor frecuencia de uso. Moya lo sigue en frecuencia, pero debido únicamente al elevado uso de dos de los cuatro sufijos y, en tercer lugar, está Sabaté Vargas, que utiliza los cuatro, pero en menor medida. En cuanto a los peyorativos, Sabaté Vargas es sin duda quien se lleva la palma, pues utiliza cinco de las once categorías de sufijos, mientras los demás apenas llegan a dos. Núñez alcanza el segundo puesto exclusivamente debido a su frecuencia de uso de *-ajo*.

En este caso es difícil extraer conclusiones relacionadas con el origen de los traductores, pues todos han nacido en Cataluña o residido allí durante años.

Como ya hemos advertido, la frecuencia de uso de diminutivos es tan alta comparada con los otros dos grupos que los resultados finales se acercan mucho a los de este. Incluso dentro del grupo de los diminutivos, hay que tener en cuenta que los resultados van a estar fuertemente influidos por los del sufijo que mayor frecuencia de uso presenta, *-ito*.

Elemento activador EN	DIM	%	AUM	%	PEY	%	Frec. total	%
Palabra de significado inherente apreciativo	139	34,1	8	15,7	17	80,9	164	34,2
Sufijo apreciativo	33	8,1	13	25,5	0	0	46	9,6
Adjetivo o adverbio con connotación apreciativa + sustantivo	94	23	10	19,6	1	4,8	105	21,9
Información contextual	44	10,8	6	11,8	1	4,8	51	10,6
Tono general del fragmento (razones estilísticas o compensación)	11	2,7	2	3,9	1	4,8	14	2,9
Equivalente funcional	34	8,3	5	9,8	0	0	39	8,1
Información distribuida en más de un elemento lingüístico	32	7,8	4	7,8	1	4,8	37	7,7
∅	21	5,1	3	5,9	0	0	24	5
Total	408	100	51	100	21	100	480	100

Tabla 183 - Frecuencia y porcentaje de uso de elementos activadores de la utilización de sufijos apreciativos en TRAD

En la tabla 183 vemos que la categoría que más activa la utilización de sufijos en textos traducidos es la de palabras con connotaciones diminutivas, aumentativas o peyorativas inherentes; sobre todo en el caso de los peyorativos, en los que suma más del 80 % de las concordancias, y algo menos en el de los aumentativos, en los que el porcentaje baja a menos de la mitad. En este último grupo, la categoría que más concordancias agrupa es la de los sufijos ingleses (especialmente *-y*, como vimos en el

apartado correspondiente), a diferencia de los otros dos, en los que es casi anecdótica. Habíamos comentado en el primer apartado de este capítulo que no existe la sufijación aumentativa en inglés. Sin embargo, sabemos que los aumentativos, como los diminutivos, además de indicar tamaño, se utilizan en español para denotar significados apreciativos; y, como indican Huddleston y Pullum (2002: 1677) refiriéndose al inglés, «The term diminutive applies to affixes which indicate small size and also, by extension, ones which (additionally or instead) mark the off-spring of animals, affection or informality, resemblance or imitation».

El segundo lugar en los datos agrupados lo ocupan los adjetivos y adverbios con connotaciones apreciativas: en diminutivos y aumentativos está en torno al 20 % de las concordancias, y en el grupo de los despectivos, en un 4,8 %, pero ninguna categoría alcanza a superar ese porcentaje, puesto que ya hemos dicho que el 80 % de los sufijos de este grupo se activan con palabras de significado despectivo inherente y solo hay un caso dentro cada una de las siguientes categorías que cuentan con frecuencias de uso.

El tercer lugar lo ocupa la información contextual (en torno al 10 %). Luego vienen, como ya hemos mencionado, los sufijos, y detrás, los equivalentes funcionales, la distribución de la información en más de un elemento, ningún elemento activador y el tono general del fragmento.

6.4. Conclusiones

Lo que pretendíamos en este estudio era comparar el uso que se realiza de palabras derivadas mediante sufijos apreciativos (diminutivos, aumentativos y peyorativos) en textos escritos originalmente en español y textos traducidos a este idioma. Los sufijos apreciativos son, según las fuentes que hemos consultado, un elemento típico del español que no tiene equivalente formal en muchas otras lenguas, incluido el inglés, o al menos no lo tiene en medida similar a la que puede encontrarse en español, dado que se trata de un recurso morfológico mucho más productivo en nuestro idioma. Por este motivo, nuestra hipótesis inicial era que, al no encontrar estas estructuras en los originales ingleses, los traductores tenderían a utilizar otros recursos lingüísticos de similar contenido semántico y funcional pero más cercanos formalmente

a los de la lengua de partida y, por tanto, no utilizarían los sufijos apreciativos con la frecuencia que sería habitual si escribieran en español.

Para averiguar si nuestra hipótesis es válida o no, lo primero que hicimos fue buscar todos los casos de sufijos apreciativos en español original y español traducido. Al tratarse, a fin de cuentas, de textos escritos, lo cierto es que el repertorio que hemos recogido en nuestro corpus no es de los más variados ni originales —como pueden serlo, por ejemplo, los diminutivos mencionados en Nájiz Fernández (1973: 37-38), extraídos de la lengua oral—. Sin embargo, hemos encontrado suficientes casos para permitirnos extraer ciertas conclusiones con respecto a determinadas características del español traducido de los textos de nuestro corpus.

En primer lugar, es interesante fijarse en el hecho de que, en la mayor parte de los casos, no hace falta normalizar las frecuencias para darnos cuenta de que, efectivamente, la frecuencia de uso de sufijos apreciativos en los textos originales es más alta que la de los traducidos. Esto nos habla de una diferencia importante, pues el subcorpus de traducciones contiene mayor cantidad de palabras. Los resultados del test de significación estadística han respaldado nuestros resultados en 10 de los 23 sufijos o agrupaciones de sufijos analizadas, pero hay que tener en cuenta que otras seis mostraron diferencias a favor del subcorpus de originales a pesar de que estas diferencias no hayan resultado significativas desde el punto de vista estadístico. Otro factor importante es que ninguna de las siete restantes presentó el fenómeno opuesto, es decir, resultó sobrerrepresentada en el subcorpus de traducciones; por lo tanto, en todos los casos en los que existen diferencias que los tests estadísticos nos permiten tener en cuenta, hallamos un fenómeno de infrarrepresentación. Además, aquellas que no confirman nuestra hipótesis presentan, por lo general, unas frecuencias de uso muy bajas. Los sufijos apreciativos están, en conjunto, infrarrepresentados en los textos traducidos de nuestro corpus.

En segundo lugar, podemos afirmar que no solo hay menor cantidad de sufijos apreciativos en traducciones, sino que las bases sobre las que se forman los derivados apreciativos son menos variadas. Siempre teniendo en cuenta que el subcorpus de traducciones cuenta con mayor número de palabras (*tokens*) y, por tanto, cabría esperar que presentara más variedad de formas (*types*), es sorprendente encontrar que la

cantidad de bases diferentes que encontramos en TRAD supera a las de ES solo en tres ocasiones: en los sufijos *-ajo*, *-ango* y variantes y *-ucho*.

La conclusión es, por lo tanto, que este recurso morfológico, tan rico y productivo en español, está siendo utilizado por debajo de sus posibilidades en los textos traducidos. Por supuesto, hay diferencias individuales entre los distintos escritores y traductores. Por ejemplo, Madrid es quien hace el mayor uso de diminutivos de ambos subcorpus si nos basamos en las frecuencias normalizadas (265,5), pero solo de tres: *-ito*, *-illo* y *-ete*. En cambio, Sabaté Vargas, que traduce a Ellroy, cuenta con una frecuencia normalizada muy inferior (de 84,6) y, sin embargo, utiliza prácticamente toda la variedad de diminutivos a su alcance.

- ✓ A través de este estudio no podemos hacer afirmaciones sobre la prototipicidad de este recurso en español original, pues habría que comparar las frecuencias de uso de los sufijos apreciativos con toda la gama de recursos disponibles en nuestro idioma para expresar los mismos significados, y este es un grupo de elementos muy diverso y, además, abierto. Por lo tanto, en este caso no sabemos qué fuerza está ejerciendo el factor 1 de la Gravitational Pull Hypothesis. Aunque sí podríamos valorar la prototipicidad de determinados sufijos dentro de cada grupo, si nos atenemos a lo expresado por los gramáticos sobre su uso y a las frecuencias que arroja nuestro análisis. Por ejemplo, *-illo* e *-ito* serían las formas prototípicas del diminutivo, o al menos las más frecuentemente utilizadas, y, sin embargo, dos de las únicas tres infrarrepresentadas del grupo.
- ✓ Con respecto a la influencia que pueden ejercer la lengua o los textos originales en este fenómeno, es indudable que no existen prácticamente elementos formalmente equivalentes a nuestros sufijos apreciativos en ellos, lo que dificulta su uso en las traducciones. Apenas hay sufijos de este tipo en inglés y hemos comprobado que estos no suelen traducirse por palabras sufijadas en español. Los elementos que más activan la utilización de sufijos apreciativos son aquellas palabras que ya llevan incorporadas las connotaciones apreciativas y no necesitan de ningún recurso morfológico para expresarlo. La segunda categoría que más aparece como equivalente en inglés de los apreciativos son los adjetivos que modifican al sustantivo añadiéndole matices de significado. Hemos visto que en las traducciones se utilizan en mayor medida que en español original

sustantivos modificados por los adjetivos *grande* y *pequeño*, un fenómeno que puede deberse a la mayor presencia de estructuras formalmente equivalentes en inglés como *small/little/big* + sustantivo, pero no podemos asegurar que sea así, pues nuestro análisis indica que la utilización de dichos adjetivos en español surge de la traducción de estas estructuras de un 14 % a un 57 % de las veces: una media de solo el 38,7%. Por lo tanto, creemos que el factor 2 de la Gravitational Pull Hypothesis, la representación del concepto en la lengua de partida, favorece el fenómeno de la infrarrepresentación, pues no existe un elemento prototípico en inglés que pueda ejercer suficiente influencia en la elección de la utilización de estos sufijos.

- ✓ Sin embargo, este fenómeno no se limita a las palabras *grande* o *pequeño*, sino que, al igual que en el caso de los términos de significado apreciativo inherente, en nuestro análisis han aflorado multitud de palabras y expresiones diferentes cuyo sentido, como hemos visto a lo largo del capítulo, puede transmitirse en español de maneras muy diversas. Por lo tanto, en primer lugar, creemos que el factor 3 también favorece la infrarrepresentación, pues la variedad de elementos que pueden transmitir significados apreciativos es tal, tanto en inglés como en español, que los patrones de conexión entre una lengua y otra en este campo no están claramente marcados y delimitados y los enlaces no son directos ni mucho menos. Consideramos, además, que la utilización del diminutivo, aumentativo o peyorativo es, en gran medida, una elección libre del traductor, motivada probablemente por factores estilísticos o idiolectales. Los sufijos apreciativos son elementos con los que el escritor puede expresar matices, caracterizar maneras de hablar y personalidades y, sobre todo, jugar con el lenguaje y ser creativo. Quizás uno de los factores que tengan que ver con su menor utilización en traducciones sea justamente el menor grado de libertad con el que cuenta el traductor para jugar con la lengua si el escritor del texto original no lo hace (o no lo hace mediante este recurso, pues no pretendemos decir que los escritores estadounidenses sean menos creativos que los españoles, sino que, debido a las características intrínsecas del inglés, su creatividad reside en otras parcelas del lenguaje).

Para finalizar, podemos mencionar el hecho de que, según nuestros datos, los diminutivos son el grupo más utilizado tanto en español original como traducido y que

los peyorativos, a pesar de ser el grupo con más número de formas, es, como habíamos adelantado, el menos productivo. Sin embargo, el análisis de las concordancias nos dice que muchas de las veces en las que se usa un apreciativo es para transmitir matices despectivos o humorísticos; sobre todo en el caso de los diminutivos pero también en aumentativos como *-ón*. Por lo tanto, como advertía ya Alonso hace décadas, la sufijación no se encuentra tanto en casos en los que se busca transmitir el tamaño de algo, sino más bien cuando se intenta, además, valorar ese algo, y muy frecuentemente, según nuestros resultados, de manera negativa.

En cuanto a los elementos incluidos en la definición de *unique item* de Hareide podemos concluir lo siguiente:

- Has no direct counterpart in the source language: esta condición se cumple sin duda pues en inglés apenas hay sufijos que puedan expresar los mismos valores que los españoles.
- Has no salient or prototypical source language structure that may impact its choice in the target language: podría considerarse que los significados diminutivos, aumentativos y peyorativos están en los textos, sobre todo en forma de adjetivos, pero no es posible valorar qué grado de prominencia poseen.
- It is characterized by being infrequent and non-prototypical: hay que decir que este punto no es un buen predictor del grado de infrarrepresentación de un elemento en traducciones, pues los sufijos menos frecuentes y prototípicos (casi todos los peyorativos, *-udo* en el caso de los aumentativos y formas como *-uco* o *-uelo* en el caso de los diminutivos) son justamente los que no muestran diferencias de uso entre subcorpus. Por el contrario, formas prototípicas del primer grupo, como *-illo* e *-ito* muestran una clara infrarrepresentación en traducciones. Ahora bien, si consideramos el conjunto de los sufijos, no es posible decir a través de este estudio si son o no formas prototípicas en español para expresar valores apreciativos.

7. Conclusiones generales

El objetivo principal de este estudio era replicar investigaciones realizadas en el pasado en torno a la hipótesis de la infrarrepresentación de elementos únicos en el lenguaje traducido. Nuestro propósito era confirmar los resultados de estudios previos cuyas conclusiones apuntan a que los elementos típicos de una lengua que no cuentan con equivalentes directos o fácilmente identificables en la lengua de partida tienden a aparecer de manera menos frecuente en los textos traducidos a la lengua meta que en textos escritos originalmente en esa lengua. Esta característica del texto traducido ha sido propuesta como candidato a universal de traducción debido a que es muy posible que sea inherente al proceso traductor y no el resultado de los factores externos que pueden influir en él. Mediante la realización de un estudio con una metodología similar a la de aquellos llevados a cabo por autoras como Tirkkonen-Condit (2004) o Mauranen (2000), pero utilizando un conjunto de textos distinto, una combinación de lenguas diferente y elementos específicos del español, intentamos alterar los factores externos al proceso de traducción para averiguar si los resultados seguían siendo similares. Una respuesta afirmativa a este interrogante apoyaría la idea de que la Unique Items Hypothesis podría ser efectivamente una característica general de los textos traducidos.

Por otra parte, el trabajo realizado por Hareide (2013) para poner a prueba la Gravitational Pull Hypothesis de Halverson (2003) dejó abiertos varios interrogantes relativos a la definición de *unique item* y a las características necesarias para que un elemento se encuentre efectivamente infrarrepresentado en traducciones. Los resultados de Hareide contradijeron la Unique Items Hypothesis incluso en aquellos casos en los que no existían en la lengua original equivalentes inmediatos para los elementos estudiados que, por otra parte, eran muy frecuentes en la lengua meta: «My thesis thereby indicates that very frequent or prototypical target-language features appear to be over-represented in translations in accordance with Baker's (1993) hypothesis, even when they constitute unique items» (2003: 275). Por lo tanto, la autora se pregunta qué hace falta para que un supuesto elemento único se encuentre infrarrepresentado en traducciones. Un objetivo secundario de este trabajo era, por consiguiente, en los casos en los que la hipótesis principal se confirmara, identificar los factores que hacían posible dicha infrarrepresentación.

7.1. Compilación del corpus

Partimos de la compilación de un corpus de textos literarios contemporáneos de género policíaco formado por obras escritas originalmente en español peninsular y traducciones del inglés estadounidense al español, complementado con un componente paralelo de originales en inglés. La creación y explotación de este corpus nos ha permitido reflexionar sobre las ventajas y las limitaciones de la metodología de corpus, que ofrece al investigador posibilidades insospechadas hasta hace unas décadas, gracias a la enorme cantidad de datos que permiten manipular, de manera cada vez más refinada, las herramientas de análisis textual disponibles en la actualidad. No obstante, estas herramientas requieren una correcta comprensión de su funcionamiento y sus limitaciones para aprovechar al máximo sus capacidades.

Como hemos comprobado, los programas de análisis de corpus solo mostrarán los datos que el investigador les haya solicitado y solo podrán basarse en la información que se les haya proporcionado previamente. Por consiguiente, tanto la selección inicial de los textos que conformarán nuestro corpus, como el grado y la forma de preprocesamiento de dichos textos influirá en las respuestas que obtengamos a través de nuestras búsquedas. El ejemplo más ilustrativo en nuestro trabajo ha sido la comprobación de que el etiquetado puede afectar claramente a los resultados, no solo a causa de los errores del propio programa, que modificaban el cómputo final de elementos, sino también debido al desconocimiento de su funcionamiento por nuestra parte: en el caso de los signos de apertura de interrogación y exclamación, dado que no éramos conscientes al principio de que estos formaban una unidad con las palabras a las que precedían, estas palabras no aparecían en nuestras búsquedas. Finalmente, una de las limitaciones de los estudios basados en corpus que más se suelen mencionar es el hecho de que estos solo pueden proporcionar las respuestas a las preguntas que realice el investigador, por lo que los resultados serán siempre subjetivos. Esto es así, por supuesto, pero hemos comprobado a través de nuestro trabajo que muchas veces las respuestas a las preguntas iniciales nos llevan a hacer nuevas preguntas o los resultados nos muestran fenómenos que no nos habíamos planteado en un primer momento, por lo que, finalmente, todo estudio basado en corpus, creemos, se convierte en *corpus-driven*.

7.2. Selección de elementos y resultados

La compilación de este corpus nos ha permitido realizar un trabajo de comparación de frecuencias de una serie de elementos lingüísticos entre español traducido y no traducido —elementos que han servido para operacionalizar la hipótesis de Tirkkonen-Condit en español—, y de análisis de concordancias paralelas para comprobar qué grado de influencia tiene el texto original en las decisiones lingüísticas del traductor. A través de un trabajo previo de investigación en gramáticas y trabajos contrastivos y especializados, hemos intentado identificar elementos que fueran típicos del español y únicos en el sentido de que no contaran, en principio, con elementos equivalentes en inglés fácilmente accesibles, bien porque no existiera en esta lengua un elemento en particular que pudiera servir de traducción y, por tanto, el significado de nuestro *unique item* se trasladase mediante una variedad de estructuras lingüísticas, o bien, si este elemento existiese, porque fuera formalmente tan diferente que lo más probable es que sugiriera a los traductores elementos alternativos en español más cercanos a la estructura inglesa. Los elementos seleccionados fueron un grupo de 27 perífrasis verbales de infinitivo y gerundio divididas en tres grupos según las posibilidades que tenían, desde nuestro punto de vista, de encontrarse representadas en mayor o menor medida en los textos traducidos (capítulo 4), el conjunto de los pronombres, determinantes y adverbios locativos demostrativos españoles (capítulo 5) y un grupo de sufijos apreciativos de valor diminutivo, aumentativo y peyorativo (capítulo 6)²²⁷.

Los resultados del análisis fueron los siguientes:

- **Perífrasis verbales:** Tanto en la primera como en la segunda versión del corpus analizado, el conjunto de los tres grupos de perífrasis seleccionadas resultó infrarrepresentado en traducciones, un hecho que no habíamos anticipado. La rotundidad de esta infrarrepresentación fue menor en la segunda versión, puesto que ni el primer ni el tercer grupo mostraron diferencias significativas entre subcorpus, lo que demuestra, por un lado, la necesidad de contar con corpus

²²⁷ En principio, 27, pero, como comentamos en dicho capítulo, el número final depende de si consideramos los elementos analizados conjuntamente sufijos diferentes o simples variantes. Tras agrupar algunas formas durante el análisis, el número de filas en la tabla de resultados quedó en 23.

comparables de tamaño similar, para que las normalizaciones de frecuencias alteren los resultados reales en la menor medida posible, y por otro lado, el efecto que pueden tener los textos individuales sobre el comportamiento de determinados elementos (nos referimos a las perífrasis que pasaron de estar infrarrepresentadas a sobrerrepresentadas a causa de su mayor utilización por parte de un autor o traductor determinado), lo que también sugiere la necesidad de contar con corpus bastante más extensos que permitan diluir la influencia de elementos idiosincráticos.

En resumen, podemos decir que la hipótesis de partida se cumplió, pues el grupo II, conformado por perífrasis sin equivalentes directos en la lengua de partida, fue el único que mostró una infrarrepresentación significativa en ambas fases del corpus. No obstante, si hilamos más fino, cuatro integrantes de este grupo no confirmaron dicha hipótesis, mientras que sí lo hicieron algunos de los integrantes de los otros dos grupos (sorprendentemente, la mayor parte de estos casos provienen del grupo I, con equivalentes supuestamente directos). El total de perífrasis infrarrepresentadas fue once de veintisiete: cinco de nueve en los grupos I y II y una en el grupo III. Como se desprende de estas cifras, nuestras previsiones con respecto a la infrarrepresentación de las perífrasis causada por la falta de equivalentes en el original no resultaron confirmadas en todas las ocasiones, lo que significa que en el comportamiento de estos elementos hay otros factores en juego además de la influencia del texto o la lengua de partida.

La infrarrepresentación en el grupo de perífrasis con equivalentes en principio inmediatos venía dada también, como vimos, por la falta de enlaces claros entre ellas y un único elemento de la lengua de partida —debida a la existencia de alternativas de función similar en español o en inglés para expresar los mismos matices—. La sobrerrepresentación de perífrasis sin equivalente claro, por otra parte, podía surgir del hecho de que la perífrasis en cuestión fuera el recurso más normal en español para expresar determinados significados. Este fenómeno también se producía, por supuesto, si el elemento equivalente en inglés sí estaba claro o era muy prominente, o si existía una gran variedad de estructuras inglesas que pudieran dar lugar a la utilización de la perífrasis, o, en alguna ocasión, también debido a la poca frecuencia de uso en español original del elemento, que magnificaba el impacto de los pocos ejemplos hallados en

traducciones. El último factor detectado fue, por supuesto, la preferencia individual del traductor por un recurso lingüístico determinado.

- **Demostrativos:** En el caso de los adverbios locativos, nuestro estudio arroja resultados similares a los obtenidos por Hareide en su trabajo sobre el gerundio, dado que existe una sobrerrepresentación general de estos elementos. Por el contrario, los pronombres y determinantes como grupo se encuentran infrarrepresentados y también cada forma básica por separado. A pesar de que los datos confirman una infrarrepresentación general de estos elementos, no podemos afirmar que nuestra hipótesis de partida se haya cumplido, puesto que lo que esperábamos encontrar era la infrarrepresentación de solo uno de los dos grados de lejanía del español. Sin embargo, sí apreciamos una mayor diferencia de frecuencias y un menor uso proporcional con respecto a las otras dos formas básicas en el subcorpus de traducciones en el caso de *aquel*, además de un uso más idiosincrático de este demostrativo por parte de los traductores, lo que acerca el tercer grado de lejanía en mayor medida que los otros dos al concepto de elemento único que buscábamos.
- **Sufijos apreciativos:** la infrarrepresentación se comprobó en diez de los veintitrés sufijos o agrupaciones de sufijos que finalmente realizamos. No obstante, ningún elemento mostró frecuencias significativamente más elevadas a favor del subcorpus de textos traducidos, y los elementos infrarrepresentados (sobre todo diminutivos y aumentativos) eran los que presentaban frecuencias de uso más elevadas, por lo que, en conjunto, el grupo de sufijos se encuentra claramente infrarrepresentado en traducciones. En inglés no existen apenas estructuras formalmente equivalentes a los sufijos apreciativos analizados: el uso de estos elementos es el que más parece depender de la voluntad del traductor.

7.3. Definición de *unique item*

Como comentábamos en la introducción teórica, Hareide proporciona una definición revisada de *unique item*, puesto que en su estudio no logra confirmar la

hipótesis de Tirkkonen-Condit a pesar de contar con un elemento sin equivalente directo en la lengua original:

A unique item is a target language feature that has no direct counterpart in the source language, and has no salient or prototypical source language structure that may impact its choice in the target language. It is characterized by being infrequent and non-prototypical. (2013: 276)

Nos proponíamos, por tanto, comprobar hasta qué punto se cumplen estas condiciones en los elementos que nosotros hemos encontrado infrarrepresentados en nuestro corpus. Por supuesto, este apartado se ha realizado de modo muy tentativo, pues, como ya hemos explicado anteriormente, en lo referente a la presunción de prototipicidad no hemos hecho estudios específicos para determinarla en ninguno de los dos idiomas, por lo que solo podemos basarnos en lo que hemos encontrado en las gramáticas y diversos estudios consultados y en las frecuencias de uso observadas en nuestro corpus. A continuación, intentaremos dilucidar, como sugería la autora, si dichos elementos deben presentar una frecuencia baja, no ser prototípicos en español, no poseer una correspondencia prominente en la lengua de partida ni tampoco conexión con ningún concepto relacionado en el léxico bilingüe del traductor.

También resumiremos lo que hemos encontrado con respecto a la influencia de los tres factores que, según la Gravitational Pull Hypothesis, pueden dar lugar a una infrarrepresentación o a una sobrerrepresentación de elementos en la lengua meta. Recordemos que estos eran: a) la frecuencia o prototipicidad del elemento en la lengua meta (cuanto mayores sean estas, más probabilidades existirán de que el elemento se encuentre sobrerrepresentado); b) la prominencia o prototipicidad de alguna estructura de la red conceptual de la lengua de partida que pueda ejercer también fuerza a favor de su utilización en la lengua meta; y c) los patrones de conexión entre los conceptos de ambas lenguas en el léxico mental del traductor (cuanto más débil sea ese enlace, más probabilidades hay de que el elemento esté infrarrepresentado).

➤ **Perífrasis verbales:**

Basándonos en el estudio de la fase II, si una perífrasis se encuentra infrarrepresentada, puede ser porque no es un elemento prototípico de la lengua meta o no se utiliza con frecuencia, ya que hay otros equivalentes de significado y funciones

similares (factor 1), porque no cuenta con elementos equivalentes lo suficientemente prominentes en la red conceptual de la lengua de partida que puedan ejercer su influencia en la selección del elemento (factor 2) o porque no existen enlaces directos o claros entre la perífrasis y los conceptos relacionados de la lengua de partida (factor 3). Los tres factores no se dan a la vez en todas las ocasiones, por lo que deducimos que en algunos casos uno o dos contrarrestan a los que no se cumplen. Hay que recordar que las perífrasis verbales son elementos muy diversos en cuanto a funciones y significado, por lo que resulta difícil extraer conclusiones sobre el grupo en su conjunto.

Tampoco podemos decir que las características contenidas en la definición de Hareide se cumplan para todas las perífrasis, dado que, como hemos visto en las conclusiones del capítulo correspondiente, algunas son muy frecuentes en español, otras tienen equivalentes directos en la lengua de partida y algunos de estos equivalentes son también muy frecuentes en inglés. Con respecto al nivel de prototipicidad en la lengua, no podemos más que aventurar conclusiones basadas en lo que hemos leído sobre los distintos elementos en estudios anteriores y los resultados de los análisis en este corpus, pero creemos que en ocasiones las formas estudiadas, tanto elementos españoles como determinados equivalentes ingleses, son prototípicas. Por lo tanto, lo que hemos hallado es una confluencia de factores que explica la infrarrepresentación o la sobrerrepresentación en cada caso concreto, como se puede observar en la tabla 58 del capítulo 4. Un factor muy importante no contemplado en la definición de Hareide que puede influir en el escaso uso de una forma en traducciones (aunque esta sea frecuente o prototípica en español) es la existencia de alternativas de traducción en lengua meta que puedan verse favorecidas por alguna razón, como motivos estilísticos o preferencias personales del traductor. No obstante, podría argumentarse que estas posibles razones siempre vendrán determinadas en última instancia por el factor 3, que relaciona los conceptos en la mente del traductor.

➔ **Demostrativos:**

En el caso de los demostrativos, recordemos que el factor de la frecuencia de uso o prototipicidad del elemento en lengua meta no ejercía suficiente fuerza para garantizar una representación equivalente a la de los originales en traducciones en el caso de los determinantes y pronombres, que se eliminaban en mayor medida que los adverbios. A pesar de ser elementos frecuentes y prototípicos dentro de la función deíctica y fórica que ejercen en español, existen alternativas de expresión que los hacen menos

necesarios que en inglés. El segundo factor, la prominencia de los elementos en lengua de partida, era claramente el más fuerte en el caso de los adverbios, que se encontraban sobrerrepresentados en traducciones, muy probablemente por la influencia de los textos originales, pero tampoco lograba contrarrestar el impulso de normalización de los traductores en el caso de pronombres y determinantes, finalmente utilizados en menor medida que en español original. Con respecto al tercer factor, la conexión de los elementos en la estructura conceptual del traductor, es evidente que esta existe, pero parece ser más pronunciada en el caso de los adverbios que en el de los otros dos elementos demostrativos, como muestran los porcentajes de traducción de sus equivalentes ingleses por dichos elementos. En el caso de *aquel*, esta conexión es más débil que las demás, puesto que se prefiere —con una diferencia de uso aún mayor que en originales españoles— la traducción de *that/those* por el otro grado de lejanía, *ese/a(s)*.

Con respecto a las características que menciona Hareide en la definición de *unique item*, observamos que tampoco se cumplen en este caso en su totalidad, pues la falta de equivalente directo se da solo en el caso del segundo y tercer grado de lejanía y, sin embargo, los tres grados resultaron infrarrepresentados. Por otro lado, sí existe en inglés una estructura, muy frecuente según las cifras obtenidas en el análisis (y prototípica por fuerza, pues hablamos de un sistema cerrado): el sistema de los demostrativos, que debería influir en la elección de una u otra forma en español. Lo que vemos es que la influencia que ejerce es muy diferente según si hablamos de adverbios o de pronombres y demostrativos. Por último, no podemos considerar que los demostrativos en español sean elementos poco frecuentes o no prototípicos, excepto en el caso de *aquel* si lo comparamos con los otros grados del sistema. No obstante, este factor tampoco parece definitorio, pues ya hemos dicho que los otros grados también resultaron infrarrepresentados.

➤ Sufijos:

En este último estudio hallamos una infrarrepresentación general del grupo de sufijos estudiados, y también de cada uno de los tres grupos, aunque individualmente no suceda lo mismo en todas las formas (especialmente en las peyorativas). El factor que más ha influido en los resultados es indudablemente el segundo, dado que existe una clara falta de elementos equivalentes en los originales. A ello se une el hecho de que en español existen numerosas otras maneras de expresar los valores apreciativos

contenidos en los sufijos estudiados, aunque recordemos que no podíamos valorar con exactitud la fuerza ejercida por el primer factor (la frecuencia o prototipicidad en la lengua meta). El tercer factor, relacionado con la conexión de las estructuras de ambas lenguas en la mente del traductor, también favorecía la infrarrepresentación de nuestros sufijos, pues los valores expresados por estos elementos pueden manifestarse de muy diversas maneras tanto en inglés como en español.

En cuanto a las condiciones de la definición de *unique item* de Hareide, recordemos, en primer lugar, que, efectivamente, consideramos que los sufijos no cuentan con elementos equivalentes directos en la lengua de partida. En segundo lugar, no podemos afirmar que no existan estructuras prominentes en la lengua de partida que puedan influir en su utilización (aunque no sean elementos formalmente parecidos a nuestros sufijos), aunque lo que sí hemos visto es que la variedad de estructuras que ejercen funciones equivalentes es muy grande. En cuanto a la tercera condición mencionada en la definición, no sabemos hasta qué punto los sufijos estudiados son o no prototípicos, pero sí hemos comprobado que las formas más frecuentes y prototípicas dentro del grupo de los diminutivos y aumentativos son las que más posibilidades tienen de estar infrarrepresentadas, al contrario de lo que postula la definición.

Por lo tanto, la primera condición («no direct counterpart in the source language») solo se cumple sin peros en el caso de los sufijos y solo en parte en los otros dos; la segunda («no salient or prototypical source language structure that may impact its choice in the target language»), parcialmente en el caso de las perífrasis y quizás la de los sufijos, pero no el de los demostrativos, y la tercera («being infrequent and non-prototypical»), también parcialmente en el caso de las perífrasis, pero no en la de los demostrativos —excepto si hablamos de la menor frecuencia de uso de *aquel* con respecto a los otros dos grados—, ni en la de los sufijos, salvo si los tomamos en conjunto. Como vemos, es muy difícil extraer conclusiones claras sobre este tema trabajando sobre grupos de elementos que han resultado ser tan poco uniformes como los nuestros. Nuestro objetivo principal era acumular evidencia procedente de varios indicadores para verificar o refutar la Unique Items Hypothesis, pero eso implica que no podemos controlar la influencia de los factores contemplados por la Gravitational Pull Hypothesis —que incidirán de modo diferente en cada uno de estos ítems— de una manera tan precisa como si analizáramos un único elemento.

Por otro lado, es evidente que, como describíamos en la tabla 58 de los factores que pueden llevar a la infrarrepresentación de las perífrasis, la infrarrepresentación de un elemento en traducciones es un fenómeno complejo, y es difícil establecer *a priori* qué elementos pueden constituir un *unique item*, pues existe una amplia variedad de factores que pueden influir en su utilización; entre ellos, la existencia de otras formas de expresión que pueden verse favorecidas en determinadas circunstancias (independientemente de que nuestro elemento sea o no frecuente o prototípico), ya sea por su similitud a las estructuras originales o incluso por preferencias personales de los traductores.

Puesto que la identificación *a priori* es una operación compleja que no siempre proporciona los resultados deseados, hemos intentado también utilizar la técnica propuesta por Chesterman para identificar elementos únicos *a posteriori* (2007: 11-12) que explicamos en el capítulo de las perífrasis verbales. Según esta, los elementos únicos identificados en cada uno de los estudios serían los siguientes:

- ✦ **Perífrasis verbales:** «ir + gerundio», «haber que + infinitivo» y «comenzar a + infinitivo».
- ✦ **Demostrativos:** las tres formas básicas de pronombres y determinantes, aunque como ya hemos visto, la que más se acerca a nuestro concepto de elemento único es *aquel*.
- ✦ **Sufijos:** los diminutivos *-ete/a*, *-illo/a* e *-ito/a* y los aumentativos *-ón/a* y *-ote/a*.

Sin embargo, creemos que un umbral de significación establecido de manera arbitraria para identificar posibles elementos únicos debe servir únicamente de guía y nunca impedir el análisis de otros elementos que puedan resultar interesantes al investigador pero no cumplan los criterios de significación estadística, como afirman Römer y Wulff (2010: 119): «While statistical significance tests are invaluable tools to quantify strong associations in the data, we would miss out on a number of interesting tendencies by discarding the results as irrelevant on the basis of the failure of the data to meet an arbitrary significance threshold».

7.4. Confluencia de fenómenos

La infrarrepresentación no es el único fenómeno que hemos observado en los textos de nuestro corpus. Además de la **sobrerrepresentación** de los adverbios demostrativos, hemos mencionado, a lo largo de este trabajo, distintos casos de normalización, explicitación y simplificación con los que nos hemos topado en el análisis. Por ejemplo, en el estudio de las perífrasis encontrábamos indicios de una tendencia hacia la **convergencia** en traducciones, dado que en ellas existía una mayor homogeneidad en la elección de determinadas perífrasis incoativas que en los originales españoles. La **explicitación** es un fenómeno que se da en los tres estudios, cuando el traductor añade alguno de los elementos estudiados con el objetivo de hacer más comprensibles ciertas relaciones sintácticas o semánticas. Observábamos en los demostrativos una menor variación de estructuras en traducciones (en las que se utilizaban en menor medida demostrativos pospuestos al sustantivo y antepuestos a oraciones de relativo), hecho que relacionamos con una posible dinámica de **normalización** del español traducido mediante el uso de las opciones más estandarizadas. Por último, también hallábamos dos posibles casos de **simplificación** en la variedad menor de verbos con los que se combina la perífrasis «echar a + infinitivo» en originales y traducciones y en la mayor variedad de bases sufijadas que dan lugar a diminutivos en el subcorpus de originales.

Todos estos fenómenos influyen seguramente en la utilización de nuestros elementos únicos. Por lo tanto, lo más probable es que nuestros resultados sean una combinación de causas de todo tipo y no puedan reducirse a los tres factores de la hipótesis de Halverson —como ella misma apunta, por otra parte, refiriéndose a la probable influencia de factores sociales y culturales además de cognitivos (2003: 233)—, ni tampoco podrán cumplirse siempre las tres condiciones de la definición de *unique item* de Hareide, pues en muchas ocasiones habrá otros fenómenos afectando su comportamiento.

Los fenómenos lingüísticos son muy complejos y es difícil establecer dónde termina uno y empieza el siguiente. Multitud de factores pueden ser responsables simultáneamente de un mismo resultado y es por este motivo por el que en este análisis hemos encontrado un entretrejimiento de infrarrepresentaciones, sobrerrepresentaciones, normalización, interferencia del texto original y convergencia entre textos traducidos.

Solo podemos extraer tendencias, y no es fácil extraer estas tendencias, como bien han señalado quienes han abordado el tema de los posibles universales a lo largo de los años, a partir de estudios de alcance limitado como el nuestro. Tengamos en cuenta, por ejemplo, el hecho de que, al variar los textos del corpus en el estudio de las perífrasis, sin siquiera modificar cuestiones como el género narrativo, obtenemos resultados contradictorios.

Cuando intentamos averiguar a qué se debe este fenómeno, son muchos los sospechosos: las preferencias lingüísticas de autores o traductores específicos, la selección de los elementos analizados y excluidos del estudio, quizás simplemente las fórmulas estadísticas, que se ven afectadas en mayor o menor medida por el número de palabras reales de cada subcorpus, o incluso la norma lingüística predominante a uno u otro lado del Océano Atlántico o de regiones determinadas de España. Una reflexión al respecto es que, a pesar de no haber establecido este criterio de comparabilidad específicamente, prácticamente todas las obras traducidas que incluimos en el corpus, tanto de la fase anterior como de la presente, estaban traducidas por profesionales de origen catalán, dado que gran parte de la industria editorial española se encuentra establecida en Barcelona. La variedad de español que hemos estudiado, por consiguiente, es la que se utiliza allí. ¿Obtendríamos los mismos resultados si intentáramos buscar traductores madrileños? Está claro que hay que realizar muchos más estudios para extraer conclusiones al respecto.

7.5. Recapitulación

El objetivo principal de este estudio era averiguar si se producía una infrarrepresentación de los elementos seleccionados en traducciones al español. Esto ha sido así en los tres casos analizados, siempre que tomemos todos los elementos de cada caso en conjunto. Las frecuencias recogidas indican que los escritores de obras escritas originalmente en español utilizan ciertas expresiones típicas de su lengua materna de manera considerablemente más frecuente que los traductores. Estos últimos parecen responder a la falta de expresiones equivalentes en la lengua origen, ya que no utilizan estas expresiones en la misma medida ni de la misma manera en que las utilizarían al escribir un texto en español sin la influencia de un texto de partida.

Si hilamos más fino y nos referimos a los integrantes de estos grupos por separado, no siempre hemos obtenido los resultados pronosticados. En el caso de las perífrasis, esperábamos estos resultados parciales porque habíamos decidido incluir un grupo de elementos «de control». No obstante, nos sorprendimos al comprobar que incluso entre estos elementos hallamos frecuencias menores en textos traducidos. En el de los demostrativos, lo sorprendente fue encontrarnos con todas las formas de pronombres y determinantes infrarrepresentadas y casi todas las formas de los adverbios sobrerrepresentadas, cuando lo que habíamos anticipado era que solo una de las formas de lejanía presentara frecuencias atípicas en traducciones. Los sufijos, por último, fueron los que más se acercaron a los resultados esperados, aunque muchas formas no se encontraran infrarrepresentadas. En este último caso creemos que es en el que más puede haber afectado los resultados el tamaño reducido del corpus, pues para muchos apenas contábamos con concordancias.

Conscientes de que el estudio que hemos realizado es de carácter empírico descriptivo y sus características nos impiden proponernos hallar causas a estos fenómenos, hemos intentado, de todos modos, identificar posibles factores a lo largo del análisis que pudieran correlacionarse con su aparición para intentar sugerir alguna explicación de naturaleza muy tentativa. Además, hemos intentado enlazar los fenómenos observados con los factores, identificados por Halverson, que pueden favorecer —o todo lo contrario— la infrarrepresentación de elementos únicos en traducciones. También hemos intentado comprobar si las condiciones propuestas por Hareide para la definición de un *unique item* se verificaban en nuestros estudios y esto no ha sido así en todos los casos. No obstante, recordemos que solo hemos podido extraer conclusiones parciales, pues para hablar de la prototipicidad de un determinado elemento en una lengua o de su prominencia en las redes conceptuales de los traductores es necesario realizar estudios específicos que no se encontraban dentro de las posibilidades de este trabajo. Además, como hemos mencionado, la variedad y falta de uniformidad de los elementos analizados no nos permiten realizar afirmaciones concluyentes sobre la incidencia de los distintos factores sobre dichos elementos. Por lo tanto, no podemos descartar que en estudios futuros de mayor alcance o con un objeto de estudio más específico puedan confirmarse las intuiciones de esta autora.

Lo que hemos comprobado es que la falta de estructuras activadoras en el original no es el único factor que puede influir en la infrarrepresentación de un elemento dado, como postula la Unique Items Hypothesis: también participan otros factores que

tienen que ver con la manera en la que los traductores utilizan la lengua meta. De lo contrario, no se explicarían casos como la infrarrepresentación en traducciones del demostrativo *este* o de ciertas perífrasis con equivalentes evidentes en inglés.

Por último, a través de los tres estudios realizados, hemos compilado un repertorio de equivalentes de traducción de los elementos analizados que puede resultar útil como fuente de inspiración para traductores o aportar información a estudios de lingüística contrastiva. Como apunta Johansson (2007: 6) acerca de los distintos estudios contrastivos que menciona en su obra: «The most striking finding is the extent to which words may differ in use cross-linguistically, although at the outset one might have regarded them as equivalent». En nuestros estudios también hemos comprobado que muchas veces los elementos que considerábamos los equivalentes más evidentes para determinadas estructuras inglesas no se utilizan en la medida esperada en las traducciones, dando lugar a una amplia variedad de equivalentes alternativos con los que no contábamos en un principio.

El interés principal de este trabajo, no obstante, radica en el hecho de que, aunque los datos no han apoyado la hipótesis en el caso de todos los elementos seleccionados para su análisis, hemos comprobado, a través de tres estudios empíricos claros y replicables, y en los que hemos tenido en cuenta algunas de las mejoras metodológicas propuestas por distintos autores a lo largo de los años —como la combinación de corpus comparables y paralelos y la utilización de tests estadísticos—, que existe un fenómeno de infrarrepresentación en textos traducidos al español de determinados elementos de esta lengua sin equivalentes formalmente similares en inglés. Por consiguiente, la Unique Items Hypothesis es un fenómeno presente no solo en traducciones del inglés al sueco, al finés, al chino y al francés, sino también en nuestra combinación lingüística. La replicación de estudios llevados a cabo en torno a esta hipótesis en condiciones diferentes (dado que los elementos analizados son distintos, al igual que el par de lenguas de trabajo y las características del corpus utilizado) es un paso más hacia la validación —o no— del estatus de la infrarrepresentación como tendencia general de traducción. Asimismo, mediante la descripción detallada del corpus utilizado y una explicación minuciosa de todos los pasos llevados a cabo en nuestro estudio, facilitamos su replicabilidad y contribuimos al progreso del conocimiento en el campo de los estudios basados en corpus.

Es un paso más, decimos, y solo uno, pues es evidente que nuestras conclusiones se aplican únicamente a la combinación lingüística inglés estadounidense-español peninsular y al género de la novela policiaca. Por otra parte, nuestro corpus no llega al millón y medio de palabras entre los distintos componentes. Estas limitaciones no nos permiten realizar generalizaciones a un nivel superior al mencionado. Hemos caracterizado el lenguaje utilizado en las traducciones al español de este género literario en concreto con respecto a los parámetros estudiados, pero para sostener que estas características constituyen tendencias generales, será necesario seguir acumulando resultados de numerosos estudios en otras combinaciones lingüísticas, otros períodos de tiempo y con elementos de diferentes categorías lingüísticas.

No obstante, los resultados de todos los trabajos como este son importantes, pues constituyen un aporte más a las pruebas recogidas hasta el momento que defienden la idea de que existen tendencias generales que se cumplen independientemente de los factores externos que afecten al proceso traductor pues pertenecen al proceso en sí mismo; ciertas características que se encuentran en los textos traducidos y parecen originarse en los procesos mentales específicos de la actividad traductora, que no hay que juzgar como errores o aciertos del traductor, sino como realidades que vale la pena estudiar en profundidad pues pueden ayudar a ampliar nuestra comprensión del proceso y los productos de esta actividad. En palabras de Eskola (2004: 86): «Studying translation universals is like trying to solve a jigsaw puzzle. Every piece of information about the use of any single pattern is part of the whole when we try to find out what translations are really like».

7.6. Futuras líneas de investigación

La diferencia de frecuencias observadas puede, por supuesto, deberse a muchos factores de naturaleza cognitiva, sociocultural y situacional (Chesterman 2003: 20), además de lingüística, que no podemos tener en cuenta en un estudio como este. Entre esos factores se encuentran los sociales, los temporales, los relacionados con el encargo, con el género textual, con la naturaleza de los propios textos —en el sentido de que quizás la misma situación planteada y narrada en algunas novelas hace que ciertas expresiones no sean necesarias y por eso aparezcan menos— o con las capacidades,

conocimientos y actitudes del propio traductor. Chesterman (2004: 45), por ejemplo, menciona entre las posibles causas de algunas características de la traducción las siguientes, relacionadas con las condiciones socioculturales específicas de la situación de traducción y el rol del traductor como mediador:

Other kinds of explanations may be sought in the nature of translation as a communicative act, and in translators' awareness of their socio-cultural role as mediators of messages for new readers (see e.g. Klaudy 1996). Translators tend to want to reduce entropy, to increase orderliness. They tend to want to write clearly, insofar as the skopos allows, because they can easily see their role metaphorically as shedding light on an original text that is obscure – usually unreadable in fact – to their target readers: hence the need for a translation. Their conception of their role may give a prominent position to the future readers of their texts; this may have been emphasized in their training, for instance. It is this conception of their mediating role that may offer some explanation for the tendency towards explicitation, towards simplification, and towards reducing what is thought to be unnecessary repetition – to save the readers' processing effort.

Todos los factores que acabamos de mencionar no pueden descartarse a menos que el estudio incluya la totalidad de las posibles variables o que se realicen sucesivos estudios que puedan ir controlando el efecto de cada una por separado.

Manteniéndonos dentro de los límites y las posibilidades de un trabajo como el nuestro, creemos que sería factible aumentar su alcance, en primer lugar, ampliando nuestro corpus actual con mayor cantidad de textos traducidos y originales en español, y añadiendo textos de diferentes géneros y épocas, para reducir la influencia de las características idiosincráticas de autores, traductores, géneros o textos específicos (o aislar la influencia que estos pueden ejercer sobre los resultados), y, en segundo lugar, añadiendo otro componente al corpus para hacerlo bidireccional: el de las traducciones de nuestras novelas españolas al inglés, si en un futuro estuvieran disponibles o si cambiáramos la composición del subcorpus de originales españoles. Esto permitiría comprobar si un mismo fenómeno se produce en ambas direcciones de traducción, lo que reforzaría la hipótesis de su naturaleza universal. Otra adición que podría hacerse al corpus es la inclusión de originales en otras lenguas, como el francés. Como defiende Baker, utilizar distintos idiomas de partida ayudaría a determinar el impacto de la lengua de origen en el español traducido y a verificar si las características que hemos

encontrado se manifiestan independientemente del par de lenguas analizado. El corpus utilizado en este estudio puede complementarse con otros ya compilados, como el corpus COVALT que hemos mencionado varias veces a lo largo de este trabajo, que incluye dos lenguas de llegada (valenciano y español) y tres de partida (inglés, francés y alemán). Por último, el análisis de diferentes candidatos en el mismo conjunto de datos, como sugiere Corpas Pastor (2008: 126), es algo que aportaría bastante perspectiva a estudios como el nuestro, dado que pasaríamos de centrarnos en un único candidato a universal a estudiar de manera sistemática los distintos fenómenos que pueden darse en la traducción y la relación que existe entre ellos.

En un nivel más práctico, sería interesante estudiar la manera más eficiente de corregir el etiquetado del corpus para poder realizar búsquedas más precisas, y no sólo de elementos léxicos sino también de categorías y estructuras sintácticas. Cabe la posibilidad de centrarnos en otros elementos típicos del español para averiguar si la hipótesis se verifica al cambiar una vez más de objeto de estudio. Algunos ejemplos en los que hemos pensado son marcadores conversacionales como *bueno* o *ya*, interjecciones y exclamaciones («¡anda ya!»), fraseología típica de ámbitos específicos del español, como el toreo, los superlativos *-ísimo* y *-érrimo*, como mencionamos en el capítulo de los sufijos, o elementos de naturaleza más gramatical, como el gerundio (estudiado por Hareide 2013), el pretérito imperfecto, las pasivas impersonales con *se* o el orden habitual de determinados elementos en la oración.

Es evidente que los estudios sobre la naturaleza de la traducción no pueden limitarse al análisis de los datos cuantitativos obtenidos mediante herramientas automáticas, sino que, a partir de dichos datos, debe intentarse profundizar en las causas que puedan estar detrás de la aparición de los fenómenos estudiados. Una posible vía investigadora sería el intentar encontrar explicaciones a dichos fenómenos, para acercarnos al modelo causal, del cual Chesterman (2003) considera que debe partir la construcción de una teoría de la traducción. En este tipo de modelo son fundamentales las causas cognitivas, es decir, las decisiones subjetivas tomadas consciente e inconscientemente por el traductor, por lo que este último objetivo, en el caso de que su consecución fuera factible, requeriría complementar los estudios textuales basados en corpus con métodos e instrumentos centrados en el proceso de traducción, no en el producto, y, por tanto, de naturaleza muy distinta a los utilizados en este trabajo (p. ej. *think-aloud protocols*, *dialogue protocols*, *keylogging*; cf. Tirkkonen Condit 2002a) o

también con estudios de base cognitiva algo más sencillos, similares a lo que sugieren o han realizado diversos autores.

Podemos mencionar el de Kujamäki (2004) sobre elementos relacionados con la nieve y la conducción en finés, que ya explicamos en la introducción teórica, o los sugeridos por Chesterman (2004: 42), que parte de la base de que, si las tendencias universales se manifiestan como rasgos característicos de los textos traducidos, entonces sería posible utilizar estos rasgos —como la presencia de una frecuencia inusual de ciertos elementos—, para determinar si un texto es original o traducción. Por lo tanto, se podría realizar un análisis de una característica de naturaleza potencialmente universal en un original y una traducción, y comprobar si dicho análisis permitiría predecir qué texto es el original y cuál la traducción, o seleccionar algunos textos que presenten marcas de la característica estudiada y otros que no las contengan, y comprobar cuáles consideran los sujetos como traducciones (véase también Toury 1995: 230 y Tirkkonen-Condit 2002b).

Estos datos proporcionarían información acerca de los factores que pueden influir en el proceso de traducción y originar tendencias generales o universales de traducción, ya que, como afirma Mauranen (2007), los universales a los que nos referimos en traducción no son necesariamente absolutos sino que se ven afectados por factores muy diversos, de naturaleza cultural, histórica, sociológica o cognitiva, entre otras. Por lo tanto, es necesario considerar todos estos tipos de variables a la hora de formular posibles tendencias generales de traducción (o universales, o como convengamos en llamarlas). A fin de cuentas, lo verdaderamente importante no es el hecho mismo de encontrar universales de traducción, sino comprobar, como hemos dicho, que existen tendencias que suelen darse en el proceso traductor, que se originan en procesos mentales específicos de esta actividad pero se ven influidas por multitud de factores externos, y cuyo estudio puede, por lo tanto, ayudarnos a mejorar nuestra comprensión del proceso de traducción y de sus productos e incluso, como sugiere Chesterman (2003: 27), a mejorar la formación de los traductores y la calidad de las traducciones, una vez hayamos identificado cuáles son las condiciones que propician, en sus palabras, «more desired effects and fewer unwanted ones».

8. Bibliografía

8.1. Fuentes primarias

- Cornwell, P. (1994). *Cruel and Unusual*. London: Warner Books (Primera edición publicada en 1993 por Little, Brown and Company).
- Cornwell, P. (1994). *Cruel y extraño* (J. Mustieles, trad.). Barcelona: Ediciones B.
- Ellroy, J. (1991). *Los Ángeles confidencial* (C. Gardini, trad.). Barcelona: Ediciones B (Obra original: *L.A. Confidential*, publicada en 1990 por The Mysterious Press).
- Ellroy, J. (1993). *Jazz Blanco* (H. Sabaté Vargas, trad.). Barcelona: Ediciones B.
- Ellroy, J. (2001). *White Jazz*. New York: Vintage Books (Primera edición publicada en 1992 por Alfred A. Knopf).
- Giménez Bartlett, A. (1996). *Ritos de muerte*. Barcelona: Grijalbo.
- Grafton, S. (1995). *L is for Lawless*. New York: Henry Holt and Company.
- Grafton, S. (1996). *L de ley (o fuera de ella)* (A. P. Moya, trad.). Barcelona: Tusquets.
- Highsmith, P. (1993). *Ripley en peligro* (I. Núñez, trad.). Barcelona: Anagrama (Primera edición publicada en 1992 por Cículo de Lectores).
- Highsmith, P. (2008). *Ripley Under Water*. New York: W. W. Norton & Company (Primera edición publicada en 1991 por Bloomsbury Publishing).
- Hunter, E. (1995). *Conversación criminal* (C. Milla Soler, trad.). Barcelona: Plaza y Janés Editores (Obra original: *Criminal Conversation*, publicada en 1994 por Warner Books).
- Leonard, E. (1994). *Rum Punch* (E. de Hériz Ramón, trad.). Barcelona: Ediciones B.
- Leonard, E. (2004). *Rum Punch*. London: Orion Books (Primera edición publicada en 1992 por Viking).
- Madrid, J. (1995). *Cuentas pendientes*. Madrid: Alfaguara.
- Martín, A. (1992). *El hombre de la navaja*. Barcelona: Plaza y Janés Editores.
- Martínez Reverte, J. (1995). *Gálvez y el cambio del cambio*. Barcelona: Anagrama.
- Mosley, W. (1994). *El demonio vestido de azul* (R. Corgatelli, trad.). Barcelona: Anagrama (Obra original: *Devil in a Blue Dress*, publicada en 1990 por Norton & Company).
- Núñez, S. (1994). *La danza de los reptiles*. Gipuzkoa: Txertoa.
- Rendell, R. (1993). *Un beso para mi asesino* (C. Camps, trad.). Barcelona: Grijalbo Mondadori (Obra original: *Kissing the Gunner's Daughter*, publicada en 1992 por Hutchinson).
- Vázquez Montalbán, M. (1996). *El premio*. Barcelona: Planeta.

8.2. Diccionarios

- Collins English Dictionary - Complete & Unabridged 10th Edition* (2005). [en línea] HarperCollins Publishers. Disponible en Lexico Publishing, LLC, *Dictionary.com*, <http://dictionary.reference.com/>
- Collins Spanish Dictionary 8th Edition* (2005). [en línea] HarperCollins Publishers. Disponible en Lexico Publishing, LLC, *Dictionary.com*, <http://spanish.dictionary.com>
- De Gámez, T. (ed.) (1973). *Simon and Schuster's International Dictionary*. Simon & Schuster. New York: Prentice Hall.
- Dictionary.com Unabridged*. [en línea]. Based on the *Random House Dictionary*, Random House. Disponible en <http://dictionary.reference.com/>
- Diccionario Espasa Concise* (2000). [en línea] Espasa Calpe. Disponible en *WordReference.com* <http://www.wordreference.com/es/>
- María Moliner (2001) *Diccionario de uso del español de María Moliner*, 2.^a edición [versión electrónica]. Madrid: Gredos.
- Merriam-Webster's Spanish-English Dictionary* (1998). [en línea] Merriam-Webster. Disponible en <http://www.merriam-webster.com/spanish>
- Oxford Superlex* (1996). [version electrónica] Oxford University Press-Hachette Livre.
- Real Academia Española (2005). *Diccionario panhispánico de dudas (DPD)* [versión en línea]. Disponible en <http://www.rae.es/recursos/diccionarios/dpd>
- Real Academia Española (2001). *Diccionario de la lengua española*, 22.^a edición [versión en línea]. Disponible en <http://lema.rae.es/drae>
- Urban Dictionary* (1999-2015). [en línea] Disponible en <http://www.urbandictionary.com>
- WordReference English-Spanish Dictionary* (2015). [en línea] Disponible en <http://www.wordreference.com/es/translation.asp>

8.3. Bibliografía científica

- Aarts, J. (2002). Does corpus linguistics exist? Some old and new issues. En L. E. Breivik y A. Hasselgren (eds.), *From the Colt's Mouth... and Others': Language Corpora Studies in Honour of Anna-Brita Stenström*, (pp. 1-17). Amsterdam: Rodopi.
- Abasolo, J. (2002). El «Personaje» en la novela policial [en línea]. *La Gangsterera*, 14. Recuperado el 10 de septiembre de 2009 de <http://gangsterera.free.fr/hispersonajesNP.htm>
- Ainaud, J., Espunya, A. y Pujol, D. (2003). *Manual de traducció anglès-català*. Vic: Eumo Editorial.

- Alarcos Llorach, E. (1995). *Gramática de la lengua española*. Madrid: Espasa-Calpe.
- Almela Pérez, R. (1999). *Procedimientos de formación de palabras en español*. Barcelona: Ariel.
- Alonso, A. (1982). *Estudios lingüísticos: temas españoles* (3.^a ed.). Madrid: Gredos.
- Alvar Ezquerro, M., Blanco Rodríguez, M. J. y Pérez Lagos, F. (1994). Diseño de un corpus español en el marco de un corpus europeo. En M. Alvar Ezquerro y J. A. Villena Ponsoda (coords.), *Estudios para un corpus del español* (pp. 9-29). Málaga: Universidad de Málaga.
- Álvarez, C. C. (2004). Mandanga [en línea]. Recuperado el 11 de agosto de 2015 de *Futuro pasado* <http://www.futuropasado.com/?p=287>
- Álvarez de Miranda, P. (2014). La tilde en los demostrativos y en *solo* [en línea]. Recuperado el 5 de junio de 2014 de *Centro Virtual Cervantes* http://cvc.cervantes.es/el_rinconete/anteriores/abril_14/04042014_01.htm
- Alvarez-de-Mon y Rego, I. (2001). Encapsulation and prospection in written scientific English [en línea]. *Estudios Ingleses de la Universidad Complutense*, 9, 81-101. Recuperado el 17 de agosto de 2015 de <http://revistas.ucm.es/index.php/EIUC/article/view/EIUC0101110081A>
- Alvarez-de-Mon y Rego, I. (2006). A Contrastive Study of Encapsulation and Prospection in Written Scientific Text. En J. Flowerdew y M. Gotti (eds.), *Studies in Specialized Discourse* (vol. 41, pp. 21-41). Bern: Peter Lang.
- Arnscheidt, G. (2002). Andreu Martín: El hombre de la navaja (1992) / L'home que tenia raor (1997) o la imposibilidad de una novela bilingüe. En P. Arnau i Segarra, P. Joan i Tous y M. Tietz (eds.), *Escribir entre dos lenguas: Escritores catalanes y la elección de la lengua literaria. Escriure entre dues llengües: Escriptors catalans i l'elecció de la llengua literària* (pp. 111-130). Kassel: Editio Reichenberger.
- Atkins, S., Clear, J. y Ostler, N. (1991). Corpus Design Criteria [en línea]. Recuperado el 29 de octubre de 2010 de *British National Corpus* <http://www.natcorp.ox.ac.uk/archive/vault/tgaw02.pdf>
- Atkins, S., Clear, J. y Ostler, N. (1992). Corpus Design Criteria. *Literary and Linguistic Computing*, 7 (1), 1-16. doi: 10.1093/lc/7.1.1
- Austermühl, F. (2001). *Electronic tools for translation*. Manchester: St. Jerome Publishing.
- Axelsson, K. (2011). *Tag questions in fiction dialogue* [en línea] (tesis doctoral defendida en mayo de 2011). Göteborg: University of Gothenburg. Recuperada el 10 de junio de 2015 de <https://gupea.ub.gu.se/handle/2077/24047>
- Balaskó, M. (2008). What Does the Figure Show? Patterns of Translationese in a Hungarian Comparable Corpus. [en línea] *trans-kom*, 1 (1), 58-73. Recuperado el 28 de agosto de 2015 de http://www.trans-kom.eu/bd01nr01/trans-kom_01_01_05_Balasko_What_Does_the_Figure_Show.20080707.pdf

- Baker, M. (1993). Corpus Linguistics and Translation Studies. Implications and Applications. En M. Baker, G. Francis y E. Tognini-Bonelli (eds.), *Text and Technology. In Honour of John Sinclair* (pp. 233-250). Amsterdam: John Benjamins.
- Baker, M. (1995). Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. *Target*, 7 (2), 223-243. doi: 10.1075/target.7.2.03bak
- Baker, M. (1996). Corpus-based translation studies: the challenges that lie ahead. En H. Somers (ed.), *Terminology, LSP and translation studies in language engineering* (175-186). Amsterdam: John Benjamins.
- Baker, M. (1998). Réexplorer la langue de la traduction: une approche par corpus. *Meta*, 43 (4), 480-485. doi: 10.7202/001951ar
- Baker, M. (2000). Towards a Methodology for Investigating the Style of a Literary Translator. *Target*, 12 (2), 241-266. doi: 10.1075/target.12.2.04bak
- Baker, M. (2004). A Corpus-based View of Similarity and Difference in Translation. *International Journal of Corpus Linguistics*, 9 (2), 167-193. doi: 10.1075/ijcl.9.2.02bak
- Baker, M. (2007). Patterns of Idiomaticity in Translated vs. Non-translated Text. *Belgian Journal of Linguistics*, 21 (1), 11-21. doi: 10.1075/bjl.21.02bak
- Barlow, M. (2003). ParaConc. A concordance for parallel texts [en línea]. Recuperado el 10 de marzo de 2012 de <http://www.athel.com/paraconc.pdf>
- Baroni, M. y Evert, S. (2006). Chapter 38: Statistical methods for corpus exploitation [versión en línea]. En A. Lüdeling y M. Kytö (eds.), *Corpus Linguistics. An International Handbook* (vol. 2, pp. 777-802). Berlin: Mouton de Gruyter. Recuperado el 8 de mayo de 2014 de <http://clic.cimec.unitn.it/marco/publications/chapter38.pdf>
- Baumgarten, N., Meyer, B. y Özçetin, D. (2008). Explicitness in Translation and Interpreting: A Critical Review and Some Empirical Evidence (of an Elusive Concept). *Across Languages and Cultures*, 9 (2), 177-203. doi: 10.1556/Acr.9.2008.2.2
- Becher, V. (2010a). Towards a More Rigorous Treatment of the Explicitation Hypothesis in Translation Studies [en línea]. *trans-kom*, 3 (1), 1-25. Recuperado el 28 de agosto de 2015 de http://www.trans-kom.eu/bd03nr01/trans-kom_03_01_01_Becher_Explicitation.20100531.pdf
- Becher, V. (2010b). Abandoning the Notion of “Translation-Inherent” Explication: Against a Dogma of Translation Studies. *Across Languages and Cultures* 11 (1), 1-28. doi: 10.1556/Acr.11.2010.1.1
- Becher, V. (2011). When and why do translators add connectives? A corpus-based study. *Target*, 23 (1), 26-47. doi 10.1075/target.23.1.02bec
- Bello, A. (1978). *Gramática de la lengua castellana*. Madrid: Edaf.
- Bernardini, S. (2011). Monolingual comparable corpora and parallel corpora in the search for features of translated language [en línea]. *SYNAPS – A Journal of Professional*

- Communication*, 26, 2-13. Recuperado el 28 de agosto de 2015 de <http://www.nhh.no/Default.aspx?ID=2242>
- Bernardini, S. y Ferraresi, A. (2011). Practice, Description and Theory Come Together - Normalization or Interference in Italian Technical Translation? *Meta*, 56 (2), 226-246. doi: 10.7202/1006174ar
- Bernardini, S. y Zanettin, F. (2004). When is a universal not a universal? Some limits of current corpus-based methodologies for the investigation of translation universals. En A. Mauranen y P. Kujamäki (eds.), *Translation Universals. Do They Exist?* (pp. 51-62). Amsterdam: John Benjamins.
- Berriatúa, M. (2006). *Manuel Vázquez Montalbán: El premio* [en línea]. Recuperado el 11 de agosto de 2015 de http://www.berriatua.com/Comentarios_de_Texto/M.Vazquez.Montalban_El.Premio.pdf
- Biber, D. (1990). Methodological Issues Regarding Corpus-based Analyses of Linguistic Variation. *Literary and Linguistic Computing*, 5 (4), 257-269. doi: 10.1093/lc/5.4.257
- Biber, D., Conrad, S. y Reppen, R. (1994). Corpus-based Approaches to Issues in Applied Linguistics. *Applied Linguistics*, 15 (2), 169-189. doi: 10.1093/applin/15.2.169
- Biber, D., Johansson, S., Leech, G., Conrad, S., y Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
- Biblioteca Nacional de España (2014). *Novela policíaca: guía de recursos bibliográficos* [en línea]. Recuperado el 3 de septiembre de 2015 de <http://www.bne.es/permalink/dd094150-af80-11df-8f54-0018fe323fc3.html>
- Blas Arroyo, J. L. (2000). *Gramáticas en contacto. Un modelo de análisis variacionista para la desambiguación de los fenómenos de contacto en el discurso bilingüe catalán-español*. Munich: Lincom.
- Blum-Kulka, S. (1986). Shifts of cohesion and coherence in translation. En J. House y S. Blum-Kulka (eds.), *Interlingual and intercultural communication* (pp. 17-36). Tübingen: Gunter Narr.
- Borja, A. (2007). Corpora for translators in Spain. The CDJ-GITRAD Corpus and the GENTT Project. En G. Anderman y M. Rogers (eds.), *Incorporating corpora: The Linguist and the Translator* (pp. 243-265). Clevedon: Multilingual Matters.
- Borrero Barrera, M. J. y Cala Carvajal, R. (2002). La carta como documento lingüístico: La deixis en el discurso epistolar [en línea]. *Revista electrónica de estudios filológicos*, n.º 4. Recuperado el 6 de abril de 2014 de <http://www.um.es/tonosdigital/znum4/estudios/discursoepistolar.htm>
- Botley, S. y McEnery, T. (2001). Demonstratives in English: A Corpus-Based Study. *Journal of English Linguistics*, 29 (7). doi: 10.1177/00754240122005170
- Bowker, L. (2001). Towards a Methodology for a Corpus-Based Approach to Translation Evaluation. *Meta*, 46 (2), 345-364. doi: 10.7202/002135ar
- Bowker, L. (2002). *Computer-aided translation technologies*. Ottawa: University of Ottawa Press.

- Bowker, L. y Pearson, J. (2002). *Working with Specialized Language: A practical guide to using corpora*. London: Routledge.
- Calderón, E. L. (7 de julio de 2004). Andreu Martín: «Los gurús de la cultura ignoran la novela negra» [en línea]. Recuperado el 27 de julio de 2009 de *Elmundolibro.com*
http://www.elmundo.es/elmundolibro/2004/07/07/narrativa_espanyol/1089201028.html
- Cantera Ortiz de Urbina, J. (2012). Por la pureza y por el esplendor de nuestro idioma [en línea]. En J. Sevilla Muñoz y M. I. T. Zurdo Ruiz-Ayúcar (coords.), *Biblioteca fraseológica y paremiológica*, serie «Monografías», 1. Recuperado el 6 de agosto de 2015 de
http://cvc.cervantes.es/lengua/biblioteca_fraseologica/n1_cantera/
- Cantos Gómez (2002). Do we Need Statistics when we have Linguistics? *D.E.L.T.A.*, 18 (2), 233-271. doi: 10.1590/S0102-44502002000200003
- Cappelle, B. (2012). English is Less Rich in Manner-of-Motion Verbs when Translated from French. *Across Languages and Cultures*, 13 (2), 173-195. doi: 10.1556/Acr.13.2012.2.3
- Chalker, Sylvia. (1984) *Current English Grammar*. London: Macmillan.
- Cheong, H. (2006). Target Text Contraction in English-into-Korean Translations: A Contradiction of Presumed Translation Universals? *Meta*, 51 (2), 343-367. doi: 10.7202/013261ar
- Chesterman, A. (2003). A causal model for Translation Studies. En S. Granger, J. Lerot & S. Petch-Tyson (eds.), *Corpus-based approaches to contrastive linguistics and translation studies* (pp. 15-28). Amsterdam: Rodopi.
- Chesterman, A. (2004a). Hypotheses about translation universals. En G. Hansen, K. Malmkjær & D. Gile (eds.), *Claims, changes and challenges in translation studies* (pp. 1-13). Amsterdam: John Benjamins.
- Chesterman, A. (2004b). Beyond the particular. En A. Mauranen y P. Kujamäki (eds.), *Translation universals: do they exist?* (pp. 33-49). Amsterdam: John Benjamins.
- Chesterman, A. (2007). What is a Unique Item? En Y. Gambier, M. Shlesinger y R. Stolze (eds.), *Doubts and Directions in Translation Studies* (pp. 3-13). Amsterdam: John Benjamins.
- Chesterman, A. (2010). Why study translation universals? [en línea]. *Acta Translatologica Helsingiensia (ATH)*, 1, 38-48. Recuperado el 28 de agosto de 2015 de
<https://helda.helsinki.fi/handle/10138/24319>
- Chesterman, A. (2014). Translation Studies Forum: Universalism in translation studies. *Translation Studies*, 7 (1), 82-90. doi: 10.1080/14781700.2013.828904
- Clements, T. et ál. (23 de febrero de 2008). 50 crime writers to read before you die [en línea]. Recuperado el 27 de septiembre de 2015 de *The Telegraph*
<http://www.telegraph.co.uk/culture/books/3671363/50-crime-writers-to-read-before-you-die.html>
- Close, R. A. (1985). *A reference grammar for students of English*. Longman: England.

- Coe, R. (2002). It's the Effect Size, Stupid: What effect size is and why it is important [en línea]. Artículo presentado en la Annual Conference of the British Educational Research Association, University of Exeter, 12-14 de septiembre de 2002. Recuperado el 28 de septiembre de 2015 de <http://www.leeds.ac.uk/educol/documents/00002182.htm>
- Conte, R. (5 de agosto de 1984). Policías y ladrones o el juego que quería ser real [en línea]. *El País, Libros*, pp 1-2. Recuperado el 11 de julio de 2009 de <http://www.vespito.net/mvm/conte.html>
- Corpas Pastor, G. (2008). *Investigar con corpus en traducción: los retos de un nuevo paradigma*. Frankfurt am Main: Peter Lang.
- Corpas Pastor, G., Mitkov, R., Afzal, N. y Pekar, V. (2008). Translation universals: do they exist? A corpus-based NLP study of convergence and simplification [en línea]. En *Proceedings of the eighth conference of the Association for Machine Translation in the Americas (AMTA-08)*. Recuperado el 28 de agosto de 2015 de <http://rgcl.wlv.ac.uk/papers/AMTA2008.pdf>
- Cuenca, M. J. y Ribera, J. (2011) Deictic neutralization and overmarking: demonstratives in the translation of fiction (English-Catalan) [en línea]. En M. L. Carrió Pastor y M. Á. Candel Mora, *Actas del III Congreso Internacional de Lingüística de Corpus. Las tecnologías de la información y las comunicaciones: presente y futuro en el análisis de corpus* (pp. 371-380). Editorial Universitat Politècnica de València. Recuperado el 14 de enero de 2013 de <http://www.aelinco.es/es/actas-publicaciones-relacionadas>
- Cruzado, A. A. (1982). En R. Nash y D. Belaval (eds.), *Readings in Spanish-English Contrastive Linguistics* (pp. 66-92). San Juan: Inter American University Press.
- Davies, M. (2009). The 385+ million word *Corpus of Contemporary American English* (1990–2008+). Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics* 14 (2), 159-190. doi: 10.1075/ijcl.14.2.02dav
- Dai, G. y Xiao, R. (2011). “SL Shining Through” in Translational Language: A Corpus-based Study of Chinese Translation of English Passives [en línea]. En Xiao, R. (ed.), *Proceedings of The International Symposium on Using Corpora in Contrastive and Translation Studies 2010 Conference (UCCTS2010)*. Lancaster: Lancaster University. Recuperado el 28 de agosto de 2015 de http://www.lancs.ac.uk/fass/projects/corpus/UCCTS2010Proceedings/papers/Dai_Xiao.pdf
- Dayrell, C. (2007). A quantitative approach to compare collocational patterns in translated and non-translated texts. *International Journal of Corpus Linguistics*, 12 (3), 375-414. doi: 10.1075/ijcl.12.3.04day
- Dayrell, C. (2008). Investigating the Preference of Translators for Recurrent Lexical Patterns: A Corpus-based Study [en línea]. *trans-kom* 1 (1), 36-57. Recuperado el 28 de agosto de 2015 de http://www.trans-kom.eu/bd01nr01/trans-kom_01_01_04_Dayrell_Investigating_the_Preference.20080707.pdf

- De Haan, P. (1996). More on the Language of dialogue in fiction [en línea]. *ICAME Journal*, 20. 23-40. Recuperado el 20 de noviembre de 2012 de <http://clu.uni.no/icame/ij20/dehaan.pdf>
- De Kock, J. y Gómez Molina, C. (1992). Los pronombres demostrativos en registros análogos y diferentes. En J. de Kock, C. Gómez Molina y R. Verdonk, *Gramática española: Enseñanza e Investigación. II. Gramática. 5. Los pronombres demostrativos y relativos* (pp. 9-90). Salamanca: Ediciones Universidad de Salamanca.
- De Miguel, A. (22 de septiembre de 2010). El habla coloquial [en línea]. Recuperado el 12 de agosto de 2015 de *Libertad Digital* <http://www.libertaddigital.com/opinion/amando-de-miguel/el-habla-coloquial-56366/>
- Deira, J. M. (3 de diciembre de 2008) ¿Qué es un bandarra? [en línea]. Recuperado el 12 de agosto de 2015 de *Andalucía Información* <http://andaluciainformacion.es/andalucia/9803/qu-es-un-bandarra/>
- Delaere, I., De Sutter, G. y Plevoets K. (2012). Is translated language more standardized than non-translated language? Using profile-based correspondence analysis for measuring linguistic distances between language varieties. *Target*, 24 (2), 203-224. doi: 10.1075/target.24.2.01del
- Denturck, K. (2012). Explicitation Vs. Implicitation: A Bidirectional Corpus-Based Analysis Of Causal Connectives In French And Dutch Translations. *Across Languages and Cultures*, 13 (2), 211–227. doi: 10.1556/Acr.13.2012.2.5
- Ebeling, J. (1998). Contrastive Linguistics, Translation, and Parallel Corpora. *Meta*, 43 (4), 602-615. doi: 10.7202/002692ar
- Eder, M. (2013). Mind your corpus: systematic errors in authorship attribution. *Literary and Linguistic Computing*, 28 (4), 603-614. doi:10.1093/lc/fqt039
- Eguren Gutiérrez, L. J. (1999). Pronombres y adverbios demostrativos. Las relaciones deícticas. En Bosque, I. y Demonte, V. (dirs.). *Gramática descriptiva de la lengua española* (volumen I, pp. 929-972). Madrid: Espasa Calpe.
- Erichsen, G. (2007a). Spanish Diminutives. Suffixes Often Indicate More Than Size [en línea]. Recuperado el 31 de julio de 2015 de *About.com* <http://spanish.about.com/od/nouns/a/diminutives.htm>
- Erichsen, G. (2007b). Augmentative Suffixes. Word Endings Can Expand Vocabulary, Provide Negative Connotation [en línea]. Recuperado el 31 de julio de 2015 de *About.com* http://spanish.about.com/od/nouns/a/augment_suffix.htm
- Eskola, S. (2004). Untypical Frequencies in Translated Language: A Corpus-Based Study on a Literary Corpus of Translated and Non-Translated Finnish. En A. Mauranen y P. Kujamäki (eds.), *Translation Universals. Do They Exist?* (pp. 83-99). Amsterdam: John Benjamins.
- Evert, S. et ál. (2010). The IMS Open Corpus Workbench (CWB): CQP Query Language Tutorial. CWB version 3.0. [en línea]. Recuperado el 5 de marzo de 2012 de http://cwb.sourceforge.net/files/CQP_Tutorial.pdf

- Evert, S. y Hardie, A. (2011). Twenty-first century Corpus Workbench: Updating a query architecture for the new millennium [en línea]. En *Proceedings of the Corpus Linguistics 2011 conference*. Birmingham: University of Birmingham. Recuperado el 12 de septiembre de 2015 de http://eprints.lancs.ac.uk/62721/1/Paper_153.pdf
- Fente Gómez, R. (1971). *Estilística del verbo en inglés y español*. Madrid: Sociedad General Española de Librería.
- Fernández de Castro, F. (1999). *Las perífrasis verbales en el español actual*. Madrid: Gredos.
- García Fernández, L. (2012). *Las perífrasis verbales*. Barcelona: Castalia (Edhasa).
- Fantinuoli, C. y Zanettin, F. (2015). Creating and using multilingual corpora in translation studies [en línea]. En C. Fantinuoli, y F. Zanettin (eds.), *New directions in corpus-based translation studies* (pp. 1-10). Berlin: Language Science Press. Recuperado el 28 de agosto de 2015 de <http://langsci-press.org/catalog/book/76>
- Field, S. (1994). *El libro del guión*. Madrid: Plot Ediciones.
- Frankenberg-García, A. (2009). Are translations longer than source texts? A corpus-based study of explicitation. 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* (pp. 47-58). Amsterdam: John Benjamins.
- Frankenberg-García, A. y Santos, D. (2003). Introducing COMPARA: the Portuguese-English Parallel Corpus. En F. Zanettin, S. Bernardini y D. Stewart (eds.), *Corpora in Translator Education* (pp. 71-87). Manchester: St. Jerome.
- Frawley, W. (1984). Prolegomenon to a theory of translation. En W. Frawley (ed.), *Translation literary, linguistic, and philosophical perspectives* (pp. 159-175). Newark: University of Delaware Press.
- Gabrielatos, C. y Sarmiento, S. (2006). Central Modals in an Aviation Corpus: Frequency and Distribution [en línea]. *Letras de Hoje*, 41 (2), 215-240. Recuperado el 10 de junio de 2015 de <http://eprints.lancs.ac.uk/529/1/LH-2006-25.pdf>
- Gale, W. A. y Church, K. W. (1993). A Program for Aligning Sentences in Bilingual Corpora. *Computational Linguistics*, 19 (1), 75-102. doi: 10.3115/981344.981367
- García Izquierdo, I. (2005). Corpus electrónico, género textual y traducción: metodología, concepto y ámbito de la Enciclopedia electrónica para traductores GENTT. *Meta*, 50 (4). doi: 10.7202/019918ar
- Gellerstam, M. (1996). Translation as a source for cross-linguistic studies. En K. Aijmer, B. Altenberg y M. Johansson (eds.), *Languages in Contrast: Papers from a symposium on text-based cross-linguistic studies in Lund, 4-5 March 1994* (pp. 53-62). Lund: Lund University Press.
- Gile, D. (1998). Observational Studies and Experimental Studies in the Investigation of Conference Interpreting. *Target*, 10 (1), 69-93. doi: 10.1075/target.10.1.04gil

- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8 (3), 183-206. doi: 10.1515/cogl.1997.8.3.183
- Gili Gaya, S. (1981/1961). *Curso superior de sintaxis española* (13.^a ed.). Barcelona: Biblograf.
- Goethals, P. (2007). Corpus-Driven Hypothesis Generation in Translation Studies, Contrastive Linguistics and Text Linguistics: A Case Study of Demonstratives in Spanish and Dutch Parallel Texts. *Belgian Journal of Linguistics*, 21 (1), 87-103. doi: 10.1075/bjl.21.07goe
- Goethals, P. (2013). Demonstratives on the move: What translational shifts tell us about demonstrative determiners and definite articles in Spanish and Dutch [en línea]. *Linguistics* 51 (3), 517-553. doi: 10.1515/ling-2013-0019
- Gómez Torrego, L. (1988). *Perífrasis verbales: sintaxis, semántica y estilística*. Madrid: Arco/Libros.
- Gómez Torrego, L. (1997). *Gramática didáctica del español*. Madrid: Ediciones SM.
- Gómez Torrego, L. (1999). Los verbos auxiliares. Las perífrasis verbales de infinitivo. En I. Bosque y V. Demonte (eds.), *Gramática descriptiva de la lengua española* (tomo. 2, pp. 3323-3389). Madrid: Espasa.
- Gómez Torrego, L. (2002). *Nuevo manual de español correcto*. Madrid: Arco/Libros.
- Gooch, A. (1970). *Diminutive, Augmentative and Pejorative Suffixes in Modern Spanish (A Guide to their Use ad Meaning)* (2.^a ed.). Oxford: Pergamon Press.
- Granger S. (2003). The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies? En: S. Granger, J. Lerot y S. Petch-Tyson, (eds.), *Corpus-based Approaches to Contrastive Linguistics and Translation Studies* (pp. 17-29). Amsterdam: Rodopi.
- Granger, S. y Tribble, C. (1998). Learner corpus data in the foreign language classroom: form-focused instruction and data-driven learning. En S. Granger (ed.). *Learner English on Computer* (pp. 199-209). Harlow: Longman.
- Gries, S. T. (2005). Null-hypothesis significance testing of word frequencies: a follow-up on Kilgarriff. *Corpus Linguistics and Linguistic Theory*, 12, 277-294. doi: 1613-7027/05/00010277
- Gries, S. T. (2010). Corpus linguistics and theoretical linguistics. A love-hate relationship? Not necessarily... *International Journal of Corpus Linguistics* 15 (3), 327-343. doi: 10.1075/ijcl.15.3.02gri
- Gumul, E. (2006). Explication in Simultaneous Interpreting: A Strategy or a By-Product of Language Mediation? *Across Languages and Cultures*, 7 (2), 171-190. doi: 10.1556/Acr.7.2006.2.2
- Halverson, S. (2003). The Cognitive Basis of Translation Universals. *Target*, 15 (2), 197-241. doi: 10.1075/target.15.2.02hal
- Halverson, S. (2010). Cognitive translation studies: Developments in theory and method. En G. M. Shreve y E. Angelone (eds.), *Translation and Cognition* (pp. 349-369). Amsterdam: John Benjamins.

- Hardie, A. (2012) CQPweb — combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, 17 (3), 380-409. doi: 10.1075/ijcl.17.3.04har
- Hardie, A. (2014a). Modest XML for Corpora: Not a standard, but a suggestion. *ICAME Journal*, 38, 73-103. doi: 10.2478/icame-2014-0004
- Hardie, A. (2014b). Log Ratio – an informal introduction [en línea]. Recuperado el 27 de septiembre de 2015 de *ESRC Centre for Corpus Approaches to Social Science (CASS) (Lancaster University)* <http://cass.lancs.ac.uk/?p=1133>
- Hareide, L. (2013). *Testing the Gravitational Pull Hypothesis in translation: A corpus-based study of the gerund in translated Spanish* (tesis doctoral defendida en mayo de 2014). Bergen: University of Bergen.
- Hareide, L. y Hofland, K. (2012). Compiling a Norwegian-Spanish parallel corpus: Methods and challenges. En M. P. Oakes y Ji, M. (eds.), *Quantitative Methods in Corpus-Based Translation Studies: A practical guide to descriptive translation research* (pp. 75-113). Amsterdam: John Benjamins.
- Heaton, J. B. (1979). *Prepositions and Adverbial Particles*. London: Longman.
- Holmes, J. S. (2000/1972). The Name and Nature of Translation Studies. En L. Venuti (ed.), *The Translation Studies Reader* (pp. 172-185). London: Routledge.
- House, J. (2008). Beyond Intervention: Universals in Translation? [en línea]. *trans-kom*, 1, 6-19. Recuperado el 13 de septiembre de 2009 de http://www.trans-kom.eu/bd01nr01/trans-kom_01_01_02_House_Beyond_Intervention.20080707.pdf
- House, J. (2011). Using Translation and Parallel Text Corpora to Investigate the Influence of Global English on Textual Norms in Other Languages. En A. Kruger, K. Wallmach y J. Munday (eds.), *Corpus-Based Translation Studies: Research and Applications* (pp. 187-208). London: Continuum.
- Huddleston, R. D; Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Izquierdo, M., Hofland, K. y Reigem, Ø. (2008). The ACTRES parallel corpus: an English–Spanish translation corpus. *Corpora*, 3 (1), 31-41. doi: 10.3366/E1749503208000051
- Jakobsen, A. L. (1986). Lexical Selection and Creation in Translation. En I. Lindblad, y M. Ljung (eds), *Proceedings from the Third Nordic Conference for English Studies, Hässelby, Sept 25–27, 1986, Volume 1*. (pp. 101-112). Stockholm: Almqvist and Wiksell International.
- Janerka, M. (2010). *La novela policiaca española (1975-2005) ante los problemas de la sociedad española contemporánea*. Vigo: Editorial Academia del Hispanismo.
- Jantunen, J. H. (2004). Untypical patterns in translations: Issues on corpus methodology and synonymy. En A. Mauranen & P. Kujamäki (eds.), *Translation universals: Do they exist?* (pp. 101-126). Amsterdam: John Benjamins.

- Jääskeläinen, R. (2004). The fate of “The Families of Medellín”: Tampering with a potential translation universal in the translation class. En A. Mauranen y P. Kujamäki (eds.), *Translation universals: Do they exist?* (pp. 205-214). Amsterdam: John Benjamins.
- Jiménez-Crespo, M. A. (2011). The future of general tendencies in translation: Explicitation in web localization. *Target* 23 (1), 3-25. doi: 10.1075/target.23.1.01jim
- Johansson, S. (marzo de 1978). *Newsletter of the International Computer Archive of Modern English (ICAME), 1* [en línea]. Bergen: NAVF's EDB-senter for humanistisk forskning, University of Bergen. Recuperado el 25 de marzo de 2013 de http://clu.uni.no/icame/archives/No_1_ICAME_News_index.pdf
- Johansson, S. (1998). On the role of corpora in cross-linguistic research. En S. Johansson y S. Oksefjell (eds.), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies* (pp. 3-24). Amsterdam: Rodopi.
- Johansson, S. (2007). *Seeing through Multilingual Corpora: On the use of corpora in contrastive studies*. Amsterdam: John Benjamins.
- Johansson, S., Ebeling, J. y Oksefjell, S. (1999/2002). *English-Norwegian Parallel Corpus: Manual* [en línea]. Recuperado el 16 de septiembre de 2015 de <http://www.hf.uio.no/ilos/english/services/omc/enpc/ENPCmanual.pdf>
- Johansson, S. y Hofland, K. (1998). The Translation Corpus Aligner: A program for automatic alignment of parallel texts. En S. Johansson y S. Oksefjell (eds.), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies* (pp. 87-101). Amsterdam: Rodopi.
- Kaplanová, V. (2010). *Clasificación de los sufijos apreciativos españoles y su uso en el cuento de hadas Blancanieves y los siete enanitos gigantes* [en línea]. Tesis de fin de carrera, Universidad de Masaryk, República Checa. Recuperada el 3 de agosto de 2015 de https://is.muni.cz/th/263223/ff_b/?lang=en;id=100028
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. London: Longman.
- Kenny, D. (1998). Creatures of Habit? What Translators Usually Do with Words. *Meta*, 43 (4), 515-523. doi: 10.7202/003302ar
- Kenny, D. (2001). *Lexis and Creativity in Translation. A Corpus-based Study*. Manchester: St. Jerome.
- Kilgariff, A. (2001). Comparing Corpora. *International Journal Of Corpus Linguistics*, 6 (1), 97-133. doi: 10.1075/ijcl.6.1.05kil
- Kilgariff, A. (2005). Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory*, 1 (2), 263-275. doi: 10.1515/clt.2005.1.2.263
- Klaudy, K. (2009). Explicitation. En M. Baker y G. Saldanha (eds.), *Routledge Encyclopedia of Translation Studies* (pp. 104-108). London: Routledge.
- Klee, C. A. y Lynch, A. (2009). *El español en contacto con otras lenguas* [en línea]. Washington: Georgetown University Press. Consultado el 3 de marzo de 2013 en

<https://books.google.es/books?id=KQ98DtEoqFkC&lpg=PP1&hl=es&pg=PP1#v=onepage&q&f=false>

- Konšalová, P. (2007). Explicitation as a Universal in Syntactic De/Condensation. *Across Languages and Cultures* 8 (1), 17–32. doi: 10.1556/Acr.8.2007.1.2
- Kruger, H. y van Rooy, B. (2012). Register and the Features of Translated Language. *Across Languages and Cultures*, 13 (1), 33-65. doi: 10.1556/Acr.13.2012.1.3
- Kujamäki, P. (2004). What happens to "unique items" in learner's translations? "Theories" and "concepts" as a challenge for novice's views on "good translation". En A. Mauranen & P. Kujamäki (eds.), *Translation universals: Do they exist?* (pp. 187-204). Amsterdam: John Benjamins.
- Labrador, B. (2007). Contrasting ways of expressing restriction in English and Spanish and suggesting translational options into Spanish. *Languages in Contrast* 7 (1), 29-52. doi: 10.1075/lic.7.1.03lab
- Labrador, B. (2011). A corpus-based study of the use of Spanish demonstratives as translation equivalents of English demonstratives [en línea]. *Perspectives: Studies in Translatology*, 19 (1), 71-8. doi: 10.1080/0907676X.2010.481047
- Lang, M. F. (1990). *Spanish Word Formation. Productive Derivational Morphology in the Modern Lexis*. London: Routledge.
- Lanstyák, I. y Heltai, P. (2012). Universals in Language Contact and Translation. *Across Languages and Cultures*, 13 (1), 99-121. doi: 10.1556/Acr.13.2012.1.6
- Lapshinova-Koltunski, E. (2015). Variation in translation: Evidence from corpora [en línea]. En C. Fantinuoli, y F. Zanettin (eds.), *New directions in corpus-based translation studies* (pp. 93-113). Berlin: Language Science Press. Recuperado el 28 de agosto de 2015 de <http://langsci-press.org/catalog/book/76>
- Laviosa, S. (1997). How Comparable Can 'Comparable Corpora' Be? *Target*, 9 (2), 289-319. doi: 10.1075/target.9.2.051av
- Laviosa, S. (1998a). The Corpus-based Approach: A New Paradigm in Translation Studies. *Meta* 43 (4), 474-479. doi: 10.7202/003424ar
- Laviosa, S. (1998b). Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose. *Meta*, 43 (4), 557–570. doi: 10.7202/003425ar
- Laviosa, S. (2000). TEC: a resource for studying what is "in" and "of" translational English. *Across Languages and Cultures* 1 (2), 159–177. doi: 10.1556/Acr.1.2000.2.2
- Laviosa, S. (2002). *Corpus-based translation studies: theory, findings, applications*. Amsterdam: Rodopi.
- Laviosa, S. (2003). Corpora and the Translator. En H. Somers (ed.), *Computer and Translation. A Translator's Guide*. (pp. 105-117). Amsterdam: John Benjamins.

- Laviosa, S. (2004). Corpus-Based Translation Studies: Where does it come from? Where is it going? *TRADTERM*, 10, 29-57. doi: 10.11606/issn.2317-9511.tradterm.2004.47039
- Laviosa, S. (2007). Studying Anglicisms with Comparable and Parallel Corpora. *Belgian Journal of Linguistics*, 21 (1), 123-136. doi: 10.1075/bjl.21.09lav
- Laviosa, S. (2011). Corpus-Based Translation Studies: Where does it come from? Where is it going? En A. Kruger, K. Wallmach y J. Munday (eds.), *Corpus-based Translation Studies: Research and Applications* (pp. 13-32). London: Continuum.
- Laviosa-Braithwaite, S. (1997). Investigating Simplification in an English Comparable Corpus of Newspaper Articles. En K. Klaudy y J. Kohn (eds.) *Transfere Necesse Est. Proceedings of the 2nd International Conference on Current Trends in Studies of Translation and Interpreting, 5-7 September, 1996, Budapest, Hungary* (pp. 531-540). Budapest: Scholastica.
- Lázaro Mora, F. A. (1999). La derivación apreciativa. En I. Bosque y V. Demonte (eds.) *Gramática descriptiva de la lengua española* (tomo 3, pp. 4647-4682). Madrid: Espasa.
- Lee, D. Y. W. (2001). Genres, Registers, Text Types, Domains, and Styles: Clarifying the Concepts and Navigating a Path through the BNC Jungle [en línea]. *Language Learning & Technology*, 5 (3), 37-72. Recuperado el 9 de septiembre de 2015 de <http://www.llt.msu.edu/vol5num3/pdf/lee.pdf>
- 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* (pp. 105-122). Berlin: Mouton de Gruyter.
- Leech, G. (1998). Preface. En S. Granger (ed.). *Learner English on Computer* (pp. XIV-XX). Harlow: Longman.
- Leonetti, M. (1999). El artículo. En I. Bosque y V. Demonte (eds.). *Gramática descriptiva de la lengua española* (vol. 1, pp. 787-890) Madrid: Espasa.
- López Merino, J. M. (2005). El primer Andreu Martín (1979-1989): variaciones y reincidencias [en línea]. *Espéculo. Revista de estudios literarios*, 29. Recuperado el 27 de julio de 2009 de <http://www.ucm.es/info/especulo/numero29/amartin.html>
- López Samaniego, A. (2010). Documentos profesionales con destinatarios no expertos. El empleo de los mecanismos referenciales en la sentencia del 11M [en línea]. *Revista Signos*, 43 (72), 99-123. doi: 10.4067/S0718-09342010000100005
- Lowry, R. (1999-2012). *Concepts and Applications of Inferential Statistics* [en línea]. Recuperado el 17 de septiembre de 2015 de <http://vassarstats.net/textbook/index.html>
- Macías Villalobos, C. (2006). *El demostrativo en Miguel Delibes* [edición digital a partir del texto original de la tesis doctoral]. Alicante: Biblioteca Virtual Miguel de Cervantes. Recuperado el 23 de febrero de 2014 de <http://www.cervantesvirtual.com/nd/ark:/59851/bmcsq989>
- Macken, L., De Clercq, O. y Paulussen, H. (2011). Dutch Parallel Corpus: A Balanced Copyright-Cleared Parallel Corpus. *Meta*, 56 (2), 374-390. doi: 10.7202/1006182ar

- Mackenzie, J. L. y Martínez Caro, Elena (2012) *Compare and contrast. An English grammar for speakers of Spanish*. Granada: Comares.
- Mahlberg, M. (2006). Lexical cohesion: Corpus linguistic theory and its application in English language teaching. *International Journal of Corpus Linguistics* 11 (3), 363-383. doi: 10.1075/bct.17.07mah
- Mahlberg, M. (2006). *English General Nouns: A corpus theoretical approach*. Amsterdam: John Benjamins.
- Malmkjær, K. (1997). Punctuation in Hans Christian Andersen's stories and in their translations into English. En F. Poyatos (ed.), *Nonverbal Communication and Translation: New perspectives and challenges in literature, interpretation and the media* (pp. 151-162). Amsterdam: John Benjamins.
- Malmkjær, K. (2007). Norms and Nature in Translation Studies. En G. Anderman y M. Rogers (eds.), *Incorporating corpora: The Linguist and the Translator* (pp. 49-59). Clevedon: Multilingual Matters.
- Marco, J. (2012). An Analysis of Explicitation in the Covalt Corpus: The Case of the Substituting Pronoun One(s) and its Translation into Catalan. *Across Languages and Cultures* 13 (2), 229-246. doi: 10.1556/Acr.13.2012.2.6
- Marco, J. y Guzmán, J. (2007). A Corpus-Based Analysis of Lexical Items Conveying Body Language in the COVALT Corpus. *Belgian Journal of Linguistics*, 21 (1), 155-170. doi: 10.1075/bjl.21.11mar
- Marco, J., Verdegil, J. y Hurtado, A. (1999). La traducción literaria. En Hurtado, A. (ed.), *Enseñar a traducir: Metodología en la formación de traductores e intérpretes* (pp. 167-181). Madrid: Edelsa.
- Marcos Marín, F. (1975). *Aproximación a la gramática española* (3.ª ed.). Madrid: Editorial Cincel.
- Marcos Marín, F., Satorre Grau, F. J. y Viejo Sánchez, M. L. (1998). *Gramática española*. Madrid: Editorial Síntesis.
- Martínez Díaz, E. (2003). La frecuencia del uso de "haber" y "tener" en las estructuras perifrásticas de obligación: algún fenómeno de variación en el español de Cataluña [en línea]. *Interlingüística*, 14, 681-694. Recuperado el 15 de septiembre de 2009 de <http://dialnet.unirioja.es/servlet/articulo?codigo=918806>
- Martínez Vilinsky, B. (2012). On the lower frequency of occurrence of Spanish verbal periphrases in translated texts as evidence for the unique items hypothesis. *Across Languages and Cultures*, 13 (2), 197–210. doi: 10.1556/Acr.13.2012.2.4
- Martínez Vilinsky, B. (2014). Analysis of the different features and functions of dialogue in a comparable corpus of crime novels. En S. M. Cadera y A. Pavić Pintarić (eds.), *The Voices of Suspense and Their Translation in Thrillers* (pp. 239-255). Amsterdam: Rodopi.

- Mason, I. y Şerban, A. (2003). Deixis as an interactive feature in literary translations from Romanian into English. [en línea] *Target* 15 (2), 269-294. doi: 10.1075/target.15.2.04mas
- Mauranen, A. (2000). Strange strings in translated language. A study on corpora. En M. Olohan (ed.), *Intercultural faultlines. Research models in translation studies* (Vol. I. Textual and cognitive aspects, pp. 119-141). Manchester: St. Jerome Publishing.
- Mauranen, A. (2004). Corpora, universals and interference. En A. Mauranen y P. Kujamäki (eds.), *Translation universals: Do they exist?* (pp. 65-82). Amsterdam: John Benjamins.
- Mauranen, A. (2005). Contrasting languages and varieties with translational corpora. *Languages in Contrast* 5 (1), 73-92. doi: 10.1075/lic.5.1.07mau
- Mauranen, A. (2007). Universal tendencies in translation. En G. Anderman y M. Rogers (eds.), *Incorporating corpora: The Linguist and the Translator* (pp. 32-48). Clevedon: Multilingual Matters.
- May, R. (1997). Sensible Elocution: How Translation Works in and upon Punctuation. *The Translator*, 3 (1), 1-20. doi: 10.1080/13556509.1997.10798985
- McEney, A. y Xiao, Z. (1999). Domains, text types, aspect marking and English-Chinese translation. *Languages in Contrast*, 2 (2), 211-229. doi: 10.1075/lic.2.2.05mce
- McEney, T. y Xiao, R. (2007). Parallel and Comparable Corpora: What is Happening? En G. Anderman y M. Rogers (eds.), *Incorporating Corpora: The Linguist and the Translator* (pp. 18-31). Clevedon : Multilingual Matters.
- McEney, T., Xiao, R. y Tono, Y. (2006). *Corpus-Based Language Studies: an advanced resource book*. London: Routledge.
- Meyer, C. F. (2002). *English Corpus Linguistics. An Introduction*. Cambridge: Cambridge University Press.
- Mikhailov, M. (2001). Two Approaches to Automated Text Aligning of Parallel Fiction Texts. *Across Languages and Cultures*, 2 (1), 87-96. doi: 10.1556/Acr.2.2001.1.6
- Milton, J. (1998). Exploiting L1 and interlanguage corpora in the design of an electronic language learning and production environment. En S. Granger (ed.). *Learner English on Computer* (pp. 186-198). Harlow: Longman.
- Montiel Ponsoda, E. (1999). La traducció de la dixi espacial. En Lluís Martos, J. (ed.) *La traducció del discurs* (pp. 53-77). Publicacions de la Universitat d'Alacant. Edición electrónica: Espagrafic. Recuperado el 16 de enero de 2013 de http://rua.ua.es/dspace/bitstream/10045/4380/1/Martos_S%C3%A1nchez_Traducci%C3%B3_d_el_discurs.pdf
- Mora, R. (30 de mayo de 2009). Maestros de la España negra [en línea] Recuperado el 27 de septiembre de 2015 de *elpais.com* http://elpais.com/diario/2009/05/30/babelia/1243641015_850215.html

- Moropa, K. (2011). A Link between Simplification and Explicitation in English-Xhosa Parallel Texts: Do the Morphological Complexities of Xhosa Have an Influence? En A. Kruger, K. Wallmach y J. Munday (eds.), *Corpus-based Translation Studies: Research and Applications* (pp. 259-831). London: Continuum.
- Mozetič, U. (2014). Explicitation and Implication as Translation Universals and Their Occurrences in the Slovene Translations of Anglo-American Literary Texts. *ELOPE*, 11 (2), 101-114. doi: 10.4312/elope.11.2.101-114
- Müller, F., y Waibel, B. (2009). *Corpus Linguistics - an introduction*. [en línea] Recuperado el 4 de agosto de 2009 de http://www.anglistik.unifreiburg.de/seminar/abteilungen/sprachwissenschaft/ls_mair/corpuslinguistics/corpus-linguistics-an-introduction#consider
- Náñez Fernández, E. (1973). *El Diminutivo: historia y funciones en el español clásico y moderno*. Madrid: Gredos.
- Nilsson, K. (1984). Observaciones sobre el uso de los adverbios deícticos del castellano (aquí, acá etc.) y sus formas correspondientes en catalán y portugués [en línea]. *Revue Romane*, 19 (2), 219-244. Recuperado el 13 de agosto de 2014 de http://img.kb.dk/tidsskrift.dk/pdf/rro/rro_0019-PDF/rro_0019_95019.pdf
- Nilsson, P. (2002). An assessment of the translation-specificity of over-represented multi-word patterns in Swedish fiction texts translated from English [en línea]. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 1, 407-417. Recuperado el 28 de agosto de 2015 de <https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/view/28>
- Nilsson, P. (2004). Translation-specific lexicogrammar? Characteristic lexical and collocational patterning in Swedish texts translated from English. En A. Mauranen y P. Kujamäki (eds.), *Translation universals: Do they exist?* (pp. 129-141). Amsterdam: John Benjamins.
- Oakes, M. P. (1998). *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Olohan, M. (2002). Corpus Linguistics and Translation Studies: Interaction and Reaction [en línea]. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 1, 419-429. Recuperado el 28 de agosto de 2015 de <https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/view/29>
- Olohan, M. (2003). How frequent are the contractions? A study of contracted forms in the Translational English Corpus. *Target* 15 (1), 59-89. doi: 10.1075/target.15.1.04olo
- Olohan, M. (2004). *Introducing corpora in translation studies*. New York: Routledge.
- Olohan, M. y Baker, M. (2000). Reporting *That* in Translated English. Evidence for Subconscious Processes of Explicitation? *Across Languages and Cultures* 1 (2), 141–158. doi: 10.1556/Acr.1.2000.2.1

- Øverås, L. (1998). In Search of the Third Code: An Investigation of Norms in Literary Translation. *Meta*, 43 (4), 557-570. doi: 10.7202/003775ar
- Paloposki, O. (2001). Enriching translations, simplified language? An alternative viewpoint to lexical simplification. *Target*, 13 (2), 265-288. doi: 10.1075/target.13.2.04pal
- Panek, L. L. (2003). Post-war American police fiction. En M. Priestman (ed.), *The Cambridge Companion to Crime Fiction* (pp. 155-171). Cambridge: Cambridge University Press.
- Pápai, V. (2004). Explicitation: A universal of translated text? En A. Mauranen y P. Kujamäki (eds.), *Translation universals: Do they exist?* (pp. 143-164). Amsterdam: John Benjamins.
- Perea Sardón, J. I. (2005). Etiquetado de textos y su aplicación a la traducción [en línea]. Investigación tutelada sin publicar, Universidad de Granada. Recuperada el 3 de agosto de 2009 de <http://www.petraweb.org/Perea%20Etiquetado%20de%20textos%20y%20su%20aplicacion%20a%20la%20traduccion.pdf>
- Pes, C. (2014-2015). Tutorial de XML [en línea]. Consultado el 16 de septiembre de 2015 en *Abrirollave.com* <http://www.abrirollave.com/xml/>
- Philip, G. (2009). Arriving at equivalence. Making a case for comparable corpora in Translation Studies. 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* (pp. 59-73). Amsterdam: John Benjamins.
- Porter, D. (2003). The Private Eye. En M. Priestman (ed.), *The Cambridge Companion to Crime Fiction* (pp. 95-113). Cambridge: Cambridge University Press.
- Priestman, M. (ed.) (2003). *The Cambridge Companion to Crime Fiction*. Cambridge: Cambridge University Press.
- Priestman, M. (2003). Introduction: crime fiction and detective fiction. En M. Priestman (ed.), *The Cambridge Companion to Crime Fiction* (pp. 1-6). Cambridge: Cambridge University Press.
- Prieto, V. M. (2005). *Spanish Evaluative Morphology: Pragmatic, Sociolinguistic, and Semantic Issues* [en línea]. Tesis doctoral, Universidad de Florida. Recuperada el 1 de marzo de 2015 de http://etd.fcla.edu/UF/UFE0010940/prieto_v.pdf
- Puurtinen, T. (1998). Syntax, readability and ideology in children's literature. *Meta*, 43 (4), 524-533. doi: 10.7202/003879ar
- Puurtinen, T. (2003). Genre-specific Features of Translationese? Linguistic Differences between Translated and Non-translated Finnish Children's Literature. *Literary and Linguistic Computing*, 18 (4), 389-406. doi: 10.1093/lc/18.4.389
- Puurtinen, T. (2004). Explication of clausal relations: A corpus-based analysis of clause connectives in translated and non-translated Finnish children's literature. En A. Mauranen y P. Kujamäki (eds.), *Translation universals: Do they exist?* (pp. 165-176). Amsterdam: John Benjamins.

- Pym, A. (2008). On Toury's laws of how translators translate. En A. Pym, M. Shlesinger y D. Simeoni (eds.), *Beyond Descriptive Translation Studies: Investigations in Homage to Gideon Toury* (pp. 311-328). Amsterdam: John Benjamins.
- Rabadán, R. (2005-2008). Tools for English-Spanish Cross-Linguistic Applied Research [en línea]. *Journal of English Studies*, 5-6, 309-324. Recuperado el 21 de febrero de 2010 de <https://publicaciones.unirioja.es/ojs/index.php/jes/article/view/135>
- Rabadán, R. (2006). Modality and modal verbs in contrast: Mapping out a translation(ally) relevant approach English-Spanish. *Languages in Contrast* 6 (2), 261–306. doi: 10.1075/lic.6.2.04rab
- Rabadán, R. (2007). Translating the Predictive and Hypothetical Meanings English-Spanish. *Meta*, 52 (3), 484-502. doi: 10.7202/016734ar
- Rabadán, R., Labrador, B. y Ramón, N. (2009). Corpus-based contrastive analysis and translation universals: A tool for translation quality assessment English → Spanish. *Babel*, 55 (4), 303-328. doi: 10.1075/babel.55.4.01rab
- Rainer, F. (2010) Sobre polisemia en la formación de palabras [en línea]. *Hesperia. Anuario de filología hispánica*, 13 (2), pp. 7-52. Recuperado el 5 de agosto de 2015 de <http://hesperia.webs.uvigo.es/paginas/indices/articulos/volXIII-2/Rainer.pdf>
- Ramón, N. (2002). Contrastive Linguistics and Translation Studies Interconnected: The Corpus-based Approach [en línea]. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 1, 393-406. Recuperado el 28 de agosto de 2015 de <https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/view/27>
- Ramón, N. y Labrador, B. (2008). Translations of '-ly' adverbs of degree in an English-Spanish Parallel Corpus. *Target* 20 (2), 275-296. doi: 10.1075/target.20.2.05ram
- Rayson, P. (1993-2010). *Log-likelihood Calculator*. UCREL Lancaster University. Consultado el 20 de noviembre de 2012 en <http://ucrel.lancs.ac.uk/llwizard.html>
- Rayson, P. (1993-2014). *Log-likelihood and effect size calculator*. UCREL Lancaster University. Consultado el 7 de octubre de 2015 en <http://ucrel.lancs.ac.uk/llwizard.html>
- Rayson, P., Berridge, D. y Francis, B. (2004). Extending the Cochran Rule for the Comparison of Word Frequencies between Corpora [en línea]. En G. Purnelle, C. Fairon y A. Dister (eds.), *Le poids des mots: Proceedings of the 7th International Conference on Statistical Analysis of Textual Data (JADT 2004)*, Louvain-la-Neuve, Belgium, March 10-12. Presses universitaires de Louvain, 926-936. Recuperado el 13 de diciembre de 2011 de http://ucrel.lancs.ac.uk/people/paul/publications/rbf04_jadt.pdf
- Real Academia Española y Asociación de Academias de la Lengua Española (2009). *Nueva gramática de la lengua española. Morfología y Sintaxis*. Vol. 1. Madrid: Espasa Libros.
- Reddy, M. T. (2003). Women detectives. En M. Priestman (ed.), *The Cambridge Companion to Crime Fiction* (pp. 191-207). Cambridge: Cambridge University Press.

- Rich, N. (2009). James Ellroy, *The Art of Fiction* No. 201 [en línea]. En *The Paris Review*, 190. Recuperado el 30 de septiembre de 2015 de <http://www.theparisreview.org/back-issues/year/2009>
- Richardson, B. (1998). Deictic Features and the Translator. En L. Hickey (ed.) *The Pragmatics of Translation* (Topics in Translation, vol. 12). Clevedon: Multilingual Matters.
- Rojo López, A. M. (2009). *Step by Step. A Course in Contrastive Linguistics and Translation*. Bern: Peter Lang.
- Römer, U. y Wulff, S. (2010). Applying Corpus Methods to Written Academic Texts: Explorations of MICUSP. *Journal of Writing Research*, 2 (2), 99-127. doi: 10.17239/jowr-2010.02.02.2
- Rosa, A. A. (2008). Narrator profile in translation: Work-in-progress for a semiautomatic analysis of narratorial dialogistic and attitudinal positioning in translated fiction [en línea]. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 7, 227-248. Recuperado el 28 de agosto de 2015 de <https://lans-tts.uantwerpen.be/index.php/LANS-TTS/article/view/217>
- Rubio Moraiz, P. (2008). *Verbos españoles conjugados*. Madrid: Sociedad General Española de Librería.
- Saldanha, G. (2009). Principles of corpus linguistics and their application to translation studies research [en línea]. *Revista Tradumática*, 7. Recuperado el 6 de marzo de 2010 de <http://www.fti.uab.cat/tradumatica/revista/num7/articles/01/01.pdf>
- Saldanha, G. (2011). Style of Translation: The Use of Foreign Words in Translation by Margaret Jull Costa and Peter Bush. En A. Kruger, K. Wallmach y J. Munday (eds.), *Corpus-Based Translation Studies: Research and Applications* (pp. 237-258). London: Continuum.
- Sánchez Benedito, F. (1990). *Gramática inglesa*. Madrid: Alhambra.
- Sánchez Zapatero, J. y Martín Escribà, À. (2010). Teoría e historia de las sagas policíacas en la literatura española contemporánea (1972-2007) [en línea]. *Dicenda. Cuadernos de Filología Hispánica*, 28, 289-305. Recuperado el 10 de septiembre de 2015 de <http://revistas.ucm.es/index.php/DICE/article/view/DICE1010110289A/11434>
- Santibáñez Sáenz, F. (1999). Conceptual Interaction and Spanish Diminutives [en línea]. *Cuadernos de Investigación Filológica*, 25, 173-190. Recuperado el 1 de marzo de 2015 de <https://publicaciones.unirioja.es/ojs/index.php/cif/article/view/2247>
- Santorini, B. (1990). Part-of-Speech Tagging Guidelines for the Penn Treebank Project (3rd Revision) [en línea]. *Technical Reports (CIS)*, University of Pennsylvania, paper 570. Recuperado el 2 de marzo de 2013 de http://repository.upenn.edu/cis_reports/570
- Santos, D. y Frankenberg-García, A. (2007). The corpus, its users and their needs. A user-oriented evaluation of COMPARA. *International Journal of Corpus Linguistics* 12 (3), 335-374. doi: 10.1075/ijcl.12.3.03san
- Santos Muñoz, A. (2006). Una propuesta didáctica para la enseñanza / aprendizaje de los sufijos apreciativos en E/LE [en línea]. Memoria de máster. *Biblioteca virtual redELE*, 8. Recuperado

- el 11 de julio de 2015 de http://www.mecd.gob.es/dctm/redele/Material-RedEle/Biblioteca/2007_BV_08/2007_BV_08_21Santos.pdf?documentId=0901e72b80e2d98b
- Scaggs, J. (2005). *Crime Fiction*. London: Routledge.
- Shlesinger, M. (1991). Interpreter Latitude vs. Due Process: Simultaneous and Consecutive Interpretation in Multilingual Trials. En S. Tirkkonen-Condit (ed.), *Empirical Research in Translation and Intercultural Studies: Selected Papers of the TRANSIF Seminar, Savonlinna, 1988* (pp. 147-155). Tübingen: Gunter Narr Verlag.
- Schmid, H. (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees [en línea]. *Proceedings of International Conference on New Methods in Language Processing*, Manchester, UK. Recuperado el 18 de octubre de 2015 de <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger1.pdf>
- Schmid, H. (1995). Improvements in Part-of-Speech Tagging with an Application to German [en línea]. *Proceedings of the ACL SIGDAT-Workshop*. Dublin, Ireland. Recuperado el 18 de octubre de 2015 de <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger2.pdf>
- Schneider, K. P. (2003). *Diminutives in English* [en línea]. Tübingen: Niemeyer. Consultado el 31 de julio de 2015 en <https://books.google.es/books?id=xVwgAAAQBAJ&lpg=PP1&dq=isbn%3A3110929554&hl=es&pg=PP1#v=onepage&q&f=false>
- Seago, K. (2014). Introduction and overview: crime (fiction) in translation [en línea]. *JoSTrans*, 22, 2-14. Recuperado el 4 de septiembre de 2015 de http://www.jostrans.org/issue22/art_seago.php
- Seco, M. (1989). *Gramática esencial del español: introducción al estudio de la lengua* (2.^a ed.). Madrid: Espasa-Calpe.
- Selva, J. B. (1949). Sufijos americanos [en línea]. *Thesaurus*, 5 (1, 2 y 3), 192-213. Recuperado el 12 de agosto de 2015 de *Centro Virtual Cervantes* http://cvc.cervantes.es/lengua/thesaurus/pdf/05/TH_05_123_202_0.pdf
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2003). *Reading Concordances: An Introduction*. Harlow: Pearson Education.
- Sinclair, J. et ál. (ed.) (1996). *Collins Cobuild English Grammar*. London: HarperCollins.
- Solana, T. (28 de noviembre de 2011). Hernán Sabaté, la voz de James Ellroy en español [en línea]. Recuperado el 27 de septiembre de 2015 de *elpais.com* http://elpais.com/diario/2011/11/28/necrologicas/1322434803_850215.html
- Soler Espiauba, D. (1996). Españolitos, curritos, amiguetes, coleguillas y demás gente de mal vivir (o la expresividad en el sistema español de sufijación) [en línea]. *Cuadernos Cervantes de la Lengua Española*, 8, 28-34. Recuperado el 5 de agosto de 2015 de J. Feernández López, *Hispanoteca* <http://hispanoteca.windows2.webhome.at/Wortbildung/XIV.htm>

- Stangroom, J. (2015). Fisher Exact Test Calculator [en línea]. Consultada el 7 de octubre de 2015 en *Social Science Statistics* <http://www.socscistatistics.com/tests/fisher/default2.aspx>
- Stefanowitsch, A. (2005). *Quantitative Thinking for Corpus Linguistics* [en línea]. Consultado el 10 de marzo de 2013 en http://www-user.uni-bremen.de/~anatol/qnt/qnt_dist.html
- Stewart, L. (8 de marzo de 1996). Packer Has Backers, Some Detractors after His Iverson Remark [en línea]. Recuperado el 13 de agosto de 2015 de *Los Angeles Times* http://articles.latimes.com/1996-03-08/sports/sp-44682_1_billy-packer
- Stockwell, R. P., Bowen, J. D. y Martin, J. W. (1965). *The Grammatical Structures of English and Spanish*. Chicago: The University of Chicago Press.
- Stubbs, M. (1993). British Traditions in Text Analysis, from Firth to Sinclair. En M. Baker, G. Francis y E. Tognini-Bonelli, *Text and Technology: In Honour of John Sinclair* (pp. 1-34). Amsterdam: John Benjamins.
- Taylor, C. (2008). What is *corpus linguistics*? What the data says [en línea]. *ICAME Journal*, 32, 179-200. Recuperado el 10 de junio de 2015 <http://clu.uni.no/icame/ij32/index.html>
- Teich, E. (2003). *Cross-Linguistic Variation in System and Text: A Methodology for the Investigation of Translations and Comparable Texts*. Berlin: Mouton de Gruyter.
- Teng, M. (2012). Using a Hybrid Methodology in the Analysis of the Topic and Non-Topic Markers *Wa* and *Ga* in Chinese-to-Japanese Translation. *Across Languages and Cultures*, 13 (2), 247-261. doi: 10.1556/Acr.13.2012.2.7
- Teubert, W. (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics*, 10 (1), 1-13. doi: 10.1075/ijcl.10.1.01teu
- Thomson, A. J. y Martinet, A. V. (1987). *A Practical English Grammar* (4.^a ed.). Oxford: Oxford University Press.
- Tirkkonen-Condit, S. (2002a). Process research: state of the art and where to go next. *Across Languages and Cultures*, 3, 5-19. doi: 10.1556/Acr.3.2002.1.1
- Tirkkonen-Condit, S. (2002b). Translationese – A Myth or an Empirical Fact? A Study into the linguistic Identifiability of Translated Language. *Target*, 14 (2), 207-220. doi: 10.1075/target.14.2.02tir
- Tirkkonen-Condit, S. (2004). Unique items — over- or under-represented in translated language? En A. Mauranen y P. Kujamäki (eds.), *Translation universals: Do they exist?* (pp. 177-184). Amsterdam: John Benjamins.
- Tirkkonen-Condit, S. (2005). Do Unique Items Make Themselves Scarce in Translated Finnish? En K. Károly y Á. Fóris (eds.), *New Trends in Translation Studies: In Honour of Kinga Klaudy* (pp. 177-189). Budapest: Akadémiai Kiadó.
- Tirkkonen-Condit, S. y Mäkisalo, J. (2007). Cohesion in Subtitles: A Corpus-Based Study. *Across Languages and Cultures*, 8 (2), 221-230. doi: 10.1556/Acr.8.2007.2.4
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.

- Toury, G. 1977. *Translational Norms and Literary Translation into Hebrew, 1930-1945*. Tel-Aviv: The Porter Institute for Poetics and Semiotics, Tel Aviv University [en hebreo].
- Toury, G. (1995). *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins.
- Toury, G. (2004a). Probabilistic explanations in Translation Studies. Universals – or a challenge to the very concept? En G. Hansen, K. Malmkjær y D. Gile (eds.), *Claims, changes and challenges in translation studies* (pp. 15-25). Amsterdam: John Benjamins.
- Toury, G. (2004b). Probabilistic explanations in translation studies: welcome as they are, would they qualify as universals? En A. Mauranen y P. Kujamäki (eds.), *Translation universals: do they exist?* (pp. 15-32). Amsterdam: John Benjamins.
- Tymoczko, M. (1998). Computerized corpora and the future of translation studies. *Meta*, 43 (4), 652-659. doi: 10.7202/004515ar
- Tymoczko, M. (2005). Trajectories of Research in Translation Studies. *Meta*, 50 (4), 1082-1097. doi: 10.7202/012062ar
- Ulašín, B. (2006). *Registro coloquial del español peninsular contemporáneo* [en línea]. Bratislava: Univerzita Komenského. Recuperado el 6 de agosto de 2015 de https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/krom/dizertacka_bohdan.pdf
- Utka, A. (2004). Phases of translation corpus: Compilation and analysis. *International Journal of Corpus Linguistics*, 9 (2), 195-224. doi: 10.1075/ijcl.9.2.03utk
- Valenzuela, J. (8 de noviembre de 2008). La negra vocación de Juan Madrid [en línea]. Recuperado el 27 de septiembre de 2015 de http://elpais.com/diario/2008/11/08/babelia/1226105411_850215.html
- Vanderauwera, R. (1985). *Dutch Novels Translated into English. The Transformation of a "Minority" Literature*. Amsterdam: Rodopi.
- Villayandre Llamazares, M. (2010). El desarrollo de un corpus (ii): codificación y anotación [en línea]. En *Lingüística Computacional II. Curso monográfico sobre Lingüística de corpus* (pp. 45-68). León: Universidad de León. Recuperado el 10 de junio de 2015 de <http://fhyc.unileon.es/Milka/LCII/Corpus6.pdf>
- Volansky, V., Ordan, N. y Wintner, S. (2015). On the features of translationese. *Digital Scholarship in the Humanities*, 30 (1), 98-118. doi:10.1093/llc/fqt031
- Whittaker, S. (2004). Etude contrastive des syntagmes nominaux démonstratifs dans des textes traduits du français en norvégien et des textes sources norvégiens: stratégie de traduction ou translationese? *FORUM - International Journal of Interpretation and Translation*: 2 (2), 221-240.
- Wilkinson, M. (2011). WordSmith Tools: The best corpus analysis program for translators? [en línea]. *Translation Journal*, 15 (3). Recuperado el 23 de marzo de 2013 de <http://translationjournal.net/journal/57corpus.htm>

- Xiao, R. y Dai, G. (2014). Lexical and grammatical properties of Translational Chinese: Translation universal hypotheses reevaluated from the Chinese perspective. *Corpus Linguistics and Linguistic Theory*, 10 (1), 11-55. doi: 10.1515/cllt-2013-0016
- Yllera, A. (1999). Las perífrasis verbales de gerundio y participio. En I. Bosque y V. Demonte (eds.), *Gramática descriptiva de la lengua española* (Vol. 2, pp. 3392-3441). Madrid: Espasa Calpe.
- Zacarías Ponce de León, R. F. (2008). Morfemas apreciativos del español: entre la flexión y la derivación [en línea]. *Núcleo* 25, 221-237. Recuperado el 1 de marzo de 2015 de http://saber.ucv.ve/ojs/index.php/rev_n/article/view/5170
- Zanettin, F. (2000). Parallel corpora in translation studies: issues in corpus design and analysis. En M. Olohan (ed.), *Intercultural Faultlines. Research Models in Translation Studies* (Vol. I. Textual and cognitive aspects, pp. 105-118). Manchester: St Jerome Publishing.
- Zanettin, F. (2012). *Translation-Driven Corpora*. Manchester: St. Jerome Publishing.
- Zanettin, F. (2013). Corpus Methods for Descriptive Translation Studies. *Procedia - Social and Behavioral Sciences*, 95, 20-32. doi: 10.1016/j.sbspro.2013.10.618
- Zubillaga, N., Sanz, Z. y Uribarri, I. (2015). Building a trilingual parallel corpus to analyse literary translations from German into Basque [en línea]. En C. Fantinuoli, y F. Zanettin (eds.), *New directions in corpus-based translation studies* (pp. 71-92). Berlin: Language Science Press. Recuperado el 28 de agosto de 2015 de <http://langsci-press.org/catalog/book/76>

Introduction and Conclusions (as required for the International Doctorate Mention)

1. Introduction

1.1. Background and research question

Are translations special? Is there something in the nature of a translated text that characterises it as such? Although many researchers have attempted to answer this question throughout the years, it has become particularly relevant during the last couple of decades. Corpus-based Translation Studies has turned it into its main research question. Comparing source texts and translations in order to make evaluative statements about translation solutions is no longer the main goal of descriptive studies. The focus of attention has shifted to a search for the aspects in which translated texts differ from texts produced spontaneously in the same language without the stimulus of a source text; what researchers are looking for are features which distinguish translated texts and allow us to identify them as such, trying not to judge them as correct or incorrect.

This idea had been put forward as early as the 1970s and 1980s by authors such as Toury (1977) or Frawley (1984), but became widespread in the mid-nineties, after the publication of a series of articles by Mona Baker (1993, 1995, 1996). These articles proved extremely influential and laid the foundations for what we now know as Corpus-based Translation Studies. In them, Baker claimed that translation, as a process with its own particular goals, strains and contexts of production, possesses certain intrinsic features which occur independently from situational variables such as language pair or translation brief. She called these features *translation universals* and argued that the best way of uncovering them was by means of the tools and techniques that Corpus Linguistics had been using for some time. Her idea was therefore to use the

methodology of Corpus Linguistics to study the special nature of translation, which, in her opinion, would manifest itself in linguistic regularities or patterns specific to translated texts and different from those found in original texts written in the same language.

Drawing on certain studies which had analysed the nature of translated language, although on a smaller scale than is nowadays possible through the use of large corpora of text, Baker suggested some translation universal candidates. Some of the best known candidates are the tendency of translated text to add information in order to make the message clearer (explicitation), to code information in a simpler way than the source text (simplification), to use target-language less creative or more conventional structures (normalisation), and to avoid extremes by gravitating around the middle ground (levelling out). Since then, many researchers have tried to confirm or refute the existence of these so-called universals through corpus-based studies. Evidence for and against them has been found; some candidates have received more attention than others and some have been shown to have a less solid base than was initially thought.

New candidates have also been put forward by other researchers. This study will focus on one of these new candidates: the under-representation of target-language unique items in translated texts, also known as the Unique Items Hypothesis. This hypothesis originated in the work of Scandinavian authors such as Mauranen (2000), Eskola (2004) and Tirkkonen-Condit (2002). In her 2004 study, Tirkkonen-Condit found that certain typical features of Finnish were less frequent in translated texts than in non-translated ones. She called these features *unique items* and defined them as linguistic elements with no straightforward equivalents in the source language, which are therefore not easily activated when translating. She believed that certain items which are unique or typical of a particular language, since they are not represented in a similar way in other languages, find no obvious stimulus in the source texts that can trigger their use. Consequently, they are unlikely to occur in translations, even though they would probably be used in spontaneous text production.

The Unique Items Hypothesis has been supported mainly by Finnish and Swedish researchers. The question that we wish to answer in this study is based on their conclusions, since what we intend to find out is whether this phenomenon can also be observed for the English-Spanish language pair. Our initial hypothesis can thus be formulated as follows: *In texts translated from English into Spanish, unique or typical*

elements of the target language tend to be under-represented when they lack an obvious counterpart in the source language.

1.2. Objectives and methodology

Baker claimed that in order to prove that a feature is in fact universal, empirical studies need to be carried out in several language pairs and with texts from different genres and periods. Consequently, our aim is to find out whether this under-representation of certain target-language unique items or structures in translations (under-represented when compared to non translated texts in the same language), confirmed by Tirkkonen-Condit for academic and fiction texts translated from English into Finnish, can also be observed in literary texts translated from English into Spanish, and whether it is thus safe to consider this phenomenon a translation universal candidate. We therefore intend to replicate previous studies changing variables such as the language pair, the characteristics of the corpus and the elements analysed, and see if our results confirm the Unique Items Hypothesis.

The methodology used in this work is that which is standard in corpus-based translation research – the study of the linguistic features found in big quantities of text compiled in different kinds of electronic corpora (in our case, comparable and parallel) with the aid of automatic data extraction and lexical analysis tools. The frequencies of occurrence of the selected items in natural text production in Spanish are established through a subcorpus of Spanish non-translations (ES); these frequencies are then compared to those found for the same items in a subcorpus of translations from English into Spanish (TRAD), in order to determine if they are indeed under-represented in translated language. The English source texts corresponding to the subcorpus of translations constitute our third subcorpus (EN). Through the extraction of parallel concordances from EN and TRAD it is possible to look for correlations and potential explanations to the phenomena observed in translations.

Our first practical objective is therefore to compile an electronic comparable corpus (comparable in terms of genre, function, linguistic variety and other criteria established beforehand) of contemporary Spanish fiction (crime fiction from the nineties), written originally in Spanish and translated from English into Spanish, and a parallel component of source texts in English.

Secondly, it is essential to find, through consultation of different monolingual grammars and contrastive studies, an element or structure in Spanish that fits the definition of *unique item*, namely, an element of a lexical, phraseological or grammatical nature without a ready-made equivalent in the source language; a feature which can be considered a singularity of the Spanish language because other languages must code its meaning in a different way, through a paraphrase or other linguistic means.

Our third practical objective is the analysis of our comparable corpus in relation to the presumed unique items. Once their frequency of use in original texts is established by means of the semiautomatic extraction of concordances with a lexical analysis program, we can find out whether they are used differently in the subcorpus of translations and, if this is the case, whether the differences can be considered statistically significant. If the frequencies are significantly lower in the subcorpus of translations then it can be safely stated that these elements are indeed under-represented.

If a phenomenon of under-representation is found in the subcorpus of translations, the next step is to find out whether it is due to the influence of the source language or to the fact that we are dealing with a feature inherent in the translation process. The subcorpus of English originals allows us to compare the source texts with their translations, and through a careful analysis of the parallel concordances extracted from the aligned subcorpora it is possible to determine whether the observed difference in frequencies and use is due to the lack of parallel structures in the source language.

The quantitative analysis of frequencies is thus complemented by a qualitative analysis, since the results yielded by the software are sifted manually, and the systematic study of patterns of use of the unique elements in both languages can provide us with clues about possible correlations or potentially causal relations that could explain the phenomena observed and described in our work. This information will allow us to answer the following research questions: Are there differences between the frequencies of use of certain Spanish unique elements in translated texts and non-translated texts? Are these differences significant? Are they due to a lack of triggering elements in English? Are they a consequence of the structural differences between the two languages or a feature inherent in the translation process?

As will be explained in chapter 2, when we explore the state of the art of Corpus-based Translation Studies, authors like Halverson (2003) or Hareide (2013) have related the phenomenon of under-representation with another hypothesis called

Gravitational Pull Hypothesis, which predicts that certain cognitive factors will exert a stronger or weaker influence on the selection of a specific linguistic item by the translator. A secondary objective of this thesis is therefore to try to identify, on the basis of the conclusions of our analysis and only to the extent possible in a study of this kind, the role that these factors play in the under-representation of unique items. Finally, we will see which of those features included in Hareide's definition of a unique item are confirmed by our results.

1.3. Structure of the thesis

To achieve these objectives it is necessary, first of all, to analyse the origin of Corpus-based Translation Studies and to review the state of the art in the field. Secondly, the concept of translation universal, which is surrounded by much controversy, needs to be defined and explained, and it should be determined if its use is beneficial or superfluous to translation theory. We devote **chapter 2** of this thesis, "Theoretical framework and state of the art", to these issues and to a review of the different studies carried out on the subject of the under-representation of unique items in translations.

The compilation of a comparable and parallel corpus of literary texts in the language pair English-Spanish is described in **chapter 3** ("Methodology"). In this section, as well as in **Appendix 1**, we describe in detail the stages of designing the corpus and of selecting, compiling and pre-processing the texts included in it, the software and techniques used to carry out the analysis, and the difficulties that arose throughout the whole process.

Chapters 4, 5 and 6 are dedicated to the analysis of the three unique items selected for this study. At the beginning of each chapter a review of the main literature on the linguistic element in question is presented and a justification of its selection as a unique item is given. After the results section of each chapter we draw some preliminary conclusions.

In **chapter 4** we explore a group of 27 verbal periphrases which, according to several authors, are a singularity of the Spanish language, both because their number is much larger than in other languages such as English and because of the way they

qualify the verbal conjugation. **Chapter 5** studies the tripartite structure of the system of demonstrative determiners and pronouns in Spanish (*este, ese* and *aquel*, and all their inflected and gender-neutral forms) and the locative adverbs *aquí, ahí* and *allí* and alternative forms, as opposed to the formal dichotomic division of English (*this* and *that; here* and *there*). In **chapter 6** we analyse a group of appreciative diminutive, augmentative and pejorative suffixes, a morphological device used to express quantification or affective evaluation which is much less frequent in English than in Spanish.

Chapter 7 brings together the partial conclusions of the previous three chapters and tries to answer the research questions posed in this introduction in order to determine whether it is possible to validate the Unique Items Hypothesis with the data gathered in this study.

1.4. Rationale and significance of the study

What we ultimately want to find out is if it is true that certain elements peculiar to Spanish are used less frequently in translated texts than in non-translated texts in the same language and, if so, if this is due to a lack of readily available equivalent elements in the source language or if it depends on the type of element one is dealing with or on other factors. The translation universal candidate under scrutiny here is particularly interesting, given that so far it has received little attention, let alone in Spanish. Most studies have been carried out with Finnish – a language which has no direct relationship with most other European languages – as a target language, and even though positive results have been obtained with French (Cappelle 2012) and Chinese (Xiao and Dai 2014), the only work with which we are familiar to this day that studies the Unique Items Hypothesis as one of its main objectives (Hareide 2013) finds no evidence to support it. With this thesis we wish to open new avenues of research on this subject.

Studying the possible existence of features shared by all translated texts is especially interesting and relevant because of the methodological and pedagogical consequences its corroboration – or otherwise – may bring about. Through the acquisition of wider knowledge about those features that may be inherent in the process

of transferring a message from one language to another we will become more conscious of what we do and refrain from doing when translating, of our resources and our limitations, and we will thus be able to learn and teach in more efficient ways. Likewise, if we discover that translations have features of their own, different from those that can be found in other communicative situations, we will be able to continue building our discipline on the basis of data of real language use in the context of this particular activity.

If the under-representation hypothesis were not to be confirmed after all, the analysis of a corpus of translations such as the one we have built for this study can yield very interesting data regarding other phenomena, universal or otherwise, present in the texts, and contribute to a better knowledge of these phenomena and of the selected linguistic items. Through the analysis of the parallel concordances, it is also possible to obtain valuable information about the various translation equivalents of these items, which may be of use in contrastive studies and in the area of translation teaching.

7. General conclusions

The main objective of this study was to extend the results of prior research done on the subject of under-representation of unique items in translated language, a hypothesis that suggests that elements which are specific to a certain language and lack a direct or easily identifiable triggering element in the source language tend to appear less frequently in texts translated to that language than in texts written originally in it. This feature of translated text has been put forward as a translation universal candidate, since it is believed to be inherent to the translation process and not a result of the specific factors influencing it. By carrying out a study similar in method to those by, for example, Tirkkonen-Condit (2004) or Mauranen (2000), but on a different set of texts, with a different language combination and different indicators, the intention was to alter external factors – those that depend on situational variables rather than on the

process itself of mediating between languages – to find out whether results remained similar and under-representation could be considered a general feature of translated text.

In addition, the study carried out by Hareide (2013) to put Halverson's (2003) Gravitational Pull Hypothesis to the test left several questions unanswered regarding the definition of unique item and the conditions that favour an under-representation of these items in translations. Hareide's results contradicted the Unique Items Hypothesis even in those cases in which there were no clear equivalents in the source language for the elements analysed (which were incidentally very frequent in the target language): "My thesis thereby indicates that very frequent or prototypical target-language features appear to be over-represented in translations in accordance with Baker's (1993) hypothesis, even when they constitute unique items" (2003: 275). Consequently, the author asks herself what it would take for a supposedly unique item to be under-represented in translations. A secondary objective of this thesis was therefore to identify the factors that made this under-representation possible whenever the main hypothesis was confirmed.

7.1. Corpus compilation

We began our research with the compilation of a corpus of contemporary crime novels made up of texts written originally in Peninsular Spanish and translations from American English into the same variety of Spanish, complemented by a parallel component of these novels' source texts in English. The process of creating and analysing this corpus enabled us to reflect on the advantages and limitations of corpus-based methodology, which nowadays offers the researcher unprecedented possibilities, thanks to the enormous amount of data that can be managed, in an increasingly refined way, by textual analysis tools. Nevertheless, these tools require a thorough understanding of their functioning and limitations in order to be able to fully exploit their capacities.

As we have seen time and again, concordancers will only show what the researcher has asked for and will base their searches solely on the information that has been provided to them. This means that factors such as the initial selection of texts or the level of pre-processing of those texts will influence the answers that we get through

our searches. A very clear example of this in our work was the realization that the tagging of the texts can certainly affect the results, not just because of the tagging errors of the program itself, which modified the total count of items, but also because of our lack of knowledge about the way the program works. In the case of opening exclamation and interrogation marks, the fact that we were not aware that they formed a unit with the word they preceded meant that these units did not appear in our searches. Last but not least, one of the most widely acknowledged limitations of corpus-based studies is the fact that corpora will only answer the questions posed by the researcher, which means that the results will always be subjective. Nevertheless, through our study we have realised that, more often than not, the initial questions will lead to new questions and the results will show things we had not anticipated. We thus believe that every corpus-based study eventually turns into a corpus-driven one.

7.2. Selection of items and results

The compilation of our corpus has permitted us to compare the frequencies of use of a set of linguistic elements in translated and non-translated Spanish. Through these elements we have operationalised Tirkkonen-Condit's hypothesis in Spanish. It has also allowed us to extract parallel concordances of these items in translations and their source texts in order to find out to what extent the source text influences the translator's linguistic decisions. Through the consultation of grammars and specialised works and contrastive studies, we have tried to identify a set of elements which can be considered typical of Spanish and unique in the sense that they lack a *prima facie* equivalent in English, either because there is no element in that language that can translate their meaning (and so this meaning is transferred through different linguistic structures) or because that element, if it exists, is so formally different to our unique item that it would probably lead the translator to other alternatives in Spanish, more similar to the English structure.

The chosen items were a group of 27 infinitive and gerund verbal periphrases divided into three groups of nine periphrases each, according to our expectations of finding them under-represented or not in the subcorpus of translations (chapter 4), the entire system of Spanish demonstrative pronouns, determiners and locative adverbs

(chapter 5), and a group of appreciative suffixes of diminutive, augmentative and pejorative meaning (chapter 6)²²⁸. The results of the analysis were the following:

- ◆ **Verbal periphrases:** The entire set of periphrases, as a whole, turned out to be under-represented in translations both in the first and the second version of our corpus, a fact that we had not anticipated. This under-representation was less evident in the second version, since neither the first nor the third group showed significant differences between subcorpora. This confirms the need to work with corpora of similar size, so that the normalisation of frequencies affects the real figures as little as possible, and also the effect that individual texts can have on the behaviour of specific elements (we are referring here to those periphrases which were under-represented in the first version of the corpus and became over-represented in the second due to a higher than average use by a particular author or translator). This suggests, too, the need to build bigger corpora so that the influence of idiosyncratic elements gets diluted.

To sum up, we can say that our initial hypothesis was verified, since group II, made up of periphrases with no direct equivalents in the source language, was the only one that showed a significant under-representation in both versions of the corpus. However, if we dig deeper, four members of this group did not confirm the hypothesis, while some of the members of the other two groups did (surprisingly, most of these cases are from group I, periphrases with supposedly straightforward equivalents). The total number of under-represented periphrases was eleven out of twenty-seven: five out of nine in groups I and II and only one in group III. As can be inferred from these figures, our initial predictions about the under-representation of certain periphrases due to the lack of direct counterparts in English turned out not to be very accurate. This means that there are other factors at play, apart from source-text or source-language influence, affecting the behavior of these elements.

Under-representation in the group of periphrases with what we considered straightforward equivalents resulted from the lack of clear links between them and a single element in the source language, probably due to the

²²⁸ They were 27 in principle, but, as we mentioned in that chapter, the final number depends on our considering the elements analysed jointly as different suffixes or mere variants. After grouping some of these forms for the analysis, the number of rows in the table of results was 23.

existence of other alternatives to express the same meanings in Spanish or English. Over-representation of periphrases without an obvious counterpart could be caused by the fact that some of them are the most usual option in Spanish to express certain meanings. This phenomenon was also observed if the equivalent element in English was in fact clear or very salient, or if a wide variety of English structures could trigger the use of a specific periphrasis, or, in some cases, because the periphrasis in question was rarely used in non-translated Spanish, so the existence of even just a small number of examples in TRAD got amplified by the lack of examples in ES. The last factor we detected was of course the translator's preference for a specific linguistic resource.

- **Demonstratives:** In the case of locative adverbs, our study yields similar results to those obtained by Hareide in her work on the Spanish gerund, since there is a general over-representation of these elements in translations. Pronouns and determiners, on the contrary, are under-represented, both the group as a whole and each individual basic form. Nevertheless, it cannot be claimed that our initial hypothesis was verified, since what we actually expected to find was an under-representation of only one of the two Spanish distal forms. There is, however, a bigger difference in frequencies and a lower proportional use with respect to the other two forms in the case of *aquel*, in addition to a more idiosyncratic use of this demonstrative by the translators. The third degree of deixis seems to be closer to the concept of unique item that we were looking for.
- **Appreciative suffixes:** Ten out of twenty-three suffixes or groups of suffixes turned out to be under-represented in TRAD. Nevertheless, no element showed a significantly higher frequency in translations, and the under-represented items (mostly diminutive and augmentative suffixes) were the ones with higher frequencies of use. The group of suffixes as a whole is therefore clearly under-represented in translations. There is barely any structure in English that can be considered formally equivalent to the analysed suffixes: their use seems to depend on the will of the individual translator, perhaps to a greater extent than in the previous two cases.

7.3. Definition of a *unique item*

As was stated in chapter 2, Hareide proposes a revised definition of unique item, since the results of her study do not confirm Tirkkonen-Condit's hypothesis, even though she analyses an element with no straightforward equivalent in the source language:

A unique item is a target language feature that has no direct counterpart in the source language, and has no salient or prototypical source language structure that may impact its choice in the target language. It is characterized by being infrequent and non-prototypical. (2013: 276)

We thus wanted to evaluate to what extent these conditions are met by the elements that turned out to be effectively under-represented in our corpus. This section has to be regarded as a very tentative exercise, since, as we have explained previously, a thorough study of the prototypicality of these elements in both English and Spanish falls outside the scope of this thesis. We have based these preliminary conclusions on what the different grammars and studies consulted say about them and on the frequencies found in our corpus. In what follows we will try to elucidate, as Hareide suggests, if unique items should be elements with a very low frequency of use, or if they should not have prototypical status in the target language, or a salient equivalent in the source language, or if they should perhaps lack any clear connection with source-language related concepts in the translator's bilingual lexicon.

We will also sum up what we have found in our analysis regarding the influence of the three factors that can lead to under- or over-representation of these items in the target language according to the Gravitational Pull Hypothesis. These were a) the frequency or prototypicality of the item in the target language (the higher its frequency or prototypicality, the higher the probability that it is over-represented in translations); b) the salience or prototypicality of a structure of the source-language conceptual network that may influence the translator to use the target-language item; and c) the patterns of connectivity between related concepts in the two languages in the translator's mental lexicon (the weaker the connection, the higher the probability of the target-language item to be under-represented in translations).

◆ **Verbal periphrases:**

According to the results of the second version of the corpus, a periphrasis may be under-represented in translations either because it does not have prototypical status in the target language or it is not frequently used due to the existence of other elements in that language with similar meanings and functions (factor 1), or because it has no source language equivalents salient enough to favour its selection in translations (factor 2), or because there are no direct or clear links between the periphrasis and its related concepts in the source language (factor 3). All three factors do not manifest themselves in the same direction on every occasion, so it has to be assumed that in some cases one or two may be cancelling out the influence of the rest. Nevertheless, we should remember that verbal periphrases are an extremely heterogeneous group of elements in terms of meanings and functions, so it is very difficult to draw conclusions about the entire group taken as a whole.

Neither can we claim that all the conditions included in Hareide's definition are met by all under-represented periphrases, since, as we have seen, some are very frequent in Spanish, others have direct equivalents in the source language and some of these equivalents are also very frequent in English. With respect to their level of prototypicality, we can only venture some conclusions based on what we have read about these elements in the literature and on the results of corpus analysis, but we believe that in some cases we are also dealing with prototypical forms in Spanish or in English. Consequently, what we have found is a confluence of factors that explain under-representation or over-representation in each particular case, as shown in table 58 of chapter 4. A very important factor not contemplated in Hareide's definition that may be responsible for the scant use of a given structure in translations – even if this structure is frequent or prototypical in Spanish – is the existence of translation alternatives in the target language which may be favoured for some reason, such as stylistic or personal preferences. It could be argued, of course, that these reasons are ultimately always determined by factor 3, which refers to how the concepts are connected in the mind of the translator.

◆ **Demonstratives:**

In the case of these elements, factor 1 – the frequency of use or prototypicality of the item in the target language – did not exert enough pull to guarantee a similar

representation in originals and in translations, at least in the case of determiners and pronouns, which were omitted much more frequently than adverbs. In spite of being frequent and prototypical in their deictic and phoric functions, in Spanish there are other means to perform those functions, and that seems to make them less necessary than in English. Factor 2 – the salience of their source-language equivalents – was clearly the strongest one in the case of adverbs, which were over-represented in translations, probably due to the influence of the source texts, but was not enough to prevent the translator's tendency to normalise in the case of determiners and pronouns, which were used less frequently than in non-translated Spanish. With regard to factor 3, it is evident that there exists a connection between the English and the Spanish items in the translator's conceptual structure, but this connection seems to be stronger in the case of adverbs than in the other two demonstrative elements, according to the percentage of times that English demonstratives are translated by their Spanish equivalents in each case. In the case of *aquel*, this connection appears to be weaker than in the rest of the demonstratives, since *that* and *those* are more frequently translated by the other distal form, *ese/a(s)*, and the difference in use between the two forms is even higher than in Spanish originals.

The features included in Hareide's definition of a unique item are, once again, not found on all cases. The lack of a direct counterpart in the source language was only present in the first and second degrees of proximity, while all three degrees turned out to be under-represented. Apart from that, there is a structure in English which is very frequent according to our figures (and undoubtedly prototypical, since we are talking about a closed system): the set of demonstratives, which should be influencing the selection of one or other form in Spanish. What we see is that the pull it exerts is very different if we talk about adverbs or about pronouns and determiners. Lastly, Spanish demonstratives cannot be considered infrequent or non-prototypical, except maybe for *aquel* when compared to the other two degrees of the system. Nevertheless, this factor does not seem to be decisive either, since all forms are under-represented in the subcorpus of translations.

◆ Suffixes:

In this last study we found a general under-representation of the entire group of suffixes analysed and also of each of the three groups taken separately. If we take each suffix individually, however, this under-representation is not verified in all instances

(especially in the case of the pejorative suffixes). The factor that has undoubtedly had more influence on the results is the second one, given that there is a clear lack of formally equivalent elements in the source texts. Moreover, in Spanish there are several other ways to express the appreciative values contained in these suffixes, although it should be remembered that the pull exerted by the first factor – frequency or prototypicality in the target language – could not be accurately measured. The third factor – the connectivity of structures in the two languages in the mind of the translator – seems to favour the under-representation of these elements too, since their meaning can be expressed through several different structures both in English and in Spanish.

With regard to Hareide's definition of unique item, in the first place let us remember that we consider these suffixes to lack a straightforward equivalent in the source language. In the second place, we cannot really say if there is any salient structure in the source language that may trigger their use (either similar or different in form to our suffixes), but what we do know is that there is a wide variety of structures that perform similar functions. As regards the third condition, we do not know to what extent these suffixes are or not prototypical, but we have seen that the most frequent and prototypical forms of each group (at least in the case of diminutives and augmentatives) are the ones that are most likely to be under-represented, as opposed to what the definition states.

To sum up, the first condition (“no direct counterpart in the source language”) is only fully met by the suffixes and only partially in the other two cases; the second condition (“no salient or prototypical source language structure that may impact its choice in the target language”) is met partially by the group of under-represented periphrases and perhaps by the suffixes, but not by the demonstratives; and the third condition (“being infrequent and non-prototypical”), partially in the case of the verbal periphrases but not of the demonstratives – unless we refer to the lower frequency of use of *aquel* with respect to the other two forms – or the suffixes, except when taken as a group. As can be seen, it is very difficult to draw clear conclusions from the analysis of groups of elements that have so little to do with each other. Our main goal was to collect evidence from different indicators in order to be able to validate or refute the Unique Items Hypothesis, but this means that it is not possible to assess the influence of the factors contemplated by the Gravitational Pull Hypothesis – which will have a

different impact on each item – as accurately as if we were working on a single item or a more uniform set of elements.

Nevertheless, it is evident that under-representation is a very complex phenomenon and, as we saw in table 58 – which describes the different factors that can lead to it in the case of verbal periphrases – it is difficult to determine *a priori* which elements may constitute unique items. The factors that can influence their use are very diverse, and one of them is the fact that there may exist other forms of expression that translators favour in certain contexts (regardless of the fact that our unique item may or may not be frequent or prototypical) due to their similarity to the source language structures or even to personal preference.

Given that *a priori* identification is a complex, often unsuccessful task, we have also tried to implement Chesterman's technique, explained in chapter 4, to identify unique elements *a posteriori* (2007: 11-12). Following those steps, the unique elements identified in each chapter would be the following ones:

- ◆ **Verbal periphrases:** “ir + gerund”, “haber que + infinitive”, and “comenzar a + infinitive”.
- ◆ **Demonstratives:** the three basic forms of determiners and pronouns, although the one that comes closest to our concept of unique item is *aquel*.
- ◆ **Suffixes:** the diminutives *-ete/a*, *-illo/a*, and *-ito/a* and the augmentatives *-ón/a* and *-ote/a*.

However, we believe that an arbitrary threshold of significance established to identify possible unique items should only be used as a guide and never prevent researchers from analyzing other elements that do not reach that threshold but nevertheless may be of interest to them. As stated by Römer and Wulff (2010: 119): “While statistical significance tests are invaluable tools to quantify strong associations in the data, we would miss out on a number of interesting tendencies by discarding the results as irrelevant on the basis of the failure of the data to meet an arbitrary significance threshold”.

7.4. Confluence of features

Under-representation is not the only phenomenon that we have found in our data. In addition to **over-representation** of demonstrative adverbs, throughout these pages we have mentioned various instances of normalization, explicitation and simplification. For example, in chapter 4 we found signs of a tendency towards **levelling out** in translations, since their use of certain inchoative periphrases was apparently more homogenous than in texts written originally in Spanish. We found instances of **explicitation** in all three studies, usually when the translator adds an element to clarify certain syntactic or semantic relations. In chapter 5 we observed a smaller variety of structures in translated text (with a lower frequency of use of demonstratives after the noun and before a relative clause), a fact that we linked to a possible dynamic of **normalisation** of translated Spanish through the use of more standard options. Finally, we also found two possible cases of simplification in the smaller variety of verbs with which the periphrasis «echar a + infinitive» combines and in the wider variety of bases to which the suffixes attach to form diminutives in the subcorpus of non-translations.

All these phenomena are bound to influence the way our unique items are used, so our results are probably a combination of all kinds of causes and cannot be boiled down just to the three factors of Halverson's hypothesis – as she herself states when she refers to the probable influence of social and cultural factors apart from the cognitive ones (2003: 233) – nor can the three conditions of Hareide's definition be met in every circumstance, since other factors will always be affecting the behaviour of these elements.

Linguistic phenomena are very complex and it is not easy to decide where one ends and the next begins. Many different factors can account for the same result and that is why in this study we have found an interweaving of under-representations, over-representations, normalization, source text interference and levelling out of translated texts. All we can do is extract tendencies, and the limitations of studies of this kind, as has been repeatedly mentioned by those who have studied the subject of translation universals before us, makes this task a difficult one. Let us bear in mind, for example, that when we changed the texts of our corpus in our study on periphrases, without even altering aspects such as genre, we obtained conflicting results.

When we try to find out why this happens, there are many suspects: the linguistic preferences of specific authors or translators, the choice of elements analysed and excluded from the study, perhaps simply the statistical tests, which may be affected by the number of real words in each subcorpora, or even the linguistic norms prevailing on each side of the Atlantic or in specific areas of Spain. In relation to the latter idea, it is interesting to note that despite the fact that we did not establish this criterion of comparability explicitly, almost all the translations included in both versions of our corpus were done by translators born, raised or living in Catalonia. This is hardly surprising, since an important part of the Spanish publishing industry is located in Barcelona. The variety of Spanish that we have studied is therefore the one that can be found there. Would the results be the same if we were to include translators from Madrid? More research needs to be carried out in order to answer this question.

7.5. Recapitulation

The main objective of this study was to find out if the group of elements selected for our analysis was under-represented in translations into Spanish. This under-representation has been confirmed, if we take all elements of each study as a whole. Our results suggest that authors of Spanish texts use certain expressions typical of their mother tongue more frequently than translators, who seem to respond to the lack of straightforward equivalents in the source language that may work as stimuli, since they do not use these expressions to the same extent and in the same way that they would when writing a text in Spanish without the influence of a source text.

If we dig deeper and consider each member of the three groups of elements individually, not all of them confirm our hypothesis. In the case of verbal periphrases we expected only partially positive results, since we had decided to include a “control group”. Nevertheless, frequencies of use in translated texts were lower even among some of the members of this control group. In the case of demonstratives what we did not expect was to find all forms of pronouns and determiners under-represented and almost all forms of locative adverbs over-represented, when what we had anticipated was that only one of the distal forms would show untypical frequencies in translations. The results obtained in the case of suffixes came closer to our expectations, even though

many forms were not under-represented. We believe that the size of the corpus may have affected the results more than in the previous two cases, since for many of them CQP yielded practically no concordances.

While aware that our study is mainly empirical and descriptive and its limitations do not allow us to look for the causes of these phenomena, we have nonetheless attempted to identify throughout the analysis possible factors that may correlate with the use of our unique items in order to try to suggest very tentative explanations for our results. We have also tried to link our observations with the factors identified by Halverson that may favour – or otherwise – the under-representation of unique items in translations. Moreover, we have tried to verify whether the conditions included in Hareide’s definition of a unique item were met by the elements analysed in our study, and this has not always been the case. Nevertheless, it should be borne in mind that the conclusions on this subject drawn from this study can only be partial, since more specific studies that lie outside the scope of this work need to be carried out in order to be able to talk about the prototypicality of a certain element in a given language or of its salience in the translator’s conceptual networks. In addition, as was previously mentioned, the diversity and lack of uniformity of the elements analysed in this thesis prevents us from making conclusive statements about the incidence of the different factors on them. This means that we cannot rule out the possibility that in future studies with a wider scope or carried out on a more specific object of research, Hareide’s intuitions may be confirmed.

What we have found is that the lack of triggering structures in the source text is not the only factor that may cause an under-representation of a given item, as stated by the Unique Items Hypothesis. Other factors that have to do with the way translators use the target language are also at play. Otherwise, the under-representation of elements such as the demonstrative *este* or certain periphrases with straightforward equivalents in English would make no sense.

Finally, the three studies carried out in this thesis have allowed us to compile a catalogue of translation equivalents for our unique items that may be of use to translators or to contrastive studies. As Johansson (2007: 6) claims when discussing the different contrastive studies included in his book: “The most striking finding is the extent to which words may differ in use cross-linguistically, although at the outset one might have regarded them as equivalent”. Through our analysis we have also found that in many cases the elements that we considered the most evident equivalents for certain

English structures turned out not to be used to the extent that we expected in translations, and were replaced by a wide variety of alternative equivalents that we had not thought of at the beginning.

The main interest of this work lies in the fact that, even though the data have not supported our hypothesis for each and every element among those selected for analysis, we have confirmed, through three clear and replicable empirical studies to which we have incorporated some of the methodological improvements suggested by several authors through the years – such as combining comparable and parallel corpora or using statistical tests –, that the phenomenon of under-representation of unique elements of Spanish without formally similar equivalents in English can be attested in texts translated into Spanish. The Unique Items Hypothesis is therefore a phenomenon present not only in translations from English into Swedish, Finnish, Chinese and French, but also in our language pair. Studying this hypothesis in different conditions (different elements of study, different language combinations and different corpora) is a step forward towards validating or rejecting the status of under-representation as a general feature of translation. Through a detailed description of the corpus used in this study and a thorough explanation of every step of the process we ensure its replicability and we contribute to the progress of knowledge in the field of corpus-based studies.

It is a step forward, as we say, but only one, since our conclusions apply solely to the language pair American English-Peninsular Spanish and to the genre of crime novels. In addition, our corpus has only about a million and a half words in total. These limitations do not allow us to make generalizations at a higher level. We have described the language used in this specific genre in translations into Spanish with respect to certain parameters, but in order to claim that these features are general tendencies of translation, more data should accumulate about other languages, other periods, and other linguistic elements.

Nevertheless, the results from every study such as this one are important, since they add to the evidence collected up to this moment in favour of the idea that there are certain general tendencies, independent of external factors that may affect the process of translation, since they belong to the process itself; certain features that are found in translated texts and seem to originate in the mental processes specific to translation, that we must not judge as correct or incorrect, but as realities worth analysing, since they may help us deepen our understanding of the process and products of this activity. As

Eskola puts it (2004: 86): “Studying translation universals is like trying to solve a jigsaw puzzle. Every piece of information about the use of any single pattern is part of the whole when we try to find out what translations are really like”.

7.6. Future research

The observed differences in frequencies may of course be due to several different factors of a linguistic, cognitive, sociocultural or situational nature (Chesterman 2003: 20) that cannot be taken into account in a study like ours. These factors can be social, temporal, related to the translation brief, to the genre of the texts, to the nature of the texts themselves – in the sense that some of the situations included in a novel may render certain expressions unnecessary and therefore less frequent – and even to the abilities, knowledge and attitudes of the translators. Chesterman (2004: 45), for example, mentions the following as possible causes for certain features of translation related to the specific sociocultural conditions of the translation situation and to the role of the translator as mediator:

Other kinds of explanations may be sought in the nature of translation as a communicative act, and in translators’ awareness of their socio-cultural role as mediators of messages for new readers (see e.g. Klaudy 1996). Translators tend to want to reduce entropy, to increase orderliness. They tend to want to write clearly, insofar as the *skopos* allows, because they can easily see their role metaphorically as shedding light on an original text that is obscure – usually unreadable in fact – to their target readers: hence the need for a translation. Their conception of their role may give a prominent position to the future readers of their texts; this may have been emphasized in their training, for instance. It is this conception of their mediating role that may offer some explanation for the tendency towards explicitation, towards simplification, and towards reducing what is thought to be unnecessary repetition – to save the readers’ processing effort.

The factors just mentioned cannot be ruled out unless our study includes every possible variable or several studies are carried out to control the effect of each variable in turn.

Keeping to the possibilities and limitations of a study like this one, we believe that it would be possible to widen its scope by enlarging our corpus with more originals and translations and by adding texts from different times and genres, in order to reduce the influence of any idiosyncratic element introduced by specific authors, translators, genres or texts (or to isolate the influence each of them may exert over the results). It would also be possible to add a fourth component to the corpus to make it bidirectional: the translations of our Spanish novels to English, if they were available in the future or if we changed the texts of the subcorpus of Spanish originals. This would allow us to check whether the features that we are analysing can be found in both translation directions, a fact that would strengthen the hypothesis of their universal nature. Another addition that could be made to the corpus is a set of translations from languages other than English – such as French. As Baker suggests, by including different source languages it may be possible to determine their impact on translated Spanish and to verify if the features found in translated texts are actually independent of the language pair. The corpus that we have compiled can be complemented by other ones such as the COVALT corpus, which includes two target languages (Catalan and Spanish) and three source languages (English, French and German). Finally, the analysis of different translation universal candidates in the same group of texts, as suggested by Corpas Pastor (2008: 126), is something that would add quite a lot of perspective to this kind of studies, since instead of focusing on just one candidate, we would begin to study in a systematic manner the different phenomena that can be found in translations and the way they relate to one another.

On a more practical level, it would be interesting to find the most efficient way to correct the tagging of our texts to be able to carry out more precise searches, and not only of lexical elements, but also of syntactic categories and structures. It would be possible to select other typical elements of Spanish to see if the hypothesis still stands after changing once again the object of study. Some examples that we have come up with are discourse markers such as *bueno* or *ya*, interjections and exclamations such as “¡anda ya!”, typical phraseology pertaining to specific areas of Spanish culture, such as bullfighting, the superlatives *-ísimo* and *-érrimo*, as mentioned in chapter 6, or elements of a more grammatical nature, such as the gerund (studied by Hareide 2013), the imperfect past tense, impersonal passive constructions with *se* or the standard order of elements in a clause.

Studies on the nature of translation cannot be limited to the analysis of quantitative data obtained through automatic analysis tools. We believe it is necessary to try to delve into the reasons that may lie behind the phenomena observed. One possible line of research would be to look for these explanations, in line with Chesterman's causal model (2003). Chesterman believes a theory of translation should build upon this model, in which cognitive causes – the subjective decisions taken consciously and unconsciously by the translator – are of fundamental importance. This means that if we were to undertake this kind of research we would have to focus on the analysis of the translation process, not its products, and it would be necessary to complement textual corpus-based studies with methods and instruments of a very different nature, such as think-aloud protocols, dialogue protocols, or keylogging (cf. Tirkkonen Condit 2002a). Another possibility would be designing simpler cognitive-based studies such as the ones suggested or carried out by authors such as Kujamäki (2004), whose study on snow-related elements and driving in Finnish is described in chapter 2, or Chesterman (2004: 42), who believes that if universal tendencies manifest themselves as specific features of translated text, then it should be possible to use these features (such as untypical frequencies of use of certain elements) to decide if a text is an original or a translation. We could therefore analyse a potentially universal feature in a translated and a non-translated text, and find out whether this analysis allows us to guess which one of the texts is the translation. We could also select a set of texts that contain instances of this feature and another set that does not and see which texts are defined as translations by the subjects of our experiment (see Toury 1995: 230 and Tirkkonen-Condit 2002b).

These data would provide information about the factors that may influence the process of translation and general or universal tendencies, since, as Mauranen (2007) explains, the universals that we are referring to are not necessarily absolute. They are affected by various factors of a cultural, historical, sociological or cognitive nature, among others. It is therefore necessary to take all these variables into account when postulating possible general (or universal, or whatever we decide to call them) translation tendencies. After all, finding translation universals is not the most important goal; what really matters is the fact that there may be certain tendencies in translating that originate in the mental processes specific to this activity but are influenced by various external factors and that their study might help us improve our understanding of the translation process and its products and even, as Chesterman suggests (2003: 27), to

improve translator training and the quality of translated texts, once we have identified those conditions that bring about, in his words, “more desired effects and fewer unwanted ones”.

Apéndice 1: Proceso de búsqueda y selección de las novelas del corpus y reseña biográfica de los autores y traductores incluidos

Proceso de búsqueda y selección de las novelas del corpus

En muchos de los trabajos consultados, la descripción de los criterios de selección de los textos o elementos que conforman el corpus es a veces una simple enumeración de características. Numerosos autores, no obstante, abogan por una mayor explicitación del proceso de selección, que debe tener en cuenta el criterio de representatividad del que hablábamos antes:

For the purpose of corpus construction, the conception of the object which a discipline more or less agrees on provides the motivation for defining a target population. The specification of a target population then provides the basis for selecting a sampling frame. The sample taken from that frame constitutes the corpus, which is thus a representation of a larger set of phenomena. (Halverson 1998: 5)

Para identificar dicha población meta, Biber (1993: 245) propone una serie de parámetros de muestreo estratificado. Con respecto a los referidos a un corpus de texto escrito, que es el que nos interesa aquí, el autor comenta lo siguiente:

Within writing, the first important distinction is publication. This is because the population of published texts can be operationally bounded, and various catalogues and indexes provide itemized listings of members. For example, the following criteria might be used for the operational definition of 'published' texts: (1) they are printed in multiple copies for distribution; (2) they are copyright registered or recorded by a major indexing service. In the United States, a record of all copyright registered books and periodicals is available at the Library of Congress.

Como hemos podido comprobar, es complicado encontrar registros completos de libros publicados en estas grandes bases de datos si queremos seleccionar un género en concreto. La base de datos del ISBN posee un índice de materias²²⁹. De este índice seleccionamos las que más se aproximaban al género buscado: «Género Policíaco Y De Misterio», «Obra De Misterio Y Suspense», «Clásicos policíacos», «Misterios históricos», «Obra de espionaje y espías» y «Obra de suspense político y judicial». Si las juntáramos todas y seleccionáramos una década concreta (en este ejemplo, la de los noventa), apenas obtendríamos 26 registros. Autores como Andreu Martín se encuentran clasificados bajo la etiqueta más general de «Ficción y Temas Afines», que devuelve más de 1000 resultados (por lo que no es posible visualizarlos). Aun así, esta base de datos nos resultó de mucha utilidad en pasos posteriores, al buscar qué obras estaban traducidas y publicadas en España o de qué fecha databan las primeras ediciones de ciertas novelas. En cuanto al ámbito anglosajón, la página web de la Library of Congress de Estados Unidos proporciona la posibilidad de buscar por palabras clave bastante acotadas²³⁰, pero aun así hemos comprobado que no todos los autores u obras que podemos considerar parte del género aparecen catalogados con alguna palabra que apunte a ello, como *mystery*, *crime* o *detective*. En el artículo de Lee (2001: 61) en el que el autor describe los diferentes sistemas de clasificación por géneros de algunos corpus y catálogos, se nombra la categoría «detective and mystery» de la British Library. Accedimos a este recurso²³¹ y, después de varias búsquedas, averiguamos que existía también la etiqueta «Detective and mystery stories, American». También es posible acotar los años de publicación de 1986 a 2004. De todas formas, dado que nuestra impresión general es que la fiabilidad de estos sistemas en cuanto a clasificación por géneros no es muy elevada, tuvimos en cuenta estas listas pero nos guiamos principalmente por los medios que explicaremos más adelante²³².

²²⁹ Base de datos de libros editados en España del Ministerio de Educación, Cultura y Deporte (búsqueda avanzada). [en línea] Consultada el 27 de septiembre de 2015 en http://www.mcu.es/webISBN/tituloSimpleFilter.do?cache=init&prev_layout=busquedaisbn&layout=busquedaisbn&language=es

²³⁰ Catálogo electrónico de la *Library of Congress* (buscador avanzado). [en línea] Consultado el 27 de septiembre de 2015 en <http://catalog.loc.gov/vwebv/searchAdvanced>

²³¹ Catálogo electrónico de la *British Library* (buscador avanzado). [en línea] Consultado el 27 de septiembre de 2015 en http://explore.bl.uk/primo_library/libweb/action/search.do?mode=Advanced&ct=AdvancedSearch&dscnt=0&fromLogin=true&dstmp=1442312090105&vid=BLVUI

²³² De hecho, una de las autoras que finalmente incluimos en nuestro corpus, Sue Grafton, no aparecería tampoco en el catálogo de la British Library si utilizáramos este método de búsqueda.

En los trabajos referidos a la compilación de corpus se menciona también otro tipo de recursos como la consulta a listas de libros más vendidos, críticas, catálogos de editoriales, estadísticas de préstamos, listas de obras premiadas o bases de datos especializadas (por ejemplo, Hareide y Hofland (2011: 78) utilizaron la base de datos *Norbok*, que recoge los textos de autores noruegos publicados en el extranjero)²³³. Pero no es fácil encontrar catálogos completos de otras décadas en recursos como, por ejemplo, las páginas web de las editoriales, y mucho menos filtrados por género —estos autores suelen referirse a la compilación de corpus más grandes y menos especializados—, por lo que lo que decidimos fue utilizar este tipo de recursos también como fuentes de consulta secundarias para confirmar los datos que hallamos en nuestra búsqueda bibliográfica principal a partir de las fuentes que detallamos a continuación.

En nuestro caso, éramos conscientes de que el universo de la literatura policiaca en España era más restringido que el de las traducciones, puesto que, como hemos comentado en el apartado 3.3.1.6, existe una tradición del género en los países anglosajones muy arraigada —se considera que la primera historia de este tipo fue *Los asesinatos de la rue Morgue*, de Poe, publicada en 1841 (Sánchez Zapatero y Martín Escribà 2010: 289)—, un modelo que en España sobre todo se importaba a través de traducciones o se imitaba, y cuya producción nacional solo empezó a florecer y a legitimarse de alguna manera en el panorama literario nacional a partir del año 1975 (Janerka 2010: 14, 63; Sánchez Zapatero y Martín Escribà 2010: 295). De hecho, en los catálogos de colecciones de novelas policiacas publicadas por editoriales españolas que hemos consultado, la mayor parte de las obras son traducciones (muchas veces clásicos del inglés o el francés, probablemente por temas de derechos de autor). Por ese motivo, la compilación del corpus comparable comenzó a la inversa de cómo suele suceder cuando lo que se busca es estudiar las características de los textos traducidos (Laviosa 1997: 310). Aunque tampoco se podría decir que modelamos nuestro corpus de traducciones en base al de originales: el proceso fue bastante paralelo, puesto que investigamos ambos universos literarios a la vez y escogimos autores que pudieran ser

²³³ Janerka (2010: 50-62), por ejemplo, proporciona un listado de editoriales de varias épocas que se dedicaron a la literatura policial, algunas de las cuales siguen existiendo hoy en día (aunque muchas han sido absorbidas por grandes grupos editoriales): en los años treinta, Dédalo, Juventud, Aguilar, Molino o Hymosa, Sopena; de los cuarenta a los cincuenta, Cliper, Pedro Janés o Plaza; en los setenta comienzan a publicar novela policiaca las grandes editoriales, como Planeta, Grijalbo o Seix Barral, aunque siguen existiendo editoriales pequeñas más especializadas en el género como Laia, Júcar, Bruguera o Sedmay. En la actualidad podemos mencionar las editoriales Planeta, Anagrama, RBA y Espasa Calpe.

representativos de ambos, basándonos en criterios de popularidad y aceptación como exponentes del género en sus respectivas culturas.

Los pasos que seguimos, entonces, para la selección de autores fueron los siguientes:

- a) Extrajimos un listado de escritores del género basado, para las novelas españolas, en la obra de Janerka (2010) —que proporciona un excelente resumen de la historia y evolución del género en nuestro país— y en artículos de investigación como los de Sánchez Zapatero y Martín Escribà (2010) o López Merino (2005).

Dado que Index Translationum tampoco filtra por géneros, para el corpus de traducciones (y, posteriormente, paralelo), seguimos el mismo proceso que para el de originales españoles, basando nuestro listado de autores inicial, esta vez, en la cronología de autores y obras de *The Cambridge Companion to Crime Fiction* (Priestman 2003) y en el libro *Crime Fiction* (Scaggs 2005).

- b) A partir de los nombres que extrajimos de estas obras, acudimos a las fuentes que hemos considerado secundarias para averiguar hasta qué punto estos escritores eran tan populares o representativos como afirmaban los autores. Esto nos llevó a consultar páginas web especializadas en literatura y en el género policiaco en concreto (incluyendo entrevistas a los autores, biografías y bibliografías, librerías y premios de narrativa policiaca), reportajes sobre el género en suplementos literarios de periódicos, guías de lectura temáticas en diferentes bibliotecas y listados de obras pertenecientes a colecciones de novela policiaca publicadas en España²³⁴ e incluso listados de obras

²³⁴ Algunos de los recursos consultados se incluyen en la bibliografía o en notas a pie de página, dado que los mencionamos en el cuerpo del texto de este apéndice y del capítulo de metodología. El listado de las demás páginas web y recursos a los que acudimos en el momento de la compilación del corpus, de su modificación y también recursos más recientes no mencionados explícitamente en dichas secciones se incluyen en una sección aparte al final de este apéndice. Con respecto a las colecciones, a continuación las enumeramos [última fecha de consulta para todos los recursos: 27/09/2015]:

- Serie Negra de El País (<http://www.elpais.com/corporativos/elpais/coleccionables/2004/serienegra/paginas/entregas.html>),
- Etiqueta Negra de Ed. Júcar (<http://laflechanegrasigloxv.blogspot.com.es/2009/09/coleccion-etiqueta-negra.html>),
- Crimen & Cía. de Ed. Versal (<http://laflechanegrasigloxv.blogspot.com.es/2009/08/coleccion-crimen-cia.html>)

preferidas por los lectores en diferentes plataformas como blogs o foros. Seleccionamos los más nombrados en dichos recursos, puesto que la recepción de sus obras por parte del público y la crítica puede ser un criterio de comparabilidad (no olvidemos que probablemente serán los autores mejores recibidos en el mundo anglosajón los que tendremos a nuestra disposición, traducidos al castellano).

- c) Revisamos su bibliografía y creamos listados con sus obras, fijándonos, sobre todo, en la descripción y clasificación de las mismas (algunos cultivan también otros géneros que no nos interesan, como la literatura juvenil) y los años de publicación (autores como Patricia Highsmith publican desde los años cincuenta).
- d) A partir de dichos listados, extrajimos un grupo de autores potenciales, que habían escrito obras del género que buscábamos en un periodo de tiempo lo más limitado posible y cercano al presente, lo que, para el subcorpus de traducciones, elimina la mayor parte de los autores de las listas —como Hammett y Chandler, los padres de la novela negra—. En este corpus de dimensiones reducidas preferimos incluir autores que escribieron y fueron traducidos en torno a la misma época, y de manera contemporánea a los autores del corpus español. Nuestro objetivo era seleccionar un conjunto de publicaciones que abarcara la menor cantidad de años posible. Nos guiamos sobre todo por los autores españoles, porque, como hemos explicado, tenemos menos de donde escoger.
- e) Para averiguar si existen traducciones al español de las obras preseleccionadas en inglés, y en qué año fueron publicadas, consultamos la base de datos del ISBN.
- f) Finalmente obtuvimos un listado de autores que han publicado novelas policíacas (y, en su caso, han sido traducidos también) en el período

-
- Colección Cosecha Roja de Ediciones B (<http://laflechanegrasiqloxy.blogspot.com.es/2009/08/coleccion-cosecha-roja-ediciones-b.html>)
 - Best Seller Serie Negra de Ed. Planeta (<http://laflechanegrasiqloxy.blogspot.com.es/2009/08/coleccion-best-sellers-serie-negra.html>)

comprendido entre los años 1990 y 1996. Los autores seleccionados fueron los incluidos en el siguiente apartado.

Reseña biográfica de los autores y traductores²³⁵

Fase I - ES

➤ Alicia Giménez Bartlett: nació en Almansa (Albacete) en 1951, estudió Filología Española en la Universidad de Valencia y desde 1975 vive en Barcelona, donde se doctoró en Literatura²³⁶. En la década de los noventa creó la serie de la inspectora de policía barcelonesa Petra Delicado, que, a pesar de no ser la primera detective femenina de las novelas policíacas españolas, es la más popular; tanto que hasta fue protagonista de una serie de televisión en 1999 (Sánchez Zapatero y Martín Escribà 2010: 302-303). Hasta la fecha lleva publicados diez libros de dicha serie, que ha sido muy bien recibida en Italia y Alemania. «Sus libros poseen muchos de los rasgos de las novelas de enigma, aunque sería mejor clasificarlos como representantes del subgénero llamado *police procedural*²³⁷, cuyos protagonistas son miembros de las fuerzas de seguridad para quienes la investigación constituye su labor cotidiana» (Janerka 2010: 70). El personaje de Petra Delicado nació del deseo de crear una mujer protagonista de novela negra, porque, por lo general, como comenta Giménez Bartlett, «la mujer en la novela negra o es la víctima, que aparece muerta en la primera página, o es la ayudante de alguien» (Mora 2009). En 2008 ganó el premio Raymond Chandler y en 2015, el premio Pepe Carvalho «por haber renovado la novela policíaca española, aportándole una perspectiva femenina y

²³⁵ No en todos los trabajos de compilación de corpus se realiza una investigación exhaustiva sobre los antecedentes de los traductores. Por ejemplo, Ebeling (1998: 3) admite con respecto al componente de ficción del ENPC: «Not much is known about the translators. What we do know is that all the fiction material has been published by reputable publishers, and we have no reason to believe that they did not employ professional translators».

²³⁶ La autora (2015). [en línea] Recuperado el 27 de septiembre de 2015 de Alicia Giménez Bartlett http://www.aliciagimenezbartlett.es/es/La_Autora/

²³⁷ Sánchez Zapatero y Martín Escribà (2010: 299), refiriéndose a la serie *Brigada Central*, de Juan Madrid, comentan: «También denominada “novela de procedimiento”, esta variedad narrativa del género se caracteriza por su estética realista, por la utilización de técnicas documentales y por el protagonismo exitoso de los agentes de la ley. Como su propio nombre indica, estas novelas mostraban los procedimientos oficiales que se han de seguir para llevar a cabo una investigación y reflejaban la cotidianeidad del trabajo policial, recurriendo a la combinación narrativa de varios casos simultáneos, a la descripción de técnicas de la policía científica y del trabajo de oficina y, en definitiva, las rutinas de trabajo de los cuerpos de seguridad».

feminista que ha resultado pionera en este ámbito. Por haber creado dos personajes de ficción –la inspectora Petra Delicado y su inseparable Fermín Garzón– que a lo largo de sucesivas novelas se han ido consolidando hasta alcanzar un reconocimiento propio de las figuras clásicas del género, así como difusión internacional»²³⁸.

➤ Juan Madrid: nació en Málaga en 1947 y en 1959 se trasladó a Madrid. Fue boxeador aficionado, se licenció en Historia en la Universidad de Salamanca y ha trabajado como periodista, novelista, guionista de cine y de televisión, y profesor en distintas universidades e instituciones²³⁹. Sobre sus dos series de novela policiaca más famosas, la de *Brigada Central* (llevada a la televisión en 1989) y las ocho novelas protagonizadas por Toni Romano (ex-boxeador y vigilante de un club nocturno), dice Janerka (2010: 67): «Sin duda, pertenecen a la novela negra, con más acción que investigación y deducción, y narran los pormenores de la vida urbana en Madrid, representando (elaborando) de forma muy personalizada y diversa el lenguaje de los bajos fondos de la ciudad». Valenzuela (2008) comenta sobre él: «(...) ninguno como Juan Madrid tan fiel al género. En forma de cuentos, novelas y guiones de cine y televisión, no lo ha abandonado jamás, y ya va para treinta años. Y sobre todo, lo practica con extrema pureza, con intransigencia casi: escenarios situados en esa zona de sombra donde coinciden la alta sociedad y la marginalidad, y uso estricto de la descripción y el diálogo como recursos narrativos».

➤ Andreu Martín: Nació en Barcelona en 1949 y allí se licenció en psicología y trabajó como guionista y traductor de cómics y redactor para distintas revistas. Desde 1979, año en que se editaron sus primeras novelas, no ha parado de escribir y publicar, tanto novela policiaca (cuenta ya con más de cuarenta títulos) como juvenil²⁴⁰. Entre otros premios, ha recibido el del Círculo del crimen por su novela policiaca *Prótesis* (1980) y, en 1993, el premio Hammett de la Asociación

²³⁸ Alicia Giménez Bartlett gana el Premio Pepe Carvalho del 2015. (29 de octubre de 2014). [en línea] Recuperado el 27 de septiembre de 2015 de *Todoliteratura.es* <http://www.todoliteratura.es/noticia/7577/novela-negra/alicia-gimenez-bartlett-gana-el-premio-pepe-carvalho-del-2015.html>

²³⁹ Madrid, Juan (2009). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/madrid-juan-1259>

²⁴⁰ Martín, Andreu (2013). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/martin-andreu-1290>

Internacional de Escritores Policiacos por *El hombre de la navaja*. «Sus novelas, cuya trama casi siempre se sitúa en Barcelona, presentan casos de robos y asesinatos privados, cuyas causas, leídas entre líneas, indican que la culpa la tiene la sociedad (marginación, paro...)»; «ambos [Juan Madrid y Andreu Martín] narran la violencia con un estilo realista y, a primera vista, crudo, utilizando frases cortas y lenguaje callejero» (Janerka 2010: 61, 68). «El estilo de Martín es claro, sencillo, seco, conciso, funcional, ágil y ameno, o —si intentamos decirlo con una sola palabra— sintético» (López Merino 2005).

➤ Jorge Martínez Reverte: nació en Madrid en 1948 y es periodista, documentalista y escritor. Además de su labor periodística en distintas revistas, periódicos y emisoras de radio, ha escrito libros sobre la historia española reciente y la memoria histórica²⁴¹. En 1979 creó la serie de Gálvez, un periodista madrileño un poco torpe, que ha protagonizado ya seis novelas (la última en 2013). «Los temas políticos, contrastados con la visión del hombre común, que Gálvez encarna a la perfección, constituyen la característica principal de esta serie de novelas» (Janerka 2010: 66), que incluyen un importante componente de humor.

➤ Manuel Vázquez Montalbán: nació en Barcelona en 1939 y murió en 2003. Estudió Filosofía y Periodismo y, además de periodista, fue poeta, ensayista y novelista. En 1972 dio vida al detective privado Pepe Carvalho, un personaje que protagonizaría veinticuatro novelas (dos póstumas) y adquiriría fama mundial. Por esta serie obtuvo el premio Planeta en 1979 y el Prix International de Litterature Policère en 1981. Además, se ha reconocido su obra completa a través del Premio Nacional de las Letras Españolas (1995)²⁴². La de Carvalho fue «la primera serie policiaca española que logró un éxito generalizado y, como tal, funciona como una piedra de toque para todas las obras policiacas posteriores» (Janerka 2010: 107). «Carvalho es un gallego afincado en Barcelona, detective privado, ex-comunista, ex-agente de la CIA, cínico y desengañado de casi todo, que pone voz a las ideas del autor en el plano político, social y literario, mientras quema libros en la chimenea de

²⁴¹ Reverte, Jorge M. (2008). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/reverte-jorge-m-1575>

²⁴² Vázquez Montalbán, Manuel (2009). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/vazquez-montalban-manuel-2375>

su torre de Vallvidrera e imparte clases magistrales de alta gastronomía a lo largo de todas sus aventuras» (Berriatúa 2006). Cada libro de la serie presenta sus características particulares, pero suele identificársela con el subgénero de la novela negra. La acción transcurre sobre todo en Barcelona. Vázquez Montalbán no escribió solo novela negra²⁴³ y, de hecho, era un autor reputado y premiado en otras áreas de la literatura, por lo que, según Janerka (2010: 187), introdujo «un fenómeno nuevo en el mundo literario de España: una serie policiaca que combina el atractivo de una historia con la estética de una obra literaria de elite».

Fase I - EN y TRAD

➤ James Ellroy: nació en Los Ángeles en 1948 y sus primeros treinta años de vida fueron muy turbulentos²⁴⁴. «(...) in some of the most stylistically self-conscious writing in the field since Chandler [Ellroy] may be said to have put the ‘hard’ back into ‘hardboiled’, works, strictly speaking, in the sub-genre of the (corrupt, i.e., LAPD) police procedural» (Porter 2003: 111). Sus novelas están ambientadas en su ciudad natal, generalmente en los años cincuenta, y algunas, como *LA Confidential*, han sido llevadas a la gran pantalla. Posee un estilo muy particular; utiliza frases cortas y ambiguas pero potentes, en las que omite elementos fundamentales como los verbos (aspecto que comentaremos en el capítulo de análisis de las perífrasis). En Clements et ál. (2008) se resume así su obra: «Ellroy's labyrinthine novels chart a West Coast underworld of corruption and evil, played out against real historical events. Bent cops, nightsticks, psychopaths and seductresses».

❖ Carlos Gardini: nació en Buenos Aires en 1948. Es escritor de ciencia ficción, fantasía y terror²⁴⁵, y sus obras han ganado diversos premios literarios, como el de la Universitat Politècnica de Catalunya (UPC) de ciencia ficción²⁴⁶. Ha traducido a autores como Isaac Asimov, Henry James

²⁴³ Bibliografía (1997). [en línea] Recuperado el 27 de septiembre de 2015 de *Vespito.net* <http://www.vespito.net/mvm/biblio.html>

²⁴⁴ Ellroy, James (2012). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/ellroy-james-13>

²⁴⁵ VV. AA. (2014). Carlos Gardini. [en línea] Recuperado el 27 de septiembre de 2015 de *Wikipedia* https://es.wikipedia.org/wiki/Carlos_Gardini

²⁴⁶ Tríptico de Trinidad (2010). [en línea] Recuperado el 27 de septiembre de 2015 de *Bibliópolis Fantástica* <http://www.bibliopolis.org/editorial/bibliofan/tripticotrinidad.htm>

o D.H. Lawrence. Estudió Filosofía y Letras y también ha sido crítico literario.

➤ Patricia Highsmith: nació en Texas, en 1921, pasó su juventud en Nueva York, donde estudió periodismo y, posteriormente, se trasladó a Europa y murió en Suiza, en 1995. Su nombre real era Mary Patricia Plangman. Su primera novela, *Extraños en un tren*, publicada en 1951, fue llevada al cine por Alfred Hitchcock²⁴⁷. En 1955 escribió la primera de las cinco novelas de la serie de Tom Ripley (*El talento de Mr. Ripley*), conocidas en todo el mundo y adaptadas varias veces al cine. «Tom Ripley es sin duda uno de los personajes más fascinantes de la historia de la novela policíaca. Inteligente, cultivado, frío, neurótico y amoral, este antihéroe de Patricia Highsmith se sirve de la mentira, el robo, la estafa y, ocasionalmente, también del asesinato para lograr sus propósitos» (Biblioteca Nacional de España 2014: 27). «The novels of Patricia Highsmith examine murder as developing from ‘ordinary’ behaviour and motivation, and in this respect are concerned with the ‘why’ of crime and criminality» (Scaggs 2005: 115).

❖ Isabel Núñez: nació en 1957 en Girona y falleció en Barcelona en 2012. Estudió Ciencias de la Educación y fue profesora de traducción en la Universidad Pompeu Fabra²⁴⁸. Además de traducir a autores como Andy Warhol o Dorothy Parker, fue crítica literaria y escritora de ensayo, crónica y relatos²⁴⁹.

➤ Evan Hunter: también conocido como Ed McBain, aunque su verdadero nombre era Salvatore Lombino, nació en Nueva York en 1926 y murió en Connecticut en 2005. En 1986 recibió el premio Grand Master de la asociación Mystery Writers of America y en 1998, el Cartier Diamond Dagger Award de la Crime Writers

²⁴⁷ Highsmith, Patricia (2009). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/highsmith-patricia-983>

²⁴⁸ Valls, F. (2012). La escritora Isabel Núñez ha muerto [en línea]. Recuperado el 27 de septiembre de 2015 de *La nave de los locos* <http://nalocos.blogspot.com.es/2012/11/la-escritora-isabel-nunez-ha-muerto.html>

²⁴⁹ Arroyo, F. (2013). Isabel Núñez: "Escrivivir" [en línea]. Recuperado el 27 de septiembre de 2015 de <http://blogs.elpais.com/tormenta-de-ideas/2013/11/isabel-nu%C3%B1ez-escrivivir.html>

Association de Reino Unido²⁵⁰, ambos en reconocimiento de toda su trayectoria. En 1956 publicó la novela *Cop Hater*, que sería la primera de la serie *Distrito 87*, con la que inauguró el género del *police procedural* en Estados Unidos. «Unlike the earlier writers, McBain chose to depict a team of officers. (...) the police motifs McBain established in his books – [were] the overworked force, the hostile public, the burned out cop, the alienation from normal life, etc.» (Panek 2003: 157).

- ❖ **Carlos Milla Soler**: nació en Barcelona, en 1958. Publicó su primera traducción en 1983 y, además de traducir, trabaja en distintas áreas del mundo editorial. También ha sido vicepresidente de ACETT²⁵¹. Algunos de los autores que ha traducido son Salman Rushdie, John Irving o Jean Auel²⁵². En 2007 recibió el premio de traducción Esther Benítez por la novela *La gran marcha*, de E. L. Doctorow, traducida junto a Isabel Ferrer.

- **Walter Mosley**: nació en Los Ángeles en 1952, estudió en Vermont y en los ochenta se trasladó a Nueva York, donde trabajó como programador informático y ahora es profesor de universidad²⁵³. En los noventa comenzó a escribir, y su primera novela, *El demonio vestido de azul*, ganó tanto el Shamus Award (Estados Unidos) como el John Creasy Award (Reino Unido) a la mejor primera novela policiaca. A partir de entonces, además de otras obras, ha escrito once novelas de la serie de Easy Rawlins, un detective privado negro de Los Ángeles. Ambientadas en las décadas de los cuarenta, cincuenta y sesenta, en estas novelas «not only does he resort to the black vernacular to get his narratives told right, he also sketches in with telling realist detail the different milieux of contemporary black urban life, including such representative locales of social interaction as the often mean streets themselves along with mainly black apartment buildings, churches, clubs and bars» (Porter 2003: 112).

²⁵⁰ McBain, Ed (2009). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/mcbain-ed-1315>

²⁵¹ Milla Soler, C. y Pino Moreno, M. (2006). De te fabula narratur: Los sistemas de cómputo y el rendimiento del trabajo de traductor en el sector editorial [en línea]. *Vasos Comunicantes* n° 34. Recuperado el 27 de septiembre de 2015 de <http://revistavasoscomunicantes.blogspot.com.es/2011/04/vasos-comunicantes-numero-34.html>

²⁵² Ficha profesional de Carlos Milla Soler (s. f.). [en línea] Recuperada el 27 de septiembre de 2015 de *Ace-traductores.org* <http://www.ace-traductores.org/user/142>

²⁵³ Mosley, Walter (2010). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/mosley-walter-1384>

- ❖ Rosa Corgatelli: correctora de estilo para el ámbito editorial y empresarial y traductora del inglés, portugués, francés e italiano desde hace casi cuatro décadas, vive en Argentina y ha trabajado para importantes editoriales argentinas y españolas²⁵⁴. Ha traducido a autores como Robin Cook y Fernando Pessoa. También ha trabajado como lectora para editoriales y periodista²⁵⁵.

- Ruth Rendell: nació en 1930 en Londres y murió allí en 2015. Durante su extensa carrera como escritora, que comenzó en 1964, ganó multitud de premios, entre los que se cuentan los Diamond, Gold y Silver Dagger Awards de la Crime Writers' Association y tres premios Edgar. «Sus personajes pertenecen a una clase media aburrida, inculta, que hace de la puntualidad y las buenas costumbres un rito cotidiano. Pero detrás de las máscaras, a poco que se raspe en ellas, aparecen el asesino, el psicópata, el reprimido sexual. La prosa de Rendell es clara y concisa, no exenta de poesía. En 1997, Pedro Almodóvar filmó “Carne trémula”, basada en un libro de Rendell»²⁵⁶. Puede considerársela parte del género de *police procedural*, pues sus personajes no son investigadores privados sino policías profesionales (Scaggs 2005: 31).

- ❖ Carme Camps: nació en 1950 en Barcelona y vive en Lleida. Es licenciada en Filosofía y Letras por la Universidad de Barcelona, además de profesora de español para extranjeros y titulada en publicidad. Cuenta con treinta años de experiencia como traductora literaria del inglés, el francés y el catalán y ha traducido, entre otros muchos, a Jane Austen o Edgar Rice Burroughs²⁵⁷. También ha sido miembro de la Junta Directiva de la Asociación Colegial de Escritores de Cataluña durante seis años²⁵⁸.

²⁵⁴ Rosa C. (2012-2015). [en línea] Recuperado el 27 de septiembre de 2015 de *Workana* <https://www.workana.com/es/w/rosa-corgatelli-1>

²⁵⁵ Rosa Corgatelli (2009-2015). [en línea] Recuperado el 27 de septiembre de 2015 de *ProZ.com* <http://www.proz.com/profile/1112099>

²⁵⁶ Rendell, Ruth (2015). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/rendell-ruth-1570>

²⁵⁷ Ficha profesional de Carme Camps Monfà (s. f.). [en línea] Recuperado el 27 de septiembre de 2015 de *ace-traductores.org* <http://www.ace-traductores.org/user/486>

²⁵⁸ Carme Camps Monfà (2015). [en línea] Recuperado el 27 de septiembre de 2015 de *LinkedIn.com* <https://www.linkedin.com/in/ccampsm>

Fase II – ES

➤ **Soledad Núñez:** nació en San Sebastián. Además de *La danza de los reptiles* (1994), novela negra ambientada en su ciudad natal y protagonizada por una fugitiva llamada Celia, un comisario y un gánster, esta vasca ha escrito otras dos novelas, *Siempre el adiós* (1990) y *Del amor y otras milongas* (1990)²⁵⁹, y ha continuado la serie de Celia con una segunda novela, *Fundido en negro* (1997)²⁶⁰.

Fase II - EN y TRAD

➤ **Patricia Cornwell:** nació en 1956 en Miami. Fue reportera y analista forense, y en su obra literaria reflejó posteriormente diversos aspectos de este trabajo. El personaje principal de la serie más conocida de Cornwell, la de Kay Scarpetta, es una médica forense, profesional y bastante valiente, que nació en 1990 con la novela *Post mortem*, un éxito de crítica y ventas. «The pattern in the Scarpetta series includes highly detailed descriptions of the bodies on which Kay performs postmortems, FBI involvement in tracking serial killers, endangerment of the detective, and her ultimate triumph over the forces of evil, at least as manifest in one criminal per book» (Reddy 2003: 206). «Los libros de Patricia Cornwell han sido traducidos a 36 idiomas en más de cincuenta países, situándola entre los autores de género negro que más pasiones despiertan internacionalmente»²⁶¹, y sus novelas han ganado, entre otros, los premios Edgar, John Creasey, Anthony, Macavity y el Prix du roman d'aventures (todos por *Post mortem*) y el Gold Dagger Award por *Cruel y extraño*.

❖ **Jorge Luis (Jordi) Mustieles Rebullida:** Ha traducido a P. D. James, Noah Gordon y Margaret Atwood, entre otros escritores, y es autor de un libro de texto de psicología general. Su primera traducción en el ISBN data del año

²⁵⁹ Bueno Martínez, M. (2002) Quince años de literatura vasca en castellano (1985-2000) [en línea]. *Rev. int. estud. vascos*, 47, 1, 11-34. Recuperado el 27 de septiembre de 2015 de <http://www.euskomedia.org/PDFAnlt/riev/47/47011034.pdf>

²⁶⁰ Abasolo, J. J. (2009). Vascos y criminales: Soledad Núñez [en línea]. Recuperado el 27 de septiembre de 2015 de *Nadie es inocente* <http://elblogdeabasolo.blogspot.com.es/2009/06/vascos-y-criminales-soledad-nunez.html>

²⁶¹ Especial Patricia Cornwell (2012). [en línea] Recuperado el 27 de septiembre de 2015 de *serienegra.es* http://www.serienegra.es/articulo/articulos/especiales/1259/especial_patricia_cornwell.html

1984, trabajó diez años como traductor y asesor literario independiente y, al parecer, actualmente se dedica en gran medida a la creación de contenidos web y otras actividades de naturaleza diferente a la traducción²⁶².

➤ Sue Grafton: nació en Louisville (Kentucky), en 1940, estudió Literatura Inglesa y, además de novelas, ha escrito guiones para el cine y la televisión²⁶³. Ha recibido diversos premios, entre ellos, tres Anthony, tres Shamus y el Diamond Dagger Award en 2008. En 1982 comenzó la serie de novelas de la detective Kinsey Millhone, cada uno de cuyos títulos empieza por una letra distinta del alfabeto. Aquella primera novela fue *A is for alibi* y la más reciente es *X*, publicada en 2015. Se trata de una de las series protagonizadas por mujeres detectives más populares, que se considera «closest to the hardboiled tradition in their language, adopting the sarcastic verbal style of the tough-talking male detectives» (Reddy 2003: 200).

❖ Antonio-Prometeo Moya: nació en Ciudad Real en 1949. Vivió en Valencia, donde tuvo diversos oficios, hasta que en 1976 publicó su primer libro y se trasladó a Barcelona, donde se licenció en Geografía e Historia²⁶⁴. Traduce del inglés, del francés y del italiano a autores como George Orwell o Martin Amis y ha publicado diez novelas de su propia cosecha²⁶⁵.

➤ Elmore Leonard: nació en Nueva Orleans en 1925 y vivió en Detroit, donde, además de publicar sus primeras obras, trabajó como redactor publicitario y guionista. Murió allí en 2013. «Su primera novela de misterio, *The Big Bounce*, se publica en 1969. A partir de entonces, sus historias, su estilo crudo y sus diálogos callejeros comienzan a convertirlo en un autor de culto. El gran golpe llega con *La Brava* (1984), ganadora del premio Edgar de la Mystery Writers Association. El mundo de Leonard es un mundo de seres extraños casi siempre excéntricos, a

²⁶² Jordi Mustieles (2015). [en línea] Recuperado el 27 de septiembre de 2015 de *LinkedIn.com* <https://www.linkedin.com/pub/jordi-mustieles/0/5/84>

²⁶³ Grafton, Sue (2009). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/grafton-sue-22>

²⁶⁴ VV. AA. (2015). Antonio-Prometeo Moya [en línea]. Recuperado el 27 de septiembre de 2015 de *Wikipedia* https://es.wikipedia.org/wiki/Antonio-Prometeo_Moya

²⁶⁵ Carreira, M. (2012) Muerte de un ciudadano por encima de toda sospecha; de Antonio-Prometeo Moya [en línea]. Recuperado el 27 de septiembre de 2015 de *Factor Crítico* <http://www.factorcritico.es/muerte-de-un-ciudadano-por-encima-de-toda-sospecha-de-antonio-prometeo-moya/>

menudo peligrosos»²⁶⁶. En 1992 recibió también el Grand Master Edgar Award. Una de las adaptaciones al cine de sus novelas más conocidas es *Jackie Brown*, de Tarantino, basada en *Rum Punch*. A diferencia de las novelas de Highsmith, «Despite Leonard's interest in, and even fascination with, criminality, there is in his novels, as Glover remarks, 'a marked lack of interest in criminal psychology' (Glover 2003: 149). In their examination of the mechanic of crime, rather than the motivation for it, they might be termed *howdunnits*, rather than *whydunnits*». (Scaggs 2005:115).

- ❖ Enrique de Hériz Ramón: nació en Barcelona, en 1964. Novelista y editor, también imparte clases de escritura y actualmente es asesor externo del departamento de traducción de la editorial Salamandra²⁶⁷. Su primera traducción fue de una novela de Elmore Leonard, pero desde entonces no ha parado de traducir a autores como Dashiell Hammett, Noah Gordon o Patricia Highsmith. En 2004 obtuvo el premio de los Libreros de Cataluña (Premio Llibreter) por su novela *Mentira*²⁶⁸.
- ❖ Hernán Sabaté Vargas (traductor de James Ellroy en la novela escogida para esta fase): nació en Barcelona, en 1953 y falleció en 2011. Comenzó a traducir en 1976 y, entre los numerosos autores a los que tradujo, se encuentran Patricia Cornwell, Isaac Asimov y James Ellroy. Se le consideraba, de hecho, «la voz de James Ellroy en español» (Solana 2011), y su última traducción de este escritor, *Sangre Vagabunda*, en colaboración con la traductora Isabel Ferrer, le valió el VI Premio Esther Benítez.

²⁶⁶ Leonard, Elmore (2013). [en línea] Recuperado el 27 de septiembre de 2015 de *Biblioteca Negra* <http://bibliotecanegra.com/autores/leonard-elmores-1129>

²⁶⁷ Enrique De Hériz Ramón (2015). [en línea] Recuperado el 27 de septiembre de 2015 de *LinkedIn.com* <https://es.linkedin.com/pub/enrique-de-h%C3%A9riz-ram%C3%B3n/18/477/235>

²⁶⁸ De Hériz, Enrique [en línea]. Recuperado el 27 de septiembre de 2015 de *edhasa.es* http://www.edhasa.es/libros/libros_personas.php?idp=119&p=De+H%C3%A9riz%2C+Enrique

Listado de recursos consultados para la compilación y modificación del corpus y la redacción de las reseñas biográficas de los autores y traductores no mencionados en el cuerpo del texto

- AbeBooks Inc. y AbeBooks Europe GmbH (1996-2015). *Iberlibro.com* [en línea]. Consultado el 28 de julio de 2011 en <http://www.iberlibro.com>
- Álamo, A. y Vallés, R. (2011). Fundamentales de la novela negra [en línea] Consultada el 28 de Julio de 2011 en *Lecturalia* <http://www.lecturalia.com/libros/fundamentales-de-la-novela-negra>
- Amazon Best Sellers of 1995 (1996-2015). [en línea] Recuperado el 28 de julio de 2011 de *Amazon.com* http://www.amazon.com/gp/bestsellers/1995/books/ref=zg_bsar_cal_ye
- Best Sellers in Mystery, Thriller & Suspense (1996-2015). [en línea] Recuperado el 28 de julio de 2011 de *Amazon.com* http://www.amazon.com/Best-Sellers-Books-Mystery-Thriller-Suspense/zgbs/books/18/ref=zg_bs_nav_b_1_b
- Biblioteca Can Milans (2015). *Guia de lectura Novel·la negra imprescindible (d'ara i de sempre, d'aquí i de tot arreu)* [en línea]. Recuperado el 3 de septiembre de 2015 de http://bibliotecavirtual.diba.cat/documents/10934/34913091/recomanats_negra+imprescindible.pdf/8178c536-cfb7-4298-93ed-1d2b4d828f6d
- Clare County Library (s. f.). *A Reader's Guide to Crime and Mystery Writers* [en línea]. Recuperado el 13 de septiembre de 2015 de http://www.clarelibrary.ie/eolas/library/services/readers_guides/readers_guides_crime_mystery_writers.htm
- Department of English, Brandeis University (1998-2015). *Annual bestsellers, 1990-1999 - Rankings from Bowker's Annual/Publisher's Weekly* [en línea]. Recuperado el 28 de julio de 2011 de <http://unsworth.unet.brandeis.edu/courses/bestsellers/best90.cgi>
- Especial Patricia Cornwell (17 de octubre de 2012). [en línea] Recuperado el 13 de septiembre de 2015 de *Serie Negra* http://www.serienegra.es/articulo/articulos/especiales/1259/especial_patricia_cornwell.html
- Galindo, J. C. (coord.) (2013-2015). *Elemental: el blog de novela negra de El País* [en línea] Consultado el 13 de septiembre de 2015 en <http://blogs.elpais.com/elemental/>
- García, Richie (s. f.). *Ellroy.com* [en línea]. Consultado el 28 de Julio de 2011 en <http://www.ellroy.com/index.html>
- Hardboiled (2009). [en línea] Recuperado el 27 de julio de 2009 de *Wikipedia, La enciclopedia libre*. <https://en.wikipedia.org/wiki/Hardboiled>
- Hawes Publications (2011). *Adult New York Times Best Seller Listings* [en línea]. Recuperado el 28 de julio de 2011 de <http://www.hawes.com/pastlist.htm>

- Heilig, G. y Heilig-Staindl, C. (2012). The Top-100 Crime Novels of All Time [en línea]. Recuperado el 20 de mayo de 2013 de *addictivebooks.com*
http://addictivebooks.com/bookrank/bookrank_1.htm
- Imprescindibles (2001-2015). [en línea]. Recuperado el 28 de julio de 2011 de *negraycriminal.com*
<http://www.negraycriminal.com/index.php?view=imprescindibles>
- Internet Archive (2008). *Open Library* [en línea]. Consultado el 27 de Julio de 2009 en
<https://openlibrary.org/>
- Jakubowski, M. (29 de octubre de 1999). A beginner's guide to crime fiction [en línea]. Recuperado el 28 de Julio de 2011 de *theguardian.com*
<http://www.theguardian.com/books/1999/oct/29/crimebooks>
- Karen (20 de abril de 2008). The Times's "50 Greatest Crime Writers" [en línea] Recuperado el 27 de julio de 2009 de *Euro Crime* <http://eurocrime.blogspot.com.es/2008/04/times-50-greatest-crime-writers.html>
- La novela negra. Juan Madrid. (20 de abril de 2009). [en línea] Recuperado el 27 de julio de 2009 de *Territorio Magenta. Foro no oficial de UPyD* <http://www.territoriomagenta.com/t623-la-novela-negra-juan-madrid#bottom>
- Lillo, A. (1996-2009). *Anika entre libros* [en línea]. Consultado el 27 de julio de 2009 en
<http://libros.ciberanika.com/>
- Listado de autores (2008). [en línea] Consultado el 27 de julio de 2009 en *QuedeLibros.com*
<http://www.quedelibros.com/autores/1.html>
- Martin, J. (31 de diciembre de 2010). Oadby Library Crime Readers' Group: History of Crime Novels [en línea]. Recuperado el 13 de septiembre de 2015 de *Leicestershire County Council*
http://www.leics.gov.uk/index/community/libraries/county_libraries/a_to_z_libraries/oadbywigtown_libraries/oadby_library/oadby_library_crime_group/crime_novels_history.htm
- Martín Cerezo, I. (2005). La evolución del detective en el género policíaco [en línea]. *Tonos: revista electrónica de estudios filológicos*, 10, 362-384. Recuperado el 3 de septiembre de 2015 de
<http://www.um.es/tonosdigital/znum10/estudios/Q-Martin.htm>
- Matuschenko, V. (2006). *Short Hard-Boiled Action Guide* [en línea]. Recuperado el 28 de julio de 2011 de <http://rraymond.narod.ru/a-z-list.htm>
- Muñoz, J. L. (2011). Novela negra y novela enigma [en línea]. Recuperado el 13 de septiembre de 2015 de *Editanet*
<http://archivos.editanet.org/14/0525419e55135dc1e/0525419e551370f23/index.html>
- MysteryNet (1998-2012). *Mystery Time Line* [en línea]. Recuperado el 13 de septiembre de 2015 de
<http://www.mysterynet.com/timeline/>
- Mystery and Thriller Book Awards (1997-2015). [en línea] Consultado el 28 de julio de 2011 en *Barnes and Noble* http://www.barnesandnoble.com/b/books/awards/mystery-thriller-book-awards/_/N-29Z8q8Z1w5k

- Mystery Bookstore-- top mystery authors and books* [en línea]. Recuperado el 13 de septiembre de 2015 de <http://www.mysterybookstore.com/>
- Mystery Writers of America (2011). *TheEdgars.com* [en línea]. Recuperado el 28 de julio de 2011 de <http://www.theedgars.com/edgarsDB/index.php>
- Narrativa policíaca. Novela negra (2015). [en línea]. Consultado el 13 de septiembre de 2015 en *casadellibro.com* <http://www.casadellibro.com/libros/narrativa-policiaica-novela-negra/126000000/2/2>
- Novela negra (2009). [en línea] Recuperado el 27 de julio de 2009 de *Wikipedia, La enciclopedia libre*. http://es.wikipedia.org/wiki/Novela_negra
- Página oficial de Manuel Vázquez Montalbán (2006). [en línea] Recuperado el 27 de Julio de 2009 de *ClubCultura.com* http://www.clubcultura.com/clubliteratura/clubescritores/vazquez_montalban/home.htm
- OCLC Online Computer Library Center (2001-2015). *WorldCat* [en línea]. Consultado el 28 de julio de 2011 en <http://www.worldcat.org>
- Red de Bibliotecas Municipales (s. f.). Guías de lectura: novel·les policíacas [en línea]. Recuperado de *Biblioteca Virtual* <http://bibliotecavirtual.diba.cat/es/guies-de-lectura>
- Sábados Negros (28 de febrero de 2009). Andreu Martín [en línea]. Recuperado el 20 de agosto de 2015 de *sabadosnegros.org* <http://www.sabadosnegros.org/archivo/archivo.html>
- The Best Mystery authors - Top 100 (2015). [en línea]. Recuperado el 13 de septiembre de 2015 de *Top Mystery* <http://www.topmystery.com/authors>
- The New York Times Company (2011). Best sellers [en línea] Recuperado el 28 de julio de 2011 de *NYTimes.com* <http://www.nytimes.com/best-sellers-books/overview.html>
- The Top 100 Crime Novels of All Time [en línea]. Recuperado el 28 de julio de 2011 de *Wikipedia, La enciclopedia libre*. http://en.wikipedia.org/wiki/The_Top_100_Crime_Novels_of_All_Time
- Torres, D. (20 de febrero de 2009). Especial Novela Policiaca: Cosecha negra [en línea]. Recuperado el 27 de julio de 2009 de *El Cultural* <http://www.elcultural.com/revista/letras/Especial-Novela-Policiaca/24788>
- USA Today (2011). *USA Today Bestselling books database* [en línea]. Consultado el 28 de julio de 2011 en <http://content.usatoday.com/life/books/booksdatabase/default.aspx>
- Zeki (2005). Fichas [en línea]. Recuperado el 27 de julio de 2009 de *La gangsterera* <http://gangsterera.free.fr/Fichas.htm>