

Programa de Doctorat de Teoria de la Traducció
Departament de Traducció i Interpretació
Universitat Autònoma de Barcelona

**Els documents digitals especialitzats:
utilització de la lingüística de corpus
com a font de recursos per a
la traducció especialitzada**

Autora: Pilar Sánchez Gijón
Directores: Dra. Anna Aguilar-Amat i Castillo
i Dra. Mariana Orozco Jutorán
Bellaterra, 2003

Índex

<u>INTRODUCCIÓ</u>	<u>9</u>
<u>PART I: Identificació i emmagatzematge de recursos textuais digitals especialitzats</u>	<u>21</u>
<u>Capítol 1. La documentació en el procés de la traducció especialitzada</u>	<u>23</u>
1. ENQUESTA SOBRE LES PRÀCTIQUES DE DOCUMENTACIÓ AMB TEXTOS PARAL·LELS DELS	
TRADUCTORS	26
1.1. Anàlisi dels resultats de l'enquesta	29
1.2. La llengua dels textos paral·lels	31
1.2.1. Textos paral·lels en la llengua de partida	31
1.2.2. Textos paral·lels originals en la llengua d'arribada	32
1.2.3. Traduccions com a textos paral·lels en la llengua d'arribada	33
1.2.4. Les traduccions i els seus originals com a textos d'arribada	34
1.3. Factors que aboquen el traductor a la cerca de textos paral·lels	35
1.4. Metodologia de cerca	36
1.4.1. Fonts de textos paral·lels	37
1.4.2. Procés de cerca dels textos paral·lels	38
1.4.3. La fiabilitat dels textos paral·lels	39
1.5. La utilitat dels textos paral·lels	44
1.6. Reutilització de les traduccions realitzades	44
1.6.1. Conservació de les traduccions	45
1.6.2. Informació que aporten les traduccions realitzades	47
1.6.3. Conservació de la informació extreta de les traduccions realitzades	48
1.7. Resultats globals	50

Capítol 2.El recurs digital textual especialitzat.....	53
1. CARACTERÍSTIQUES DEL TEXT DIGITAL: L'HIPERTEXT	53
1.1. Els orígens de l'hipertext	54
1.1.1. El MEMEX de Vannevar Bush.....	55
1.1.2. Els hereus del MEMEX.....	57
1.1.3. El World Wide Web.....	58
1.1.4. L'hipertext després del WWW.....	61
1.2. Característiques formals del text digital publicat a Internet.....	64
1.2.1. Els elements hipertextuals	65
a) Els nodes i els ancoratges	65
b) Els enllaços.....	68
c) Eines de navegació	72
1.2.2. Les estructures hipertextuals.....	76
2. EL TEXT EN LA COMUNICACIÓ ESPECIALITZADA.....	92
2.1. El llenguatge especialitzat	95
2.2. El text especialitzat.....	98
2.2.1. Classes i tipologies textuales.....	104
2.3. La comunicació especialitzada	108
2.3.1. La gradació de l'especialització	109
a) Els nivells d'especialització.....	110
b) El continuum de la comunicació especialitzada.....	113
2.4. El text especialitzat digital	115
Capítol 3.L'obtenció de recursos textuales.....	119
1. LA RECUPERACIÓ D'INFORMACIÓ: L'OBTENCIÓ DE DADES AL WORLD WIDE WEB ...	119
1.1. Avaluació de la informació al <i>World Wide Web</i>	120
1.1.1. Distribució de la informació a la World Wide Web	122
1.1.2. Tipus de recursos web i qualitat de la informació	127
1.2. La recuperació d'informació.....	136

1.2.1. La representació de la informació.....	137
1.2.2. La selecció i obtenció d'informació	140
1.3. Els sistemes de recuperació d'informació al <i>World Wide Web</i>	144
1.1.3. Els cercadors segons el seu funcionament	144
a) Els directoris o índexs temàtics	145
b) Els motors de cerca	149
c) El web invisible	156
d) Altres sistemes de cerca	162
2. LA CERCA DE DOCUMENTS DIGITALS ESPECIALITZATS.....	166
2.1. La cerca dels recursos textuais digitals especialitzats sobre els Leònids per a la posterior creació de corpus monolingües comparables	167
3. L'extracció de documents digitals de la xarxa.....	175
3.1. El procés d'identificació dels recursos textuais dels corpus	177
<u>Capítol 4. La classificació dels corpus de recursos textuais</u>	<u>183</u>
1. LA LINGÜÍSTICA DE CORPUS.....	184
1.1. ELS TIPUS DE CORPUS.....	187
1.2. LA COMPOSICIÓ D'UN CORPUS <i>AD HOC</i>	193
2. LA CLASSIFICACIÓ DELS RECURSOS TEXTUALS DIGITALS ESPECIALITZATS DEDICATS ALS LEÒNIDS.....	197
2.1. CRITERIS DE CLASSIFICACIÓ DELS RECURSOS TEXTUALS	198
2.1.1. Dades administratives	198
2.1.2. L'autor i el receptor	200
2.1.3. Dades sobre l'estructura del document hipertextual	200
2.1.4. Tipus de lloc web	201
a) Tipologia proposada per Alexander i Tate	201
b) Tipologia de llocs amb recursos textuais d'astronomia observats.....	202
2.1.5. Tipus de recurs textual	203
a) La relació dades/text.....	204
b) Tipologia observada de recursos textuais.....	205
c) Tipologia de continguts observada.....	208
2.1.6. Esquema final de classificació dels recursos digitals.....	209

2.2. DESCRIPCIÓ DEL CORPUS	211
2.2.1. Descripció dels corpus monolingües.....	212
a) Descripció en funció de l'autor i el receptor.....	215
b) Descripció en funció del tipus de lloc web.....	220
c) Descripció del corpus en funció del tipus de recurs textual	222
d) Descripció dels corpus en funció del contingut dels recursos textuais.....	225
PART II: Anàlisi dels recursos recollits i aplicació a la traducció	235
Capítol 5. El text paral·lel en la traducció especialitzada.....	237
1. LA DOCUMENTACIÓ AMB EL TEXT PARAL·LEL EN TRADUCCIÓ.....	239
1.1. LA COMPETÈNCIA TRADUCTORA EN LA TRADUCCIÓ DE TEXTOS ESPECIALITZATS	239
1.2. LA DOCUMENTACIÓ AMB TEXTOS PARAL·LELS DEL TRADUCTOR ESPECIALITZAT	242
1.3. ELS TEXTOS PARAL·LELS	246
1.3.1. La comprensió del text original.....	247
1.3.2. L'elaboració del text final	250
1.3.3. L'anàlisi dels textos paral·lels	251
2. LA TERMINOLOGIA I ELS TEXTOS PARAL·LELS: LA CERCA D'UNITATS DE SENTIT	
ESPECIALITZAT	253
2.1. ELS ENFOCAMENTS TEÒRICS.....	253
2.1.1. El terme com a unitat independent del llenguatge natural.....	254
2.1.2. El terme com a unitat pròpia del llenguatge natural.....	256
2.2. ALTRES ELEMENTS PROPIS DE LA COMUNICACIÓ ESPECIALITZADA.....	260
2.2.1. Les relacions conceptuals.....	263
2.3. LA METODOLOGIA DE TREBALL DE LA TERMINOLOGIA	266
3. L'APLICACIÓ DE LA LINGÜÍSTICA DE CORPUS EN LA TRADUCCIÓ ESPECIALITZADA I LA	
TERMINOLOGIA	269
3.1. ELEMENTS D'ANÀLISI DE LA LINGÜÍSTICA DE CORPUS	270
3.2. EXPERIÈNCIES AMB CORPUS EN LA PRÀCTICA DE LA TRADUCCIÓ.....	278

Capítol 6.L'anàlisi del corpus compilat	283
1. EINES DE GESTIÓ I ANÀLISI DEL CORPUS.....	284
1.1. EL GESTOR DE RECURSOS TEXTUALS (GERT)	284
1.1.1. La fitxa consultes	285
1.1.2. La fitxa selecció	287
1.1.3. La fitxa estadístiques.....	290
1.1.4. Edició	291
1.2. EL PAQUET D'EINES WORDSMITH TOOLS.....	292
1.2.1. L'extractor de llistes de paraules	293
1.2.2. L'extractor de concordances	293
1.2.3. L'extractor de llistes de paraules clau	295
1.3. L'EXTRACTOR D'N-GRAMS.....	295
2. COMPROVACIÓ DE L'ADEQUACIÓ DEL CORPUS MULTILINGÜE	298
2.1. COMPROVACIÓ DE L'ADEQUACIÓ DEL CORPUS ANGLÈS.....	299
2.2. PARAULES CLAU DEL CORPUS CASTELLÀ.....	302
2.3. PARAULES CLAU DEL CORPUS CATALÀ.....	304
3. ANÀLISI DEL CORPUS MULTILINGÜE COMPARABLE: CERCA SISTEMÀTICA	307
3.1. CERCA DE DEFINICIONS.....	308
3.1.1. Cerca de característiques conceptuals	312
3.2. CERCA DE RELACIONS JERÀRQUIQUES	315
3.2.1. Relacions gènere–espècie.....	315
3.2.2. Relacions meronímiques	317
3.2.3. Relacions material–objecte	319
3.3. CERCA DE RELACIONS NO JERÀRQUIQUES	320
3.3.1. Relacions infralògiques	321
a) Relacions cronològiques	321
b) Relacions genètiques	323
c) Relacions causals	326
3.3.2. Relacions argumentals.....	328
3.4. ANÀLISI COMPARATIVA DELS SUBCORPUS EXPERTS I NO EXPERTS.....	330
3.4.1. Corpus anglès	331

a) Cerca de definició	331
b) Cerca de característiques conceptuals	333
c) Relacions gènere–espècie.....	335
d) Relacions meronímiques.....	336
e) Relacions material–objecte	337
f) Relacions cronològiques.....	339
g) Relacions genètiques	340
h) Relacions causals	341
i) Relacions argumentals	342
3.4.2. Corpus castellà.....	343
a) Definició	344
b) Cerca de característiques conceptuals	345
c) Relacions gènere–espècie.....	346
d) Relacions meronímiques.....	348
e) Relacions material–objecte	348
f) Relacions cronològiques.....	349
g) Relacions genètiques	350
h) Relacions causals	351
i) Relacions argumentals	351
4. ANÀLISI DEL CORPUS MULTILINGÛE COMPARABLE: CERQUES PUNTUALS.....	355
4.1. CERQUES D'UNITATS LINGÛÍSTIQUES CONEGUDES DE L'ARTICLE DE PREMSA	356
4.1.1. El cas de PEAK	357
a) The PEAK of	361
b) Sobre la situació de PEAK	363
c) Equivalents de PEAK.....	365
4.1.2. El cas de SHOWER vs STORM	369
a) El cas de SHOWER	369
b) El cas de STORM	373
c) Shower i storm.....	376
d) Cerca d'equivalents	378
4.2. UNITATS LINGÛÍSTIQUES CONEGUDES DE L'ARTICLE ACADÈMIC	386
4.2.1. El cas de PEAK	388
a) Cerca d'equivalents.....	395
4.2.2. El cas d'ACTIVITY	401

a) ACTIVITY al subcorpus anglès acadèmic.....	401
b) ACTIVITY al subcorpus anglès expert–expert.....	405
c) Cerca d'equivalents	409
4.3. UNITATS LINGÜÍSTIQUES DESCONEGUDES DE L'ARTICLE ACADÈMIC	416
4.3.1. El cas d'FWHM	417
a) Cerca d'equivalents.....	419
4.3.2. El cas de NODAL.....	421
<u>Capítol 7.Proposta metodològica sobre l'obtenció i l'explotació de recursos textuais digitals com a textos paral·lels.....</u>	<u>431</u>
1. <i>ELS RECURSOS TEXTUALS DIGITALS PUBLICATS AL WORLD WIDE WEB COM A OBJECTE D'ANÀLISI</i>.....	432
1.1. Inconvenients causats pel format dels recursos digitals	432
1.2. Inconvenients causats per les característiques formals dels recursos digitals	436
a) Inconvenients deguts a les característiques dels documents hipertextuals.....	436
b) Inconvenients deguts a les característiques específiques d'alguns tipus de recursos textuais.....	440
2. <i>EL PROCÉS D'OBTENCIÓ D'INFORMACIÓ A PARTIR DE TEXTOS PARAL·LELS MITJANÇANT LA METODOLOGIA DE LA LINGÜÍSTICA DE CORPUS</i>	446
2.1. L'OBTENCIÓ DE RECURSOS DIGITALS ÚTILS	446
2.1.1. Els llocs web segons la seva situació comunicativa	448
2.1.2. Els llocs web segons el tractament del tema	450
2.1.3. Els llocs web segons l'activitat del responsable secundari.....	453
2.2. L'EXTRACCIÓ D'INFORMACIÓ DELS RECURSOS OBTINGUTS.....	457
<u>Conclusió i perspectives de futur</u>	<u>461</u>

Índex

<u>Bibliografia.....</u>	<u>471</u>
REFERÈNCIES BIBLIOGRÀFIQUES PER ORDRE ALFABÈTIC.....	471
REFERÈNCIES BIBLIOGRÀFIQUES PER ÀMBITS TEMÀTICS.....	513
Documentació, teoria de l'hipertext i ciències de la informació	513
Lingüística aplicada i lingüística textual	525
Lingüística de corpus.....	531
Terminologia.....	542
Traducció	550
Tradumàtica	559
<u>Llistat de figures i taules.....</u>	<u>563</u>
1. FIGURES	563
2. TAULES	566
<u>ANNEX A: ENQUESTES.....</u>	<u>579</u>
a) Empreses / Traductors en plantilla	579
b) Traductors del Servei de Traducció de la Unió Europea.....	591
c) Traductors freelance	598
<u>ANNEX B: CLASSIFICACIONS I DESCRIPCIONS GRÀFIQUES DELS CORPUS CASTELLÀ I CATALÀ.....</u>	<u>618</u>
<u>ANNEX C: CORPUS EN ANGLÈS, CATALÀ I CASTELLÀ</u>	<u>625</u>
<u>ANNEX D: SUPORT DIGITAL</u>	<u>739</u>

Introducció

El gran volum d'informació d'abast global condiona, per bé i per mal, la societat actual. L'aparició i el subsegüent èxit mundial d'Internet com a mitjà de comunicació és, en part, responsable d'aquest creixement exponencial de la informació que es produeix i es distribueix, però no n'és l'únic. De fet, no podem separar la noció de *societat de la informació*, en la qual Internet i, de manera més general, la comunicació en format digital, és un element clau, de la de *societat del coneixement*. Totes dues nocions s'impliquen, i són resultat, alhora que punt de partida, de la situació socioeconòmica de la segona meitat del segle passat i d'aquest inici de segle.

En aquest sentit, en el document elaborat pel Departament de Societat del Coneixement de la Unió Europea, es defineixen *societat del coneixement* i *de la informació* com el resultat de la transformació econòmica i social provocada per la introducció i la difusió a gran escala de les tecnologies de la informació i la comunicació (TIC), basades en microprocessadors. En aquest document també es considera que el coneixement i la informació s'han convertit en el fonament per a l'organització i el desenvolupament d'activitats econòmiques i socials, ja que les TIC han fet de la *societat del coneixement* el que són a causa de la seva capacitat de codificar informació i permeten processar i manipular coneixement per satisfer necessitats de diferent mena i transmetre'l per tot el món (Knowledge Society, 2002: 2).

Aquest marc té grans repercussions sobre la pràctica professional de la traducció, com ara el fet que els nous mitjans de comunicació impliquen l'aparició de nous formats i productes que han de ser traduïts. El nostre interès, però, se centra fonamentalment en els dos efectes següents:

1. El caràcter intersocial i intercultural de la *societat del coneixement* fomenta l'intercanvi de coneixement especialitzat, tècnic i científic, entre diferents països i, per tant, diferents llengües. Aquest intercanvi es caracteritza tant

per la rapidesa amb què es duu a terme com pel seu caràcter caduc: les innovacions tecnològiques i científiques s'estenen ràpidament pel planeta, i caduquen de seguida que sorgeix un nou avenç tecnològic o científic que les suplanta. Aquest relativament curt cicle de vigència dels coneixements, efecte dels nostres temps, fa que el traductor es trobi en un procés continu de formació i reciclatge dels seus coneixements factuais, fet que es veu accentuat per la necessitat professional del traductor de diversificar els seus àmbits d'actuació per tal de ser competitiu.

2. En aquest procés continu de formació, els mitjans propis de la *societat de la informació* hi intervenen de ple dret, en tant que mitjans que permeten un accés gairebé immediat a fonts d'informació d'arreu del món. És per això, doncs, que el traductor ha de compaginar fonts d'informació tradicionals, com ara diccionaris, enciclopèdies, manuals, etc., amb altres fonts pròpies dels mitjans de comunicació més recents, com és la documentació que es publica a Internet. L'aparició d'aquestes fonts d'informació obliga al traductor a adquirir habilitats i destreses de consulta noves, que li permetin fer-ne un ús adequat a les característiques de les fonts i a les seves pròpies necessitats.

La convergència d'aquestes dues repercussions ha estat la motivació principal d'aquesta tesi doctoral. En concret, tenim un especial interès en el procés de documentació i extracció d'informació en relació amb la traducció científica, àmbit que tractem contínuament com a docents de terminologia aplicada a la traducció.

És per això que, tot i que l'aplicació de la recerca que a continuació presentem s'emmarca en l'àmbit de la traducció, i en concret de la traducció científica, necessàriament s'aborda des d'altres disciplines que conflueixen, ja sigui en la metodologia o bé en l'objecte d'estudi. Així, molts dels plantejaments teòrics i metodològics que s'utilitzen com a punt de partida en aquest treball provenen de disciplines com la documentació i les ciències de la informació, la lingüística textual, la terminologia i la lingüística de corpus.

La documentació i les ciències de la informació són disciplines que aborden la gestió de la informació, tant des d'un punt teòric com metodològic, i ens facilitarà l'apropament a l'element textual típic d'Internet, les pàgines web, i a les diferents metodologies que permeten classificar-les i recuperar-les. La lingüística textual proporciona elements de caracterització i anàlisi del text, i en concret del text especialitzat, que és un dels principals objectes d'estudi d'aquesta tesi. Per la seva banda, la terminologia, en tant que disciplina dedicada a l'estudi del lèxic propi de cada àrea de coneixement, tant des d'una perspectiva teòrica com aplicada, aporta instruments que el traductor pot utilitzar per construir estratègies de cerca d'informació, com també caracteritza les unitats que típicament vehiculen el coneixement especialitzat. Finalment, la lingüística de corpus proporciona una metodologia d'estudi lingüístic descriptiu basat en el caràcter quantitatiu de les dades de partida que s'analitzen; aquesta metodologia ha de permetre al traductor identificar la informació lingüística i factual que requereix per dur a terme la seva tasca professional.

En conjunt, la recerca que a continuació presentem l'entendem en el marc de la tradumàtica, àrea que Vilarnau i Dalmau definí com a "conjunt de coneixements i habilitats en l'ús de les TIC que permeten desenvolupar la tasca de la traducció d'una manera més ràpida, eficaç i fiable" (2001: 18), ja que l'ús correcte de les tecnologies de la informació i la comunicació (TIC) per part del traductor roman en el rerefons de qualsevol operació relacionada amb els instruments i les eines pròpies de la comunicació en l'àmbit de la *societat de la informació*, com la que es presenta en aquesta tesi. Tal com ja ho vam indicar, "la tradumática incluye todo lo referente a la información multilingüe que el traductor recupera, genera, procesa y transfiere, así como los sistemas de producción, gestión y almacenamiento de esa información, y toda iniciativa o estudio vinculado a ese complejo" (Sánchez-Gijón *et al.*, 2002: 140).

Objectius

Aquesta tesi es desenvolupa al voltant de dos objectius principals, el primer de caràcter metodològic i el segon de caràcter teòric, que es concreten en set subobjectius:

- I. Establiment d'una metodologia que permeti al traductor professional identificar i resoldre problemes de traducció a partir de recursos textuais digitals en l'àmbit de la traducció científica. En aquest cas, l'àmbit d'aplicació serà la traducció científica en les combinacions anglès-castellà i anglès-català. Per tal d'assolir aquest objectiu, considerem necessari:
 - * Aprofundir en el concepte d'hipertext. Els recursos textuais digitals típics d'Internet, i fonamentalment del *World Wide Web* estan integrats en documents hipertextuals. El traductor ha de conèixer les seves característiques formals per tal de fer-ne un bon ús com a font d'informació.
 - * Aprofundir en la noció de text especialitzat. Ja que els recursos textuais digitals provinents del *World Wide Web*, recursos dels quals el traductor extraurà informació lingüística i factual pertinent per a la seva tasca professional, són part de documents hipertextuals, ens aproparem al seu caràcter especialitzat a partir de trets lingüístics i extralingüístics. Atès que analitzar els documents sencers resultaria una tasca difícil d'assolir, l'anàlisi no s'abordarà des de la perspectiva dels gèneres textuais.
 - * Recollir un corpus especialitzat multilingüe comparable *ad hoc*. Aquesta experiència permetrà establir estratègies de compilació i anàlisi en tant que recurs per a l'obtenció d'informació lingüística i factual per a la traducció científica. En aquest cas, i per tal que els resultats puguin ser extrapolables a altres àmbits científics, escollirem un tema científic susceptible de ser tractat en gairebé totes les situacions de la comunicació especialitzada. El tema que es tractarà és l'observació dels Leònids, la pluja de meteorits que té lloc cada mes de novembre. Aquest fenomen és observat i analitzat tant per experts com per semiexperts, per la qual cosa el ventall de formes comunicatives resultant és molt ampli. Gràcies a la diversitat de recursos textuais digitals sobre aquest tema que esperem recollir al corpus *ad hoc*, les conclusions seran susceptibles de ser extrapolables als recursos textuais

dedicats a altres àmbits especialitzats. Les llengües utilitzades seran l'anglès, el castellà i el català.

- * Caracteritzar els recursos textuais digitals especialitzats que integren el corpus *ad hoc* en tant que exemple del ventall de recursos dedicats a àmbits especialitzats. Aquesta caracterització es durà a terme partint de criteris d'avaluació de la qualitat dels recursos textuais proposats per diferents iniciatives emmarcades en la documentació i les ciències de la informació.
- * Dissenyar una eina informàtica de gestió dels recursos textuais digitals i que permeti la creació de subcorpus de recursos amb característiques comunes per tal de ser analitzats per separat.
- * Establir metodologies d'anàlisi del corpus *ad hoc*. Analitzar el corpus tot identificant estratègies que permetin, d'una banda, introduir el traductor en un àmbit temàtic especialitzat que desconeix i, de l'altra, resoldre problemes puntuals que el text de partida, o l'elaboració del text d'arribada, li pugui plantejar, problemes que poden ser tant de caràcter lingüístic com factual i que, en aquest cas, respondran a les combinacions lingüístiques anglès–castellà i anglès–català.
- * Establir estratègies de compilació d'un corpus *ad hoc*. A partir de la compilació d'un corpus *ad hoc* i la seva anàlisi, identificar els recursos textuais de major utilitat per al traductor en aquesta mena d'anàlisi.

II. Actualitzar la noció de text paral·lel a la llum de les noves necessitats informatives del traductor i els recursos que li proporcionen, així com la manera d'explotar aquest recurs de la traducció científica. A fi d'assolir aquest objectiu, considerem necessari caracteritzar prèviament els comportaments sistemàtics dels diferents col·lectius de traductors en l'àmbit científic pel que fa a la utilització de textos paral·lels.

Hipòtesis de partida

Les característiques de la *societat de la comunicació* i *de la informació* no només tenen repercussions sobre el cicle de vigència dels coneixements i la immediatesa de la comunicació, sinó que aquesta també es recolza sobre nous productes textuais que no sempre seran la versió digital d'un tipus de text tradicional i publicat en format paper, ja que el canal provoca l'alteració de condicionants de la comunicació, tant lingüístics com extralingüístics, com ara qui pot rebre el recurs textual (que no necessàriament coincidirà amb la figura a qui el document s'adreça), com s'hi arriba, i també aspectes relatius a la intertextualitat i la intratextualitat del recurs i del document a què pertany o a la seva macroestructura.

Un altre dels efectes de la immediatesa de la comunicació escrita actual és la manca de processos de revisió de molts documents del *World Wide Web* prèviament a la seva publicació. A més, en aquest mitjà sovint es publica informació sense indicar-hi les dades mínimes per identificar la seva font i, per tant, la seva fiabilitat. Tanmateix, si bé aquests aspectes afecten gran part dels recursos textuais del *World Wide Web*, la nostra hipòtesi de partida és que els afecta de manera individual, no col·lectiva. El traductor no podrà prendre decisions en funció de la informació d'un recurs textual determinat si no és suficientment fiable. Això no obstant, si el que consulta és un nombre considerable de recursos, i el conjunt recull una dada informativa determinada, l'aspecte quantitatiu la dotarà de la fiabilitat que cada recurs per separat no presenta.

Altrament, l'anàlisi quantitativa és la base de la metodologia proposada per la lingüística de corpus. Aquesta branca aplicada i descriptiva de la lingüística està tenint diverses aplicacions en l'àmbit de la traducció, principalment en la didàctica (Corpas, 2001; Rodríguez Inés, en elaboració). Segons la nostra segona hipòtesi de partida, l'estudi de corpus *ad hoc* integrats per recursos textuais digitals especialitzats mitjançant la metodologia i les eines d'anàlisi de la lingüística de corpus ha de permetre establir estratègies d'explotació d'aquests corpus com a textos paral·lels que facilitin l'obtenció d'informació tant lingüística com factual. Si és així, l'estudi del corpus compilat ha de permetre establir els tipus d'estratègies vàlides per a aquesta mena d'anàlisi, tant si el que el traductor necessita és entrar en contacte amb una àrea

temàtica que desconeix o si tot el que cerca és identificar i resoldre problemes de traducció.

Metodologia

Per tal de comprovar les dues hipòtesis indicades en el subapartat anterior, hem cercat al *World Wide Web* recursos textuais dedicats a l'àmbit temàtic de la pluja de meteorits coneguda amb el nom de Leònids. Aquest fenomen astronòmic anual és apreciable sense instruments específics, per la qual cosa acostuma a ser una cita tant d'experts com d'aficionats a l'astronomia. La cerca de recursos textuais, doncs, es duu a terme a partir de les paraules clau Leònids i astronomia en cadascuna de les llengües de treball: anglès, castellà i català.

Un cop identificats els recursos textuais que responen a aquestes paraules clau i descarregats al disc dur, classifiquem els recursos recollits en funció de criteris de caràcter formal, comunicatiu i documental, fet que ens permetrà dur a terme anàlisis separades dels diferents tipus de recursos textuais identificats i, finalment, establir quins són els tipus de recursos de major utilitat com a component d'un corpus *ad hoc* per al traductor. Aquesta classificació es durà a terme mitjançant una eina informàtica dissenyada específicament amb aquesta finalitat, programa que hem denominat GeRT. En els criteris de classificació s'han tingut en compte aspectes provinents de la documentació i les ciències de la informació, com també de la lingüística textual, en concret la vessant d'aquesta branca que estudia el text especialitzat. Durant la fase de classificació es desestimen tots aquells recursos textuais que han estat recuperats inapropiadament.

A continuació, i després d'analitzar les necessitats informatives del traductor en funció de les seves possibles mancances de coneixement, tant lingüístic com factual, se cerquen les estratègies d'estudi del corpus *ad hoc* que, partint de la metodologia proposada per la lingüística de corpus, permetin resoldre aquestes necessitats informatives. A més de les eines informàtiques originades per dur a terme estudis

Introducció

descriptius segons els principis de la lingüística de corpus, també utilitzarem alguns dels seus instruments d'anàlisi que s'adeqüen perfectament a les estratègies d'interrogació i extracció d'informació que pot desenvolupar el traductor.

L'objectiu de l'anàlisi realitzada és cobrir essencialment dos tipus de necessitats possibles del traductor: d'una banda, les d'aquell traductor que desconeix l'àmbit temàtic en què s'emmarca el text original que ha de traduir i que, per tant, necessita obtenir coneixements factuais passius per tal de dur a terme la seva traducció; de l'altra, les possibles mancances puntuals, lingüístiques i/o factuais, que tot traductor pot trobar en el text que ha de traduir.

Estructura

La tesi es divideix en dues parts diferenciades. La primera, titulada "Identificació i emmagatzematge de recursos textuais digitals especialitzats", està dedicada a la identificació, l'obtenció i la caracterització dels recursos textuais en tant que textos paral·lels per a la traducció. Aquesta part està formada pels capítols de l'1 al 4. La segona, titulada "Anàlisi dels recursos recollits i aplicació a la traducció", es dedica a l'establiment d'estratègies d'anàlisi i extracció d'informació en funció de les necessitats informatives del traductor. Aquesta part la integren els capítols 5, 6 i 7.

El capítol 1 permet contextualitzar la recerca que es du a terme en els següents capítols. Aquest capítol inicial presenta un breu estudi sobre les pràctiques de documentació, principalment a partir de textos paral·lels, de tres diferents col·lectius de traductors: els traductors en plantilla d'empresa, els que treballen en serveis públics i els traductors autònoms. La dades recollides en aquest capítol han estat extretes d'una enquesta orientativa realitzada entre traductors de tots tres col·lectius. L'objectiu d'aquesta enquesta de conèixer de primera mà les tendències documentals dels traductors professionals, aspecte central d'aquesta tesi. Els resultats de les enquestes es recullen posteriorment a l'annex A.

La noció de recurs textual digital especialitzat s'aborda detalladament al capítol 2. En primer lloc, s'estudia la seva vessant digital, en tant que part de documents hipertextuals publicats al *World Wide Web*. Partint, doncs, de la teoria de l'hipertext, es descriu l'adaptació dels documents hipertextuals al WWW i les seves repercussions. Després d'analitzar els components d'un document d'aquestes característiques, s'estudien les estructures que més habitualment presenten i els seus efectes sobre el lector. A continuació s'aborda la noció de text especialitzat des de la perspectiva de la lingüística textual, principalment de l'escola alemanya. Aquest enfocament teòric ens permetrà caracteritzar els recursos textuais especialitzats en tant que parts de documents hipertextuals, ja que no els observarem dins del context en què estan integrats: caracteritzarem, doncs, el component textual de les pàgines web, no els llocs web en la seva integritat.

Un cop identificades les característiques formals, lingüístiques i extralingüístiques dels recursos textuais digitals especialitzats, el capítol 3 es dedica a les tècniques i els instruments per a l'obtenció d'aquesta mena de recursos del WWW, tot partint del marc teòric de la documentació i les ciències de la informació, en concret de l'àmbit conegut com a recuperació de la informació (RI). Després de descriure els instruments que permeten cercar informació al WWW, i els principis que els regeixen, aquest capítol s'il·lustra amb la localització i descàrrega dels recursos que integraran el corpus *ad hoc* que utilitzarem com a conjunt de textos paral·lels per a la resolució de problemes de traducció.

La primera part, dedicada a la identificació i emmagatzematge de recursos textuais digitals especialitzats, conclou amb el capítol 4, que està dedicat a la descripció dels postulats principals de la lingüística de corpus. En aquest capítol es presenta una tipologia dels diferents corpus existents i, a continuació, es mostra el sistema de classificació de recursos textuais digitals especialitzats seguit en el nostre cas. Aquests criteris d'ordenació que s'hi presenten permetran observar la utilitat dels diferents tipus de recursos textuais recollits quan s'empren com a textos paral·lels en l'àmbit de la traducció científica. A partir dels paràmetres previstos, fonamentalment de caràcter extralingüístic, s'ofereix una descripció acurada del corpus per llengües, il·lustrada a partir de les dades relatives als recursos en llengua anglesa (la informació estadística

relativa als recursos en castellà i català es recullen a l'annex B, mentre que la classificació completa dels recursos que integren el corpus es presenta a l'annex C).

La segona part d'aquesta tesi, la relativa a l'anàlisi dels recursos recollits i la seva aplicació, s'inicia amb l'explicitació de les necessitats informatives del traductor, necessitats que ha de saber identificar i cobrir amb el procés de documentació. Després d'abordar la noció de text paral·lel, observem com es caracteritzen des de la terminologia les unitats que representen coneixement especialitzat, unitats que generalment són les que plantegen els problemes de comprensió i redacció en la llengua final més evidents per al traductor. A més, els enfocaments terminològics també poden proporcionar elements per a la creació d'estratègies de cerca i extracció d'informació útils per al traductor. Finalment, un cop establertes les necessitats informatives del traductor científic, es descriuen els plantejaments teòrics de la lingüística de corpus, i els instruments que permeten observar-los.

Al capítol 6, després de presentar els programes informàtics que ens permetran realitzar les diferents anàlisis dels corpus compilat i d'estudiar la seva adequació com a recurs del qual es pot extreure informació lingüística i factual per a la traducció de dos textos de diferents nivells d'especialitat, es presenten dues metodologies d'anàlisi i extracció de dades d'acord amb les necessitats informatives del traductor. Aquestes dues anàlisis, amb objectius ben diferenciats, dotaran el traductor d'estratègies i elements d'interrogació de corpus d'aquesta mena. Cadascuna de les anàlisis realitzades s'il·lustren amb exemples del corpus compilat, per la qual cosa en aquest capítol s'inclou un gran nombre de taules i figures, necessàries per a l'exemplificació dels estudis que es duen a terme.

Aquesta segona part de la tesi finalitza amb el capítol 7, en el qual es presenta la proposta metodològica final, tot recalcant les característiques dels recursos textuais digitals especialitzats més útils en funció de les necessitats informatives del traductor, així com resumint les estratègies que caracteritzen els dos tipus d'anàlisi duts a terme en el capítol anterior: l'anàlisi sistemàtica i l'anàlisi puntual del corpus *ad hoc*.

En les conclusions finals es realitza una recapitulació dels aspectes tractats al llarg d'aquest treball i es presenten les perspectives de futur d'aquesta recerca.

Finalment, l'apartat dedicat a la bibliografia està dividit en els mateixos àmbits teòrics que sostenen aquest treball: documentació i ciències de la informació, lingüística textual, terminologia i lingüística de corpus, a més de traducció i tradumàtica. Algunes de les referències citades apareixen en més d'un subapartat de la bibliografia a causa del seu caràcter interdisciplinari. En relació amb la bibliografia també volem recalcar que part de les referències són d'obres de recerca en format digital, per la qual cosa en citar-les textualment no hem pogut indicar-hi cap pàgina, com és habitual segons el sistema de citació de l'estil Chicago, que és el que hem seguit en aquest treball.

Qüestions terminològiques

Atès el caràcter interdisciplinari d'aquesta tesi, certs termes s'encavalquen entre àrees de coneixement diferents i poden provocar malentesos. És per això que hem volgut dedicar un breu apartat de la introducció a certes qüestions terminològiques.

Alguns dels problemes terminològics que hem detectat es deuen al solapament entre la teoria de l'hipertext i la tecnologia del *World Wide Web*. Com a criteri de diferenciació, entendrem els termes propis de la teoria de l'hipertext com a unitats lògiques, mentre que les nocions relatives al WWW són de caràcter virtualment delimitador i físic. Així, mentre que *document hipertextual* és una noció de caràcter lògic, l'estructura del qual respon a un raonament conceptual i a l'establiment de determinades relacions entre els nodes que l'integren, el *lloc web* l'entendem com un conjunt de pàgines web que es troben dins del mateix domini i, per tant, podem considerar-lo una unitat física, virtualment tangible, d'arxius. És per això, doncs, que tot i que sovint es fan servir com a formes sinònimes, en el decurs d'aquesta tesi els diferenciarem.

Introducció

D'altra banda, i seguint les indicacions recollides al mòdul Cercaterm del lloc web del Termcat¹, utilitzarem la denominació *web* en masculí per referir-nos al *World Wide Web*.

Un altre aspecte terminològicament conflictiu és el que presenta la noció de *paral·lel*. Aquesta noció, en traducció, s'utilitza generalment per qualificar els textos que es fan servir com a font d'informació per a la resolució de problemes de traducció, i generalment són originals. Altrament, en lingüística de corpus, la noció de *paral·lel* implica la confluència de dos conjunts de textos en llengües diferents, confluència caracteritzada pel fet que un dels conjunts està format per textos originals i l'altre per la seva traducció a una altra llengua. Aquest conflicte terminològic l'hem intentat resoldre evitant la variació denominativa i explicitant en totes les ocasions el nucli de la unitat terminològica a què ens referim: ja sigui *text paral·lel* o bé *corpus paral·lel*.

La intenció de la tesi és, doncs, trobar la maneres d'optimitzar la utilització de la informació existent per part del traductor científic, artífex i peça clau de la *societat del coneixement* multilingüe.

¹ El Cercaterm és el mòdul de consulta terminològica i resposta automàtica del TermCat, i es pot consultar a <http://www.termcat.es/cercaterm/>.

Capítol 1. La documentació en el procés de la traducció especialitzada

El traductor és un consumidor d'informació amb característiques molt especials. Si bé tot professional ha de conèixer les seves necessitats informatives, en el cas de la traducció encara cal ser més conscient d'aquestes necessitats atesa la seva condició final de comunicador de coneixement especialitzat entre llengües diferents (vegeu l'apartat 2 "La cerca de documents digitals especialitzats" del capítol 3, pàgina 166).

Tal com assenyala Mayoral (1997/1998), entre els diferents recursos a l'abast del traductor es troben els textos paral·lels, en els quals cercarà tant informació factual com lingüística. Els textos paral·lels, com es veurà amb més detall en l'apartat 1 "La documentació amb el text paral·lel en traducció" del capítol 5 (pàgina 239), permeten al traductor aprofundir en el domini temàtic en què s'emmarca la seva traducció alhora que augmenta la seva competència lingüística en aquell àmbit, o viceversa.

En lloc de partir d'hipòtesis sobre el comportament del traductor professional quan recorre a textos paral·lels¹ durant el desenvolupament de la seva tasca, hem decidit partir d'informació real, tot i que no representativa, sobre aquest procés. Per aquest motiu, i per tal d'obtenir més informació sobre les tècniques de documentació del traductor professional, en concret les relacionades amb l'obtenció de textos

¹ En el capítol 5 s'aprofundeix en el concepte de text paral·lel. D'entrada, i pel que fa als resultats de l'enquesta que a continuació presentem, hem considerat text paral·lel aquell text original o traduït en llengua de partida o d'arribada del qual el traductor extrau informació lingüística o factual que li permet prendre decisions en una nova traducció.

paral·lels i la seva posterior explotació, vam sol·licitar a professionals de la traducció que responguessin de manera anònima a l'enquesta que es presenta a l'apartat següent. L'objectiu d'aquest breu estudi que hem dut a terme no és el de determinar fefaentment les pràctiques documentals dels traductors professionals, sinó únicament conèixer les tendències més esteses.

Els conjunt heterogeni de professionals que es dedica a la traducció de textos científics i tècnics ha estat dividit en tres col·lectius en funció de la seva situació laboral per tal d'intentar detectar diferències sistemàtiques entre els diferents professionals. Els col·lectius que hem tingut en compte són els següents:

- * Els traductors que treballen en plantilla en una empresa de traducció.
- * Els traductors que treballen en serveis de traducció d'organismes públics.
- * Els traductors autònoms o *freelance*.

Al primer dels col·lectius, el format per traductors que treballen en empreses de traducció, s'ha arribat mitjançant la mediació de l'Agrupació de Centres Especialitzats en Traducció (ACT), com també de la llista de distribució Traducción en España de RedIRIS². No es va voler obtenir informació de traductors vinculats a empreses estrangeres per tal de vetllar per la comparabilitat dels resultats obtinguts amb les enquestes procurant que les possibles diferències no es deguessin a condicions tecnològiques o socials, o bé a pràctiques laborals no equiparables. En total, el col·lectiu format per traductors que treballen en empreses aporta 17 de les 55 enquestes computades, i representen 31% del total de la mostra.

El col·lectiu de traductors d'organismes públics està format exclusivament per traductors funcionaris dels serveis de traducció dels organismes europeus (que hem abreviat com a SdT), en concret de la Comissió Europea i del Parlament Europeu. Hem volgut diferenciar aquests col·lectius de la resta de professionals per raó de les característiques especials de l'entorn en què desenvolupen la seva tasca, sovint

² Llista moderada i dedicada a la traducció a Espanya. La seva pàgina web de presentació és <http://www.rediris.es/list/info/traduccion.es.html>.

recolzada per recursos o centres de documentació propis. Aquest col·lectiu aporta el 16,3% (9 enquestes) de la mostra total.

L'últim dels col·lectius previstos i el més extens és el dels traductors autònoms. Es tracta de professionals que poden treballar per a empreses de traducció o organismes oficials, o bé directament per a clients, sense intermediaris. Per entrar en contacte amb aquest col·lectiu hem utilitzat la llista de Traducción de Rediris, així com les llistes de les associacions de professionals de la traducció ATIC (Associació de Traductors i Intèrprets de Catalunya) i TRIAC (Traductors i Intèrprets Associats pro Col·legi). Les enquestes que aquest col·lectiu aporta el 52,7% (29 enquestes) de la mostra total.

1. Enquesta sobre les pràctiques de documentació amb textos paral·lels dels traductors

L'enquesta que hem dissenyat esta dividida en setze preguntes, que es poden englobar en tres blocs:

- * La classe de text que s'utilitza habitualment com a text paral·lel.
- * Com s'obtenen els textos.
- * Com es guarda la informació extreta de textos paral·lels o de traduccions anteriors.

Pregunta 1. Realitzeu una cerca específica de **textos paral·lels** per a cada traducció especialitzada que heu de fer?

En la llengua de partida	En la llengua d'arribada
<input type="checkbox"/> Sempre	<input type="checkbox"/> Sempre
<input type="checkbox"/> Molt sovint	<input type="checkbox"/> Molt sovint
<input type="checkbox"/> Sovint	<input type="checkbox"/> Sovint
<input type="checkbox"/> Poques vegades	<input type="checkbox"/> Poques vegades
<input type="checkbox"/> Mai	<input type="checkbox"/> Mai

Pregunta 2. Feu servir **textos en la llengua de partida** com a textos paral·lels?

Sempre	Molt sovint	Sovint	Poques vegades	Mai	Amb quina finalitat

Pregunta 3. Feu servir **textos en la llengua d'arribada** com a textos paral·lels?

Sempre	Molt sovint	Sovint	Poques vegades	Mai	Amb quina finalitat

Pregunta 4. Feu servir **traduccions en la llengua d'arribada** com a textos paral·lels?

Sempre	Molt sovint	Sovint	Poques vegades	Mai	Amb quina finalitat

Pregunta 5. Feu servir conjuntament **originals i les seves traduccions** com a textos paral·lels?

Sempre	Molt sovint	Sovint	Poques vegades	Mai	Amb quina finalitat

Pregunta 6. Quin és **el factor** que us porta a cercar textos paral·lels?

El tema del text	El gènere del text	El caràcter innovador del contingut	El grau d'especialització del contingut	L'encàrrec de traducció	Les característiques de l'emissor o el receptor	El format del text	Altres

Pregunta 7. On cerqueu habitualment textos paral·lels?

Internet	Biblioteques tradicionals (monografies especialitzades)	Revistes i altres publicacions professionals	Documentació aportada pel client	Traduccions ja fetes	Altres

Pregunta 8. Com cerqueu la documentació necessària?

Per paraules clau	Per autors representatius	Per empreses o entitats	Altres

Pregunta 9: Teniu en compte la **fiabilitat** de les fonts documentals que utilitzeu?

Sempre	Molt sovint	Sovint	Poques vegades	Mai

Pregunta 10. Marqueu aquells factors que teniu en compte per **valorar la fiabilitat** dels textos paral·lels que trobeu.

A. Que hi consti la data.	Sí
	No
B. Que la data sigui actual.	Sí
	No
C. Que hi consti l'autor.	Sí
	No
D. Que l'autoria estigui recolzada per una institució o empresa.	Sí
	No
E. Que el text estigui ben redactat o sembli un document original (i no una traducció).	Sí
	No
F. Que el text estigui publicat en format electrònic.	Sí
	No
G. Que el text estigui publicat en format paper.	Sí
	No

Capítol 1 - La documentació en el procés de la traducció

Pregunta 11. Què cerqueu o espereu trobar en els documents?

Coneixement, informació	Dades terminològiques	Fraseologia o col·locacions	Validar la terminologia que ja teniu	Altres

Pregunta 12. Conserveu les traduccions realitzades per a reutilitzar-les com a fonts d'informació?

Sempre	Molt sovint	Sovint	Poques vegades	Mai

Pregunta 13. Com conserveu les traduccions?

Pel que fa a la llengua de partida	Pel que fa a la llengua d'arribada
<input type="checkbox"/> En paper	<input type="checkbox"/> En paper
<input type="checkbox"/> En una base de dades documental	<input type="checkbox"/> En una base de dades documental
<input type="checkbox"/> En arxius de text o altres formats	<input type="checkbox"/> En arxius de text o altres formats
<input type="checkbox"/> Com a memòria de traducció	<input type="checkbox"/> Com a memòria de traducció
<input type="checkbox"/> No les conservo	<input type="checkbox"/> No les conservo

Pregunta 14. Què us aporten les traduccions fetes?

Coneixement, informació	Dades terminològiques	Fraseologia o col·locacions	Validar la terminologia que ja teniu	Altres

Pregunta 15. Com conserveu la informació que extraieu de les traduccions?

En bases de dades	En llistes o glossaris en arxius de text	Altres

Pregunta 16. En quin moment extraieu la informació de les traduccions?

Durant el procés de la traducció	Quan necessiteu la informació (després de fer la traducció)	Altres

En el proper apartat s'analitza la mostra obtinguda amb deteniment, tant de manera global com per col·lectius.³

1.1. Anàlisi dels resultats de l'enquesta⁴

Les enquestes revelen que la utilització de textos paral·lels és, en general, freqüent entre tots els col·lectius.

Així, per exemple, els traductors d'empreses acostumen a documentar-se específicament per a cada traducció a partir de textos paral·lels en la llengua d'arribada més que no pas en la llengua de partida. De fet, mentre que alguns afirmen no documentar-se mai o fer-ho poques vegades en la llengua de partida, tots ells accepten documentar-se sovint, molt sovint o sempre en la llengua d'arribada.

	Sempre	Molt sovint	Sovint	Poques vegades	Mai
Ll. de partida	5,9%	23,5%	17,6%	41,2%	5,9%
Ll. d,arribada	35,3%	35,3%	29,4%	0%	0%

Taula 1-1. Utilització dels textos paral·lels dels traductors d'empreses, resposta a la pregunta 1 (elaboració pròpia, en endavant e. p.)

La tendència similar es detecta entre els traductors de l'SdT, que utilitzen amb major assiduitat textos paral·lels en la llengua d'arribada que no pas en la llengua de partida:

	Sempre	Molt sovint	Sovint	Poques vegades	Mai
Ll. de partida	22,2%	22,2%	44,4%	0%	11,1%
Ll. d,arribada	33,3%	55,5%	11,1%	0%	0%

Taula 1-2. Utilització dels textos paral·lels dels traductors de l'SdT, resposta a la pregunta 1 (e. p.)

³ El conjunt de respostes recollides es troba a l'annex A, pàgina 579.

⁴ Totes les xifres que s'indiquen a continuació són percentatges. D'aquesta manera podem comparar les dades relatives dels diferents col·lectius entre elles.

Capítol 1 - La documentació en el procés de la traducció

I aquesta tendència es repeteix de manera més marcada en el cas dels traductors autònoms, tal com mostra la taula 1-3:

	Sempre	Molt sovint	Sovint	Poques vegades	Mai
Ll. de partida	0%	51,7%	20,7%	27,6%	0%
Ll. d'arribada	37,8%	40,6%	19,2%	1,1%	0%

Taula 1-3. Utilització dels textos paral·lels dels traductors autònoms, resposta a la pregunta 1 (e. p.)

Així doncs, es confirma la utilització de textos paral·lels com a recurs d'informació per a la traducció, i que el procés de cerca de textos s'acostuma a fer per a cada encàrrec de traducció. A continuació es mostra de manera gràfica i global aquesta tendència comuna de tots tres col·lectius.

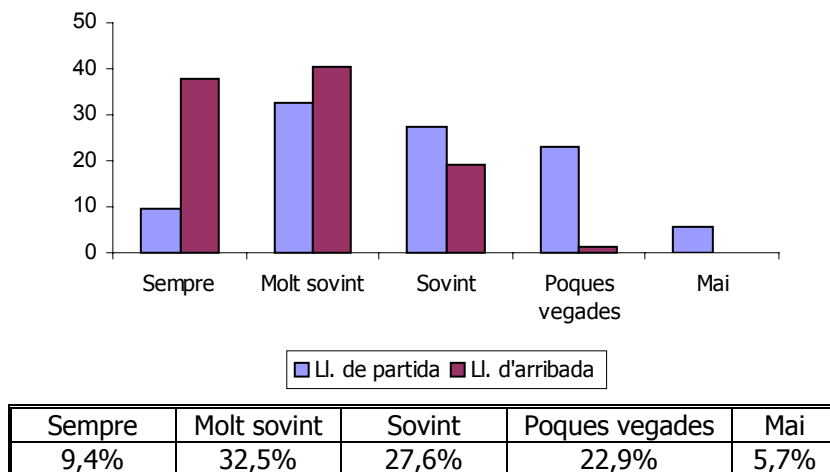


Figura 1-1. Cerques específiques de textos paral·lels en la llengua de partida, resposta a la pregunta 1 (e.p.)

1.2. La llengua dels textos paral·lels

Tal com s'ha indicat a l'apartat anterior, els traductors habitualment es documenten per a cada traducció, més en la llengua d'arribada que en la de partida. Partint d'aquest fet, ens hem interessat per la finalitat que els porta a cercar informació en textos paral·lels en totes dues llengües.

1.2.1. Textos paral·lels en la llengua de partida

Tots tres col·lectius coincideixen en el fet que cerquen tant informació lingüística com factual. Els traductors d'empreses, per exemple, obtenen informació de tots dos tipus en percentatges pràcticament equivalents:

	I. lingüística	I. factual	Lingüística i factual
Empreses	29,4%	23,5%	23,5%
SdT	55,5%	22,2%	11,1%
<i>Freelance</i>	27,6%	34,5%	24,2%

Taula 1-4. Informació que se cerca als textos paral·lels en la llengua de partida, resposta a la pregunta 2 (e. p.)

En el cas dels traductors de l'SdT, però, cerquen més dades de caire lingüístic que no pas factual, mentre que en el col·lectiu de traductors autònoms es dona la situació inversa. Si, tal com es veurà a l'apartat 1 "La documentació amb el text paral·lel en traducció" del capítol 5 (pàgina 239), la documentació factual està íntimament relacionada amb el coneixement que el traductor té del domini en què tradueix, aquesta situació cobra sentit, ja que mentre que els traductors de l'SdT acostumen a traduir regularment sobre el mateix tema, la diversitat temàtica dels encàrrecs dels traductors autònoms pot ser molt més variada. Per aquest motiu, les necessitats d'informació factual del traductor autònom poden ser més altes que les del traductor de l'SdT.

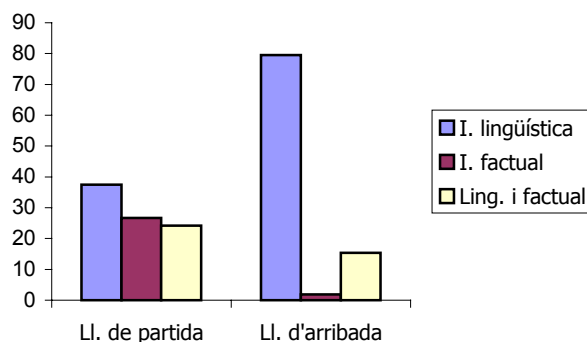
1.2.2. Textos paral·lels originals en la llengua d'arribada

D'altra banda, en el cas dels textos paral·lels en la llengua d'arribada, la finalitat és majoritàriament lingüística, tant en el cas dels traductors d'empreses com en els de l'SdT i els traductors autònoms.

	I. Lingüística	I. Factual	Lingüística i Factual
Empreses	88,2%	5,9%	0%
SdT	77,8%	0%	22,2%
<i>Freelance</i>	79,5%	0%	24,1%

Taula 1-5. Informació que se cerca als textos paral·lels en la llengua d'arribada, resposta a la pregunta 3 (e. p.)

Aquest fet fa palès, doncs, que mentre que els textos originals en la llengua de partida aporten informació factual a més de lingüística, els textos originals en la llengua d'arribada aporten pràcticament de manera exclusiva informació lingüística: denominacions de termes, expressions fraseològiques, usos restringits de paraules habituals, etc.



	I. Lingüística	I. Factual	Ling. i factual
LI. de partida	37,5%	26,7%	24,2%
LI. d,arribada	79,5%	1,9%	15,4%

Figura 1-2. Utilització de textos paral·lels en la llengua de partida i d'arribada de manera global, resposta a la pregunta 3 (e. p.)

Aquest gràfic posa de manifest les diferents aplicacions dels textos paral·lels en funció de la seva llengua. Cal assenyalar en aquest punt que la llengua materna dels traductors enquestats és el castellà i/o el català, i que la majoria tradueix de l'anglès.

1.2.3. Traduccions com a textos paral·lels en la llengua d'arribada

Els textos que el traductor té més a mà, i els que millor coneix, són les traduccions que ell mateix ha realitzat. Per aquest motiu, antigues traduccions poden acabar sent utilitzades com a textos paral·lels d'on pot extreure informació per a noves traduccions.

Sempre	Molt sovint	Sovint	Poques vegades	Mai
0%	17,6%	35,3%	35,3%	11,8%

I. Lingüística	I. Factual	Lingüística i factual	Sense resposta
11,8%	0%	0%	88,2%

Taula 1-6. Utilització de traduccions com a textos paral·lels dels traductors a empresa, resposta a la pregunta 4 (e. p.)

Tal com indica la taula anterior, els traductors a empreses utilitzen traduccions com a textos paral·lels en poques ocasions. Comparativament, aquesta mena de textos s'utilitza sensiblement menys que els textos originals en la llengua d'arribada (vegeu la taula 1-1. "Utilització dels textos paral·lels dels traductors d'empreses, resposta a la pregunta 1", pàgina 29).

Sempre	Molt sovint	Sovint	Poques vegades	Mai
22,2%	22,2%	55,5%	0%	0%

I. lingüística	I. factual	Lingüística i factual	Sense resposta
77,8%	0%	22,2%	0%

Taula 1-7. Utilització de traduccions com a textos paral·lels dels traductors de l'SdT, resposta a la pregunta 4 (e. p.)

En el cas dels traductors de l'SdT, de nou les seves circumstàncies especials de treball els porta a utilitzar traduccions fetes com a textos paral·lels. També és cert que el rang de text original que legalment tenen els textos traduïts els obliga a respectar decisions traductològiques preses anteriorment, més del que probablement es fa en altres situacions professionals.

Sempre	Molt sovint	Sovint	Poques vegades	Mai
3,4%	17,9%	17,2%	55,2%	10,3%

I. Lingüística	I. Factual	Lingüística i factual	Sense resposta
34,5%	0%	10,3%	55,2%

Taula 1-8. Utilització de traduccions com a textos paral·lels dels traductors autònoms, resposta a la pregunta 5 (e. p.)

La pràctica dels traductors autònoms en aquest cas està molt propera a la dels traductors a empreses. Utilitzen relativament poc les seves traduccions com a textos paral·lels. Tanmateix, tots tres col·lectius no cerquen informació factual a les seves traduccions, sinó fonamentalment informació lingüística.

1.2.4. Les traduccions i els seus originals com a textos d'arribada

Un ús específic que poden rebre les traduccions com a textos paral·lels és el de ser emprats conjuntament amb el seu original. En combinar original i traducció es poden cercar dades de manera bilingüe.

	Sempre	Molt sovint	Sovint	Poques vegades	Mai
Empresa	5,9%	17,6%	24,5%	52,9%	0%
SdT	33,3%	33,3%	22,2%	11,1%	0%
Autònoms	3,4%	6,9%	17,2%	51,7%	13,8%

I. lingüística	I. factual	Lingüística i factual	Sense resposta
58,8%	5,9%	0%	35,3%
100%	0%	0%	0%
34,5%	0%	6,9%	58,6%

Taula 1-9. Utilització de traduccions i originals com a textos paral·lels, resposta a la pregunta 5 (e. p.)

Igual que en el cas anterior, la utilització de textos traduïts, tenint en compte el seu original, és relativament baixa. De nou, els traductors a empreses i els autònoms utilitzen menys aquest recurs que els traductors de l'SdT.

El gràfic següent posa en relleu les lleugeres diferències que s'observen en la utilització de traduccions i traduccions amb el seu original com a textos paral·lels de manera global.

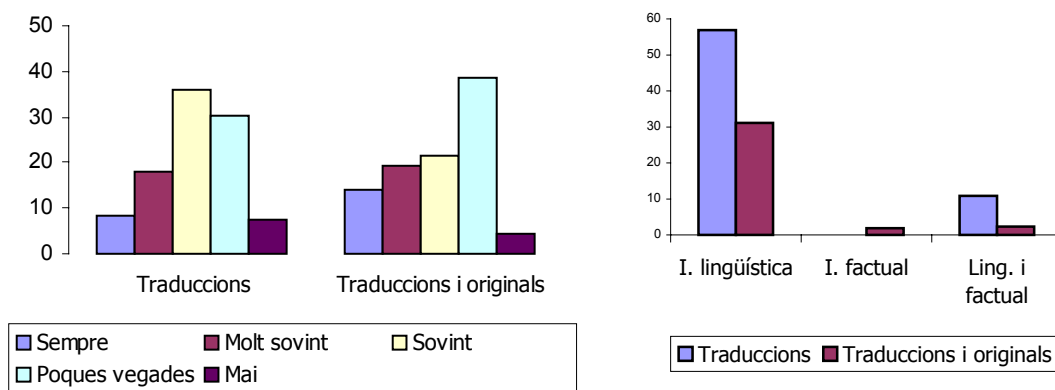


Figura 1-3. Utilització de traduccions i de traduccions amb original, resposta a la pregunta 5 (e. p.)

1.3. Factors que aboquen el traductor a la cerca de textos paral·lels

Pel que fa als factors que fan que el traductor cerqui textos paral·lels, tots tres col·lectius mostren comportaments similars.

	El tema del text	El gènere del text	El caràcter innovador del contingut	El grau d'especialització del contingut	L'encàrrec de traducció	Les característiques de l'emissor o el receptor	El format del text	Altres
Empresa	58,5%	29,4%	58,8%	100%	17,6%	41,2%	17,6%	5,9%
SdT	77,8%	22,2%	55,5%	77,8%	11,1%	11,1%	22,2%	11,1%
Autònoms	55,25	27,6%	44,8%	96,5%	6,9%	13,8%	17,2%	17,2%

Taula 1-10. Factors que provoquen la cerca de textos paral·lels, resposta a la pregunta 6 (e. p.)

Com es pot comprovar, el factor que més es té en compte és el grau d'especialització del contingut, seguit pel tema del text i el seu caràcter innovador. Aquests dos factors estan molt vinculats: tots dos fan referència a l'àmbit temàtic en què s'emmarca la traducció. El primer, però, és un factor relatiu: el grau

d'especialització del contingut es calcula en relació amb el coneixement que el traductor té sobre aquell àmbit, i és la possible diferència entre el nivell de coneixements del traductor i el del text el que provoca la cerca d'altres textos paral·lels.

El caràcter innovador del text és el tercer dels factors que els traductors han considerat més importants. Sovint, les diferents obres lexicogràfiques poden no oferir informació nova o actual, per la qual cosa els textos paral·lels es converteixen en el recurs més útil en aquests casos.

A continuació es mostra de manera gràfica la relació de factors pels quals els traductors, de manera global, cerquen textos paral·lels:

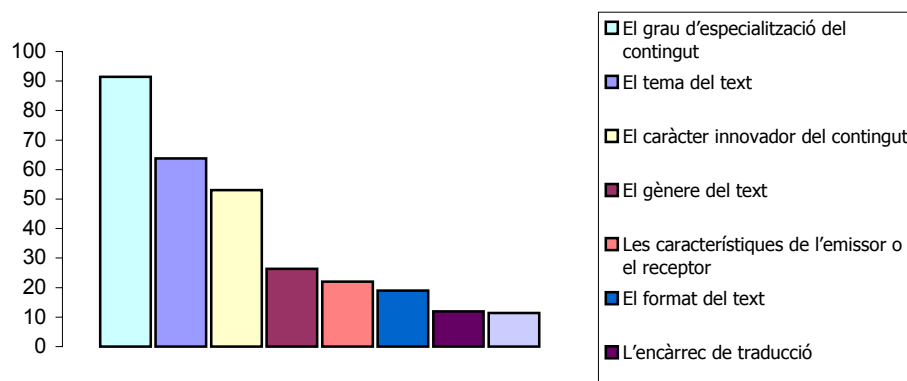


Figura 1-4. Factors globals que provoquen la cerca de textos paral·lels, resposta a la pregunta 6 (e. p.)

1.4. Metodologia de cerca

Un cop establertes les similituds i diferències pel que fa a l'ús de textos paral·lels dels col·lectius de professionals enquestats, un altre dels aspectes que hem tingut en compte en aquest breu estudi fa referència a la manera com obtenen aquests textos. L'objectiu, novament, és determinar quines són les fonts més utilitzades en aquest sentit i la manera d'obtenir els textos, i també si, de nou, es poden establir diferències significatives entre els diferents col·lectius de traductors.

1.4.1. Fonts de textos paral·lels

Tots els traductors enquestats declaren que fan servir Internet quan cerquen textos paral·lels. Internet és, de fet, l'única font de textos paral·lels en què coincideixen tots els traductors.

	Internet	Biblioteques tradicionals (monografies especialitzades)	Revistes i altres publicacions professionals	Documentació aportada pel client	Traduccions ja fetes	Altres
Empresa	100%	17,6%	29,4%	82,3%	70,6%	11,8%
SdT	100%	11,1%	33,3%	22,2%	77,8%	22,2%
Autònoms	100%	44,8%	31%	68,9%	62%	17,2%

Taula 1-11. Fonts de textos paral·lels, resposta a la pregunta 7 (e. p.)

Tanmateix, en la segona de les fonts més consultades no coincideixen els tres col·lectius: mentre que per als traductors a empreses i els autònoms la segona font d'informació és la documentació facilitada pel client, per als traductors de l'SdT ho són les traduccions ja fetes, que també són una font força rellevant per als dos col·lectius anteriors. Aquesta discrepància entre col·lectius es deu, de nou, a les peculiaritats de l'entorn de treball dels traductors de l'SdT (17,6% i 11,1% respectivament).

També resulta remarcable l'ús relativament alt que fan els traductors autònoms de les biblioteques tradicionals (el 44,8%), que no coincideix amb les pràctiques dels traductors a empreses ni dels de l'SdT.

Finalment, cal destacar la rellevància de les biblioteques personals, recollides a l'apartat "Altres", entre els traductors a empreses i els autònoms. Novament, la pràctica dels traductors de l'SdT en aquest cas també es desmarca de la dels altres dos col·lectius, ja que en l'apartat "Altres" recullen fonamentalment fonts documentals propis de les institucions per a les quals tradueixen, així com legislació pròpia de la UE i dels països membres.

En resum, les fonts a què recorren els traductors professionals per ordre de rellevància es distribueixen tal com recull el gràfic següent:

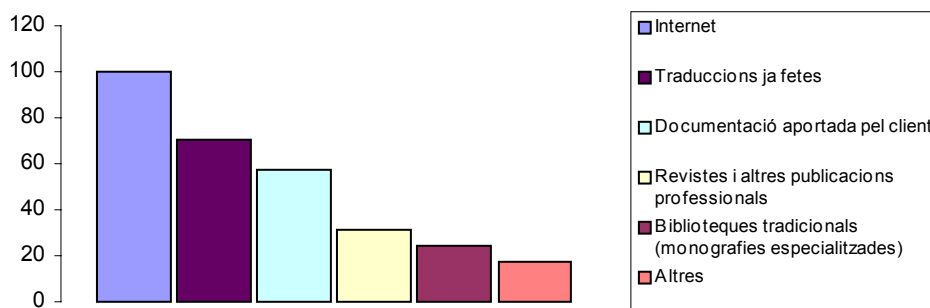


Figura 1-5. Fonts de textos paral·lels per ordre de rellevància, resposta a la pregunta 7 (e. p.)

1.4.2. Procés de cerca dels textos paral·lels

A més de les fonts que generalment s'utilitzen en la cerca de textos paral·lels, per a l'estudi que durem a terme a continuació també resulta rellevant conèixer com arriben fins als textos així com si es té en compte la seva fiabilitat (aspecte aquest que es tracta amb més detall al subapartat següent).

Tal com revela la taula 1-12, el sistema de cerca de documents més utilitzat per tots tres col·lectius de traductors professionals és el de la cerca per paraules clau. Això indica que la cerca dels textos paral·lels es fa a partir de paraules representatives de l'àmbit de la traducció que estan duent a terme, o bé a partir de les expressions que els resulten problemàtiques, ja sigui perquè no les entenen en tota la seva amplitud o perquè desconeixen com expressar-les en la llengua d'arribada. Alguns traductors, en tots tres col·lectius, han recollit aquesta darrera opció dins del camp "Altres".

	Per paraules clau	Per autors representatius	Per empreses o entitats	Altres
Empresa	94,1%	11,8%	58,8%	35,3%
SdT	100%	11,1%	22,2%	22,2%
Autònoms	93,1%	17,2%	44,8%	26,1%

Taula 1-12. Mètodes de cerca de textos paral·lels, resposta a la pregunta 8

El segon dels elements més emprats a l'hora de fer una cerca de textos paral·lels és el nom d'empreses o entitats rellevants o representatius de l'àmbit en què s'emmarca la traducció que estan duent a terme. Tanmateix, aquest aspecte resulta menys important entre els traductors de l'SdT que no pas entre els altres dos col·lectius de traductors.

El camp "Altres", en aquest cas, recull possibilitats de diversa mena: d'una banda, tal com ja s'ha indicat, recull la possibilitat de cercar a partir d'expressions desconegudes, encara que no siguin pròpiament paraules clau; d'altra banda, també s'indiquen altres possibilitats de cerca, com són la cerca temàtica o la consulta a experts. En resum, els elements a què es recorre per dur a terme una cerca de textos paral·lels per ordre d'utilització són els següents.

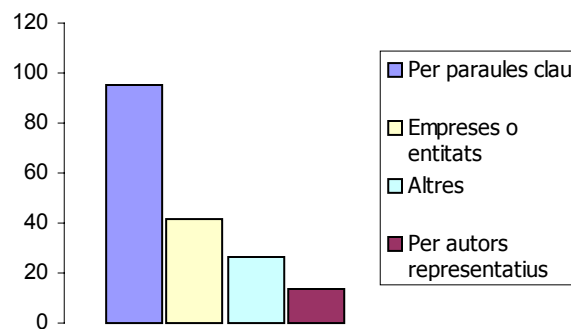


Figura 1-6. Mètodes de cerca de textos paral·lels per ordre de rellevància, resposta a la pregunta 8 (e. p.)

1.4.3. La fiabilitat dels textos paral·lels

Tal com es veurà a l'apartat 1 "La documentació amb el text paral·lel en traducció" del capítol 5 (pàgina 239), els textos paral·lels han de ser necessàriament fiables. Tanmateix, però, la fiabilitat dels textos extrets d'Internet, la font documental més utilitzada pels traductors segons s'ha indicat anteriorment, sovint no està garantida

(vegeu l'apartat 1.1 "Avaluació de la informació al *World Wide Web*" del capítol 3, pàgina 120).

És per aquest motiu que hem demanat als traductor si tenen en compte la fiabilitat dels documents que empren com a textos paral·lels, així com en quins factors es basen per calcular-la.

Els traductors gairebé sempre tenen en compte la fiabilitat dels seus textos paral·lels, tal com mostra el gràfic següent:

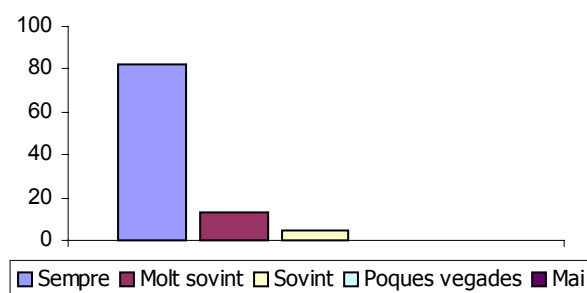


Figura 1-7. Valoració global de la fiabilitat dels textos paral·lels, resposta a la pregunta 9 (e. p.)

Pel que fa als aspectes que denoten la fiabilitat del text paral·lel, en l'enquesta proposem que es pronunciïn sobre els següents:

- * Que hi consti la data.
- * Que sigui actual.
- * Que hi consti l'autor.
- * Que l'autoria estigui recolzada per una institució o empresa (el que, en l'àmbit de la publicació digital a Internet, es coneix com a responsable secundari)⁵.
- * Que el text estigui ben redactat o sembli un document original (i no una traducció).
- * Que el text estigui publicat en format electrònic.

⁵ Vegeu l'apartat **iError! No se encuentra el origen de la referencia.** "iError! No se encuentra el origen de la referencia." del capítol 3, pàgina **iError! Marcador no definido..**

- * Que el text estigui publicat en format paper.

	Que hi consti la data		Que sigui actual	
	Sí	No	Sí	No
Empresa	88,2%	11,8%	100%	0%
SdT ⁶	66,7%		88,9%	11,1%
Autònoms	65,5%	34,5%	79,3%	20,7%

Taula 1-13. Factors que denoten fiabilitat. data i actualitat, respostes a les preguntes 10 A i 10 B (e. p.)

Tal com es desprèn de la taula anterior, el que més interessa al conjunt dels traductors és que el text paral·lel sigui d'actualitat, tant o més que el fet que hi consti la data de manera explícita. Aquest factor és, doncs, un dels que més es té en compte en el moment de decidir la utilització d'un document com a text paral·lel.

	Que hi consti l'autor		Que l'autoria estigui recolzada per una institució o empresa	
	Sí	No	Sí	No
Empresa	47,1%	52,9%	88,2%	11,8%
SdT	88,9%	11,1%	89,7%	10,3%
Autònoms	63,2%	36,8%	88,9%	11,1%

Taula 1-14. Factors que denoten fiabilitat. l'autoria i el responsable secundari, respostes a les preguntes 10C i 10D (e. p.)

Pel que fa a l'autor, la seva rellevància resulta desigual entre els diferents factors: mentre que per als traductors de l'SdT és un dels factors més destacables, per als traductors autònoms i sobretot per als que treballen a empreses no resulta tan important. Tot i així, tots tres col·lectius atorguen un alt valor al responsable secundari com a indicador de la fiabilitat del text, superior en tots tres casos que al fet que hi consti l'autor de manera explícita.

D'altra banda, i com que sovint no és una tasca senzilla establir si un text és un original o una traducció (sobretot a Internet), hem volgut demanar als traductors fins a

⁶ El 33,3% dels enquestats d'aquest col·lectiu van deixar en blanc aquesta qüestió.

Capítol 1 - La documentació en el procés de la traducció

quin punt rebutgen un document com a text paral·lel segons la qualitat de la seva redacció. Els resultats han estat els següents:

	Que el text estigui ben redactat o sembli un document original (i no una traducció)	
	Sí	No
Empresa	92,8%	7,1%
SdT ⁷	22,2%	
Autònoms	85%	15%

Taula 1-15. Factors que denoten fiabilitat. la qualitat del text, resposta a la pregunta 10E (e. p.)

De nou es demostra que aquest també és un valor important a l'hora d'establir el grau de fiabilitat d'un text, segons els traductors professionals.

Finalment, vam voler determinar si el format del text, el fet que hagi estat publicat en format digital o en format paper, afecta la percepció de la fiabilitat dels traductors.

	Que el text hagi estat publicat en format electrònic		Que el text hagi estat publicat en format paper	
	Sí	No	Sí	No
Empresa	35,7%	64,2%	14,3%	85,7%
SdT ⁸	22,2%		11,1%	
Autònoms	27,8%	72,2%	30%	70%

Taula 1-16. Factors que denoten fiabilitat. el format de publicació, resposta a les preguntes 10F i 10G (e. p.)

Segons les dades recollides a la taula anterior, ni el suport tradicional ni el digital denoten fiabilitat, de la qual cosa es desprèn que el suport del text no produeix

⁷ El 77,8% dels enquestats van deixar en blanc aquesta qüestió.

⁸ El 77,8% i el 88,9% dels enquestats, respectivament, van deixar en blanc aquestes qüestions.

cap efecte sobre la percepció de la fiabilitat que en rep el traductor. No és aquest, doncs, un factor determinant que permeti discriminar un text en favor d'uns altres, sinó que els factors més determinants són que sigui d'actualitat i el seu responsable secundari, tal com queda palès al gràfic següent:

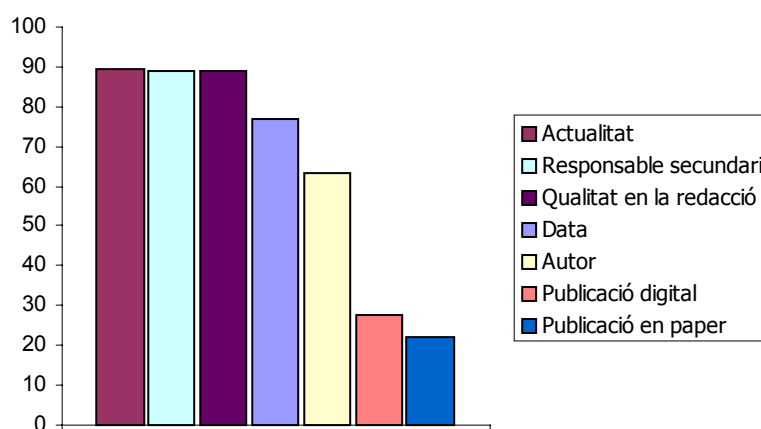


Figura 1-8. Factors que denoten fiabilitat, resposta a la pregunta 10 (e. p.)

1.5. La utilitat dels textos paral·lels

Un cop establerta la metodologia que els traductors segueixen per escollir els seus textos paral·lels, hem volgut aprofundir en la informació que esperen obtenir-ne per tal de determinar la utilitat.

	Coneixement, informació	Dades terminològiques	Fraseologia o col·locacions	Validar la terminologia que ja teniu	Altres
Empresa	88,2%	88,2%	70,6%	94,1%	11,8%
SdT	77,8%	88,9%	44,4%	100%	11,1%
Autònoms	82,7%	89,6%	75,9%	93,1%	3,4%

Taula 1-17. Dades que el traductor espera trobar al text paral·lel, resposta a la pregunta 11 (e. p.)

Tal com es pot comprovar amb la taula anterior, els textos paral·lels s'utilitzen per obtenir dades de diversa mena, des de dades factuais o coneixement, fins a dades de caràcter lingüístic. Crida l'atenció, però, que l'objectiu més rellevant sigui el de validar terminologia que ja es té, ja sigui perquè s'ha aconseguit per altres canals o perquè es vol validar una intuïció lingüística del traductor. Els textos paral·lels resulten ser, sobretot, recursos per a la validació d'informació en contextos lingüístics reals.

D'altra banda, a l'apartat "Altres" tots tres col·lectius de traductors assenyalen la cerca de cites literals com a objectiu de la utilització de textos paral·lels.

1.6. Reutilització de les traduccions realitzades

Aprofitant aquesta enquesta, hem volgut conèixer més profundament els hàbits dels traductors pel que fa a la conservació i la reutilització de les seves pròpies traduccions, ja que són els textos que més fàcilment té al seu abast i els que millor coneix.

1.6.1. Conservació de les traduccions

El primer aspecte pel qual ens vam interessar fou si acostumaven a conservar les seves traduccions. Aquestes foren les respostes:

	Sempre	Molt sovint	Sovint	Poques vegades	Mai
Empresa	88,2%	5,9%	5,9%	0%	0%
SdT	44,4%	22,2%	22,2%	11,1%	0%
Autònoms	68,3%	13,9%	11,7%	6%	0%

Taula 1-18. Conservació de les traduccions realitzades, resposta a la pregunta 12 (e. p.)

Tal com es podia suposar prèviament, la majoria dels traductors conserven les seves traduccions. En el cas dels traductors de l'SdT, les conserven en un percentatge menor que els altres col·lectius, probablement per raó de la facilitat amb què poden recuperar les seves pròpies traduccions i les d'altres traductors mitjançant els seus sistemes de documentació. Crida l'atenció, però, que un 6% dels traductors autònoms conservi les seves traduccions en poques ocasions.

A més de saber si conservaven les traduccions, ens vam interessar pel format amb què ho feien, i si també conserven els originals.

Text de partida					Traducció				
En paper	En una base de dades documental	En arxius de text o altres formats	Com a memòria de traducció	No els conservo	En paper	En una base de dades documental	En arxius de text o altres formats	Com a memòria de traducció	No els conservo
47,1%	29,4%	94,1%	52,9%	5,9%	29,4%	29,4%	100%	52,9%	0%

Taula 1-19. Conservació de traducció i original dels traductors a empreses, resposta a la pregunta 13 (e. p.)

De la taula anterior es desprèn que el col·lectiu dels traductors a empreses sempre conserven les seves traduccions en format digital, i en més del 50% de les ocasions ho fan conjuntament amb el seu original en forma de memòria de traducció. També es conserven per parells de llengües en bases de dades documentals. Altrament, el percentatge de textos originals que es conserven en format paper resulta més elevat que el de les traduccions, probablement perquè encara és habitual rebre els originals en aquest suport, i no en format digital. A més, sorprèn el fet que en pràcticament un 6% de les ocasions no es conservi l'original.

Text de partida					Traducció				
En paper	En una base de dades documental	En arxius de text o altres formats	Com a memòria de traducció	No els conservo	En paper	En una base de dades documental	En arxius de text o altres formats	Com a memòria de traducció	No els conservo
44,4%	33,3%	88,9%	55,5%	0%	33,3%	33,3%	88,9%	55,5%	0%

Taula 1-20. Conservació de traducció i original dels traductors de l'SdT, resposta a la pregunta 13 (e. p.)

Capítol 1 - La documentació en el procés de la traducció

En el cas del col·lectiu de traductors de l'SdT, segons les dades recollides, la conservació d'original i traducció canvia lleugerament respecte de la pràctica dels traductors a empreses. Així doncs, la taula 1-20 revela que, de nou, més del 50% de les traduccions es conserven en forma de memòria de traducció juntament amb el seu original. Tanmateix, no totes les traduccions es desen en format digital; de fet, en un percentatge lleugerament superior que en el cas anterior es conserven en suport paper i en bases de dades documentals, com també els seus originals. Destaca el fet que sempre es conservin els originals.

Text de partida					Traducció				
En paper	En una base de dades documental	En arxius de text o altres formats	Com a memòria de traducció	No els conservo	En paper	En una base de dades documental	En arxius de text o altres formats	Com a memòria de traducció	No els conservo
37,9%	10,3%	82,7%	31%	6,9%	17,2%	13,8%	93,1%	31%	3,4%

Taula 1-21. Conservació de traducció i original de traductors autònoms, resposta a la pregunta 13 (e. p.)

Les pràctiques dels traductors autònoms són sensiblement diferents de les dels dos col·lectius anteriors. D'una banda, aquest col·lectiu sembla que recorre als sistemes de memòria de traducció en un percentatge menor que els anterior (únicament el 31% de les ocasions), i tampoc no utilitza tan sovint bases de dades documentals per conservar les seves traduccions i els originals. També cal destacar el fet que en un baix percentatge no conservi ni l'original ni la traducció⁹.

El gràfic següent mostra de manera conjunta els resultats de les enquestes:

⁹ En una enquesta s'assenyala que aquest fet es deu a la impossibilitat de conservar tant original com la traducció per motius de confidencialitat.

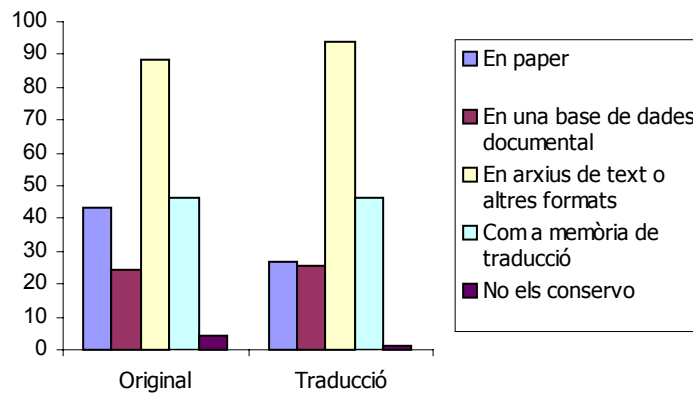


Figura 1-9. Conservació de traducció i original, resposta a la pregunta 13 (e. p.)

Tal com mostra aquest gràfic, la conservació d'originals i traduccions resulta molt similar. L'únic aspecte destacable és el fet que es conservin més originals en suport paper que no pas traduccions, probablement perquè de vegades no es facilita al traductor l'original en format digital.

1.6.2. Informació que aporten les traduccions realitzades

A més de conèixer les pràctiques de conservació de traduccions i originals, també ens vam voler interessar per les característiques de la informació que els traductors obtenen a partir de les seves pròpies traduccions. Així doncs, i per tal de poder comparar les dades amb les recollides a l'apartat 1.5 "La utilitat dels textos paral·lels" (pàgina 44) sobre els textos paral·lels originals, vam plantejar la mateixa qüestió que en aquell cas, i aquesta ha estat la resposta obtinguda:

	Coneixement, informació	Dades terminològiques	Fraseologia o col·locacions	Validar la terminologia que ja teniu	Altres
Empresa	64,7%	94,1%	64,7%	70,6%	5,9%
SdT	66,7%	88,9%	55,5%	44,4%	0%
Autònoms	68,9%	89,7%	63,1%	59%	6,6%

Taula 1-22. Dades que el traductor espera trobar a les seves traduccions, resposta a la pregunta 14 (e. p.)

Mentre que, en aquest cas, la cerca d'informació factual resulta sensiblement més baixa que en el cas dels textos paral·lels originals (vegeu la taula 1-17. "Dades que el traductor espera trobar al text paral·lel, resposta a la pregunta 11", pàgina 44), el percentatge de cerca de dades terminològiques augmenta. Pel que fa als altres dos aspectes tinguts en compte, la cerca de fraseologia o col·locacions i la validació de terminologia, també presenten dades lleugerament més baixes en el cas de la utilització de traduccions com a textos paral·lels.

Altrament, hem de recalcar un dels aspectes recollits a l'apartat "Altres", segons el qual les traduccions s'utilitzen com a textos paral·lels a fi de vetllar per una coherència en la producció de noves traduccions sobre un tema similar per a un mateix client.

1.6.3. Conservació de la informació extreta de les traduccions realitzades

Si, tal com hem vist, les traduccions es conserven i s'utilitzen com a textos paral·lels per a noves traduccions, també hem volgut conèixer en quin moment s'extreuen les dades que poden ser rellevants per a futures traduccions i com s'emmagatzemen.

La informació que es pot obtenir a partir de les traduccions realitzades generalment és producte d'una cerca terminològica puntual (vegeu l'apartat 2 "La terminologia i els textos paral·lels: la cerca d'unitats de sentit especialitzat" del capítol

5, pàgina 253) o bé d'una decisió de traducció que cal recordar per tal de mantenir la coherència en futures traduccions. El primer dels aspectes que abordarem és el format amb què els traductors conserven la informació que extreuen de les seves traduccions.

	En bases de dades	En llistes o glossaris, en arxius de text	Altres
Empresa	47,1%	70,6%	11,8%
SdT	77,8%	77,8%	0%
Autònoms	24,1%	79,3%	3,4%

Taula 1-23. Conservació de la informació obtinguda a partir de traduccions, resposta a la pregunta 15 (e. p.)

Tal com revela la taula anterior, tant el col·lectiu de traductors a empreses com els autònoms prefereixen conservar aquesta informació en llistats o glossaris en arxius digitals de text, com ara de processadors de textos. En el cas dels traductors a empreses, també cal assenyalar el fet que poden no recollir aquesta informació de manera explícita i conservar-la únicament en els textos que formen part de les seves memòries de traducció (aspecte recollit a "Altres"). D'altra banda, un percentatge molt baix de traductors autònoms, el 3,4%, afirma que conserva aquesta informació en suport paper.

De nou, els traductors de l'SdT es diferencien de la resta en el fet que conserven la informació obtinguda en bases de dades de manera més habitual que els altres dos col·lectius de professionals.

D'altra banda, també resulten reveladores les dades relatives al moment en què s'extreu informació de les traduccions per conservar-la de manera independent:

	Durant el procés de la traducció	En el moment que es necessita la informació (després de fer la traducció)	Altres
Empresa	82,3%	41,2%	5,9%
SdT	100%	55,5%	11,1%
Autònoms	82,7%	41,4%	3,4%

Taula 1-24. Moment en què es desa la informació obtinguda a partir de les traduccions, resposta a la pregunta 16 (e. p.)

Tots tres col·lectius coincideixen a realitzar aquesta operació durant la traducció, de manera que a mesura que es resolen problemes lingüístics o terminològics les solucions adoptades també es recullen als glossaris corresponents. De tota manera, tant els traductors a empreses com els autònoms reconeixen, sobretot en el cas de fer-ho un cop finalitzada la traducció, que sovint no disposen de prou temps per a realitzar aquesta operació.

A l'apartat "Altres" es recullen pràctiques addicionals, com la de realitzar aquesta tasca abans de traduir un nou text, després de fer-ne una primera lectura (en el cas dels traductors a empresa), o bé ocasionalment en llegir revistes o publicacions en diferents llengües (traductors de l'SdT). En el cas dels traductors autònoms, afirmen que en un baix percentatge de les ocasions aquesta informació mai no s'arriba a recollir.

1.7. Resultats globals

En resum, la mostra recollida sobre les pràctiques de documentació mitjançant textos paral·lels de traductors a empreses, a l'SdT i autònoms posa de manifest els fets següents:

- * Els textos paral·lels són principalment textos originals en la llengua de partida i, en segon lloc, textos originals en la llengua d'arribada, mentre que les traduccions representen un percentatge inferior.
- * D'aquests textos s'extreu més informació lingüística que factual.
- * La font de textos paral·lels a la qual més recorren els traductors és Internet.
- * Les cerques de documents es fan majoritàriament per paraules clau.
- * Els traductors tenen en compte la fiabilitat dels documents que faran servir com a textos paral·lels.
- * La percepció de la fiabilitat dels textos es basa fonamentalment en el seu nivell d'actualitat i en la presència d'un responsable secundari representatiu.

- * Els textos paral·lels s'utilitzen com a textos paral·lels sobretot per validar informació que el traductor ja té, a més d'obtenir coneixement i dades de caràcter lingüístic.
- * Els traductors conserven l'original i la traducció sobretot en format electrònic, i gairebé en la meitat de casos en forma de memòria de traducció.
- * Les traduccions també s'utilitzen com a textos paral·lels, sobretot per obtenir-ne dades terminològiques.
- * El suport més utilitzat per desar la informació obtinguda amb les traduccions és el d'arxius de text.
- * Aquesta informació s'extreu principalment durant el procés de traducció.

Els resultats obtinguts, sobretot pel que fa a la utilització massiva d'Internet com a font de recursos textuais malgrat la manca de qualitat de gran part dels continguts que s'hi troben (vegeu l'apartat 1.1 "Avaluació de la informació al *World Wide Web*" del capítol 3, pàgina 120), com també la utilitat d'establir criteris que permetin determinar la fiabilitat dels documents en aquest entorn, corroboren l'interès suscitat en l'autora d'aquesta tesi que es fa palès en aquest estudi. A més, el fet que els textos obtinguts a Internet ja es trobin en format digital ens permetrà desenvolupar una metodologia d'anàlisi i explotació dels textos paral·lels a partir de la metodologia proposada per la lingüística de corpus, tal com es veurà al capítol 6.