

Eines informàtiques de suport a la traducció: tendències actuals i camps futurs de desenvolupament

Xavier Arderiu
Traductor

En aquest article intentaré exposar breument l'estat actual d'aplicació de les tecnologies informàtiques en l'àmbit professional de la traducció. Repassaré breument els principals programes del mercat, la forma com interactuen i finalment proposaré algunes possibles tendències d'actuació futura en aquest camp.

1. Les tecnologies actuals

En l'actualitat, podem afirmar que les tecnologies relacionades amb les memòries de traducció i la gestió de terminologia tenen un grau d'aplicació elevat en el sector professional de la traducció.

Pel que fa a les memòries de traducció, les empreses, els professionals independents i els organismes públics o privats amb departaments interns de traducció i documentació tenen al seu abast diferents programes existents en el mercat (alguns dels més estesos a casa nostra serien TRADOS Translator's Workbench, STAR Transit, ATRIL Déjà Vu).¹ En alguns casos utilitzen les seves pròpies eines internes, com en el cas del programa IBM TranslationManager/2, que de ser pioner en el camp dels programes gestors de memòries de traducció ha passat a quedar gairebé exclusivament restringit a un ús intern per part dels diferents centres de traducció d'IBM repartits pel món, o fins i tot el cas de Microsoft, que progressivament està incorporant els programes de TRADOS (empresa de la qual posseeix un percentatge important de les accions) com a eines estàndards de treball. Ens trobem també amb el cas de grans proveïdors de serveis de traducció, a escala internacional, que han optat per desenvolupar les seves pròpies solucions de memòries de traducció: SDL, amb el seu programa SDLX, Alpnet, amb el seu programa WinJoust, de poca difusió externa, o Lionbridge, amb el ForeignDesk.

1. A l'article «Catàleg de sistemes de memòries de traducció», de Pilar SÁNCHEZ-GUJÓN, publicat a la revista *Tradumàtica* [en línia] <<http://www.fti.uab.es/tradumatica/revista/sumari/sumari.htm>>, trobareu una llista detallada dels programes gestors de memòries de traducció que hi ha disponibles al mercat.

Tots aquests programes incorporen eines que permeten reaprofitar traduccions antigues realitzades sense memòria de traducció, sempre que es conservi en suport informàtic la versió original i la traduïda. Són el que anomenem programes d'alineació, que aparellen cada segment del document original amb la seva traducció corresponent, i emmagatzemen els parells en una base de dades, que directament o posteriorment es convertirà en memòria de traducció.

TRADOS, amb el seu programa Translator's Workbench, va implantar la seva hegemonia en el mercat a partir de mitjan anys 90, i encara la manté; els principals competidors, STAR Transit i ATRIL Déjà Vu, malgrat ser productes d'una qualitat perfectament comparable, no han assolit la quota de mercat de TRADOS a escala mundial, tot i que sembla que progressivament guanyen adeptes.

Els programes de gestió de terminologia han estat molt relacionats tradicionalment amb el desenvolupament de les memòries de traducció. De fet, podem dir que els principals programes de gestió de memòries de traducció tenen el seu corresponent programa gestor de terminologia, per treballar de forma interactiva. Així podríem parlar de TRADOS MultiTerm, Transit TermStar i també del gestor de terminologia incorporat al programa ATRIL Déjà Vu. Sembla que en aquests últims anys aquestes solucions han perdut una mica de protagonisme davant la progressiva implantació dels gestors de memòries de traducció. Així, TRADOS ha estat endarrerint durant bastants anys la revisió de la seva versió de MultiTerm, afavorint el desenvolupament de Translator's Workbench, i podem dir que en l'actualitat l'elecció dels gestors de terminologia en l'àmbit professional generalment se subordina a l'elecció del programa de gestió de memòries de traducció.

Finalment, cal simplement comentar que qualsevol de les tecnologies, i més específicament els programes esmentats anteriorment, suporten sense cap mena de problema el treball en català.

2. Les tecnologies emergents

La implementació d'aquestes dues tecnologies (memòries de traducció i gestió de terminologia), és complementada per altres tipus d'aplicacions, que les diferents corporacions implicades en el sector de la traducció tenen també al seu abast. Esmentarem ràpidament l'existència progressiva de programes destinats a facilitar o millorar la gestió dels projectes, específics per al sector de la traducció, com el programa LTC Organiser, de Language Technology Centre, o el nou mòdul de gestió de projectes incorporat a l'última versió de TRADOS Translator's Workbench, anomenat Workspace.

Pel que fa a la traducció automàtica, sota aquesta denominació apleguem totes les tecnologies en les quals a partir d'un text original i amb la intervenció exclusiva de sistemes informàtics obtenim un text traduït. Amb unes bases teòriques i pràctiques molt anteriors a les de les memòries de traducció, les diferents iniciatives d'investigació en el camp de la traducció

automàtica han passat per diverses etapes de desenvolupament al llarg dels anys. Inicialment, la traducció automàtica es va percebre com una possible amenaça per als traductors, però en l'actualitat s'ha consolidat com un possible complement a la seva feina. Podem afirmar que les memòries de traducció han agafat el protagonisme que originàriament se suposava que assumiria la traducció automàtica. A escala internacional, i entre els productors multilingües, cal destacar les empreses Sail-Labs, empresa alemanya amb un centre de desenvolupament de parells de llengües a Barcelona, SYSTRAN, empresa francoamericana pionera en les tecnologies de traducció automàtica, amb més de trenta anys d'activitat, Softissimo, propietari de la tecnologia i el grup de productes Reverso, i IBM, que continua desenvolupant iniciatives en aquest camp, a través de la seva empresa Alphaworks i el seu centre de desenvolupament i recerca a Nova York, que recentment ha incorporat la seva tecnologia al conjunt de programes WebSphere. Cal esmentar, per la qualitat especial del seu producte, en l'àmbit europeu, l'empresa alemanya LinguaTec, hereva de la tecnologia d'IBM i propietària del software Personal Translator, reconegut àmpliament com el millor paquet de software de traducció automàtica, amb la limitació de comptar amb dues úniques combinacions de llengües (alemany-anglès i alemany-francès).

A l'àrea lingüística de parla catalana ens trobem amb quatre models diferents adreçats a elaborar solucions informàtiques de traducció automàtica bidireccional català-castellà. Es tracta de les solucions de Sail Labs, que funcionen en una plataforma Sun, la iniciativa Internostrum de la Universitat d'Alacant, accessible mitjançant un servidor d'Internet, el programa Salt de la Generalitat Valenciana (amb el seu programa corresponent adaptat al dialecte oriental, Autotrad Ara), en plataforma Windows, i la tecnologia utilitzada pel diari *El Periódico*, vinculada comercialment amb l'empresa AutomaticTrans. Les proves realitzades amb alguns d'aquests sistemes mostren que el nivell de qualitat final de la traducció és molt correcta en tots ells, facilitada enormement per la proximitat estructural d'ambdues llengües. De cara a usuaris professionals, podríem aproximar un grau d'error (frases traduïdes no utilitzables) del 15%, quan en la majoria de combinacions de llengües europees ens movem en uns graus d'error acceptables propers al 25-30%.

3. Relació entre traducció automàtica i memòries de traducció

Pel que fa a la relació entre les memòries de traducció i la traducció automàtica, bona part dels programes gestors de memòries de traducció incorporen la possibilitat d'interactuar d'alguna manera amb la traducció automàtica. Alguns dels programes de traducció automàtica permeten també que l'usuari emmagatzemi el producte d'una traducció feta per la màquina en memòries de traducció (Sail Labs, LinguaTec).

Cal destacar també que els grans grups internacionals de serveis relacionats amb la traducció han començat a interessar-se activament pel camp de la traducció automàtica. Així, SDL ha adquirit recentment l'empresa americana Transparent Language per incorporar la tecnologia

Transcend al seu programa de gestió de memòries de traducció SDLX, amb la qual cosa pot oferir a grans clients un conjunt de solucions integrades que inclou traducció automàtica, memòries de traducció i serveis de postedició, sense la intervenció d'altres empreses en el procés. Berlitz col·labora activament amb SYSTRAN i Lionbridge ha signat recentment un extens acord de col·laboració amb Sail Labs.

4. Qüestions pendents i futures tendències

Com he comentat anteriorment, la utilització de les memòries de traducció en el sector professional està ben assentada, i en aquest sentit la principal incògnita de cara al futur se centra en els possibles programes i estàndards que dominaran el mercat. Pel que fa als estàndards, sembla que el TMX, reconegut com a estàndard d'intercanvi de memòries de traducció, s'acabarà imposant, tot i que la seva implantació, avui per avui, encara no està generalitzada. Pel que fa als programes, sembla difícil que TRADOS perdi fàcilment la seva hegemonia, tot i que sembla que STAR i ATRIL retallaran progressivament la seva quota de mercat.

La situació de l'ús de la traducció automàtica és ben diferent, i ofereix una sèrie d'incògnites que alhora es converteixen en possibles camps de desenvolupament.

Els grans problemes que els usuaris professionals han trobat en la traducció automàtica han estat, habitualment, els següents:

a) **Qualitat.** Les expectatives elevades de qualitat dels potencials usuaris fan que tradicionalment es consideri que la qualitat general de les traduccions generades pels programes de traducció automàtica és molt baixa. Un estudi més aprofundit d'aquesta qüestió demostra que molta part dels problemes que s'atribueixen a la qualitat de la traducció automàtica tenen una relació directa amb la baixa qualitat dels documents originals que es tradueixen, i també que l'elevat grau d'expectatives de qualitat dels usuaris sovint és fruit d'una manca de coneixement sobre les possibilitats i els usos idonis del software de traducció automàtica.

b) **Terminologia.** Una altra queixa tradicional dels usuaris de la traducció automàtica ha estat la poca adequació terminològica o de registre de la traducció automàtica. La incorporació de lèxics específics o l'ajut de les eines d'extracció de terminologia sovint no es tradueix en una millora de la qualitat de traducció, i tampoc s'ha acabat de trobar un bon model per tractar el tema de la traducció no desitjada dels noms propis, percebut molt negativament per part dels usuaris finals.

5. Futurs reptes per al català

a) La visió global a l'àrea de llengua catalana, en definitiva, és la següent: tenim perfectament al nostre abast un ampli ventall de programes que ens permeten treballar amb memòries de tra-

ducció. Tenim també a la nostra disposició quatre diferents tecnologies de traducció automàtica per elegir, amb un grau de qualitat de traducció suficient. El primer repte, doncs, és fer que aquestes memòries de traducció i la traducció automàtica es combinin efectivament i permetin una millora de la producció i de la qualitat final de les traduccions. Tard o d'hora, les empreses i organismes públics de l'àrea lingüística catalana responsables de tecnologia relacionada amb la traducció automàtica hauran d'adaptar els seus programes per tal que incorporin memòries de traducció o per aportar als possibles clients solucions d'interacció entre la traducció automàtica i la traducció assistida.

b) Paral·lelament, cal un major desenvolupament d'eines informàtiques que permetin ajudar a controlar la qualitat dels documents originals, com a pas previ al procés de traducció. Els correctors ortogràfics i gramaticals tradicionals (corrector de l'MS Word, WordCorrect, etc.) s'han mostrat insuficients en aquest sentit. Calen eines d'anàlisi similars al que en el cas de l'anglès s'està fent des de fa bastant de temps amb el llenguatge controlat, i que en el cas d'altres llengües com el català està en un estadi molt incipient.

c) Cal, finalment, donar una empenta a la utilització de les eines d'anàlisi i extracció terminològica, i cal fer que els programes existents de traducció automàtica incorporin plenament la possibilitat d'incorporar bases de dades terminològiques de programes externs.

En definitiva, el que es planteja és una major necessitat d'investigar en camps nous, amb vista a millorar la qualitat final de les traduccions, tant si es fan exclusivament utilitzant memòries de traducció com si s'utilitza el suport de la traducció automàtica. Més que no pas insistir en el desenvolupament de més iniciatives relacionades amb la traducció automàtica entre català i castellà, o en la millora de la qualitat de traducció proporcionada pels programes existents, sembla que el camp d'investigació que queda més descobert, avui en dia, és el de la millora del procés previ a la traducció, com a clau per millorar de forma ostensible el resultat final, és a dir, els textos traduïts.