

# LA TRADUCCIÓ AUTOMÀTICA APLICADA A LA PREMSA ESCRITA. EL CAS D'EL PERIÓDICO EN CATALÀ

RICARD FITÉ LABAILA

És veritat que la màquina no podrà substituir mai del tot l'home, però també ho és que sense la màquina l'home no podrà arribar mai tan lluny. Aquesta podria ser una resposta genèrica a les preguntes que ben segur es formulen els professionals de la traducció amb la recent proliferació de programes informàtics més o menys capaços de traslladar a una altra llengua un text qualsevol. Preguntes de la mena ¿una màquina pot traduir un text de la mateixa manera que ho faria un ésser humà?, ¿el pot arribar a substituir a mesura que sigui perfeccionada?

L'home ha estat capaç de crear una màquina que el supera pel que fa a rapidesa i productivitat, que en aquest aspecte *l'limita*, i la incògnita consisteix a saber fins a quin punt l'invent liquidarà l'inventor. Però no hem d'estar preocupats pels nostres llocs de treball, perquè la màquina també té les seves limitacions, i *ilimitar-la* implica d'alguna manera *humanitzar-la*, és a dir fer-la anar més a poc a poc, i per aquest camí no podrà substituir mai l'ésser humà. Home i màquina es necessiten, depenen l'un de l'altre si s'ha d'anar frenèticament de pressa i ser lingüísticament fiables al cent per cent. La màquina pot córrer més perquè no pensa: només compara, però no decideix mai entre dues o més opcions elegibles més enllà de l'irracional i insensible càlcul, perquè per poder decidir s'ha de pensar, i per poder pensar s'ha de ser lliure, humà. Aquesta seria la conclusió. Vegem com hi hem arribat.

La versió catalana d'*El Periódico de Catalunya*, que es va començar a publicar el 28 d'octubre de 1997, ha estat i és diàriament possible d'ençà d'aquella data gràcies a un programa de traducció automàtica i a un equip de més de trenta periodistes-editors-traductors que treballen cada una de les pàgines a partir de la inicial proposta artificial fins a donar-los la forma definitiva que el lector troba l'endemà al quiosc. I aquesta unió de programa i equip és tan inseparable i indispensable per a la mera existència del producte i per a la bona marxa del negoci —i mai més ben dit, perquè *El Periódico* en català és efectivament un negoci, econòmic i social—, que si un

membre d'aquesta parella falla no hi ha producte possible: sense programa d'ajuda l'equip humà no pot en cap cas fer arribar a la rotativa l'edició acabada a una hora factible, i sense equip humà no hi ha director de diari responsable que gosi servir als seus lectors una edició del tot il·legible com la que resultaria de la versió en brut generada pel sistema.

Per això s'ha de defensar per força l'existència dels dos membres de la parella, sense perdre ni un segon a analitzar fins a quin punt perfeccionant la qualitat del programa o augmentant el nombre d'integrants de l'equip es podrien superar els inconvenients que presenten màquina (de qualitat i precisió) i factor humà (de costos econòmics) i, en conseqüència, apostar només per un d'ells. No. En aquest cas, no hi ha producte si no hi ha programa, i tampoc hi ha producte si no hi ha equip. De totes maneres, vegem els avantatges i els inconvenients de la màquina, que és un dels factors que condicionen més la traducció.

Desenvolupat pel departament de R+D d'*El Periódico de Catalunya*, el programa de traducció automàtica castellà-català que serveix de punt de partida per elaborar la versió catalana del diari tradueix en l'actualitat (juny del 1998) una pàgina de diari —unes 150 línies de 70 espais, és a dir uns cinc folis de text— en un temps increïblement breu de tres segons, absolutament espectacular i que estalvia tota mena de consideracions crítiques sobre la velocitat del procés. Aquest és precisament el seu gran avantatge, perquè permet disposar dels textos de la pàgina *presumptament* catalana en aquest registre rècord, i la feina d'edició sempre requereix menys temps que la de traducció manual.

Però aquesta velocitat traductora de vertigen, gairebé diríem que sideral, només és possible perquè el programa —malgrat la seva capacitat infinita per admetre noves entrades lèxiques sense cap mena de restricció quantitativa— està concebut de tal manera que no té la possibilitat d'aplicar regles gramaticals —el que en diríem intel·ligència artificial— que donin una mínima resposta als components morfològic, sintàctic i semàntic de la llengua; és a dir que presenta una restricció d'origen, per la qual només pot comparar —mai relacionar— els termes o seqüències prèviament introduïts.

Així doncs, dit d'una altra manera, la potència i eficàcia d'aquest programa se circumscriu únicament al nivell lèxic (i als altres nivells solament en la mesura que pot processar seqüències com a sintagmes lexicalitzats, és a dir deixant sense solució la variació morfosintàctica de cada una d'elles). Per tant, aquest programa d'ajuda deu la seva espectacular velocitat a la seva absoluta (i també espectacular) incompetència morfosintàctica i semàntica, cosa que, com òbviament es dedueix, converteix en infinita la quantitat d'errors de traducció que ha de resoldre la intervenció humana

posterior. En aquest cas, el repte informàtic consistiria a afegir l'esmentada intel·ligència al programa sense disminuir-ne la velocitat.

Per posar alguns exemples de les limitacions lingüístiques del programa, diguem que ignora tota regla de flexió (substantiva, adjectiva, verbal i pronominal) i de concordança, i que no reconeix les diferents categories gramaticals a què pot pertànyer una mateixa forma lèxica, a més a més, per descomptat, de desconèixer les que encara no li han estat introduïdes. Això comporta la impossibilitat de traduir al català un relatiu d'estructura complexa i variable com «cuyo», i genera una infinitat de propostes errònies, a vegades aparentment possibles i a vegades clarament *inhumanes*: «agredeix el seu ex» en lloc de «agredeix la seva ex» («agrede a su ex»), «una nas molt grossa» en lloc de «un nas molt gros» («una nariz muy gruesa»), «la justícia nord-americà» en lloc de «la justícia nord-americana» («la justicia estadounidense»), «va venir de la Rioja» en lloc de «vi de la Rioja» («vino de la Rioja»), «l'oposició denúncia corrupció» en lloc de «denuncia» («la oposición denuncia corrupción»), «una atempta vigilància» en lloc de «una atenta vigilància» («una atenta vigilancia»), «va arribar tarda» en lloc de «tard» («llegó tarde»), «vol que carrer» en lloc de «vol que calli» («quiere que calle»), o bé «un nen fos de la terrassa» en lloc de «un nen fora de la terrassa» («un niño fuera de la terraza»), entre moltes altres d'una infinita llista de disbarats.

Per evitar casos com aquest últim, i com que el programa permet ser alimentat constantment, el mateix equip humà encarregat de l'edició del diari va introduint paral·lelament en la base de dades solucions seqüencials específiques per reduir en el futur aquesta mena d'errors. Així, reprenent l'últim exemple, al sintagma castellà «fuera de» se li pot assignar la traducció equivalent «fora de», amb què el risc de la confusió entre adverbi i verb disminueix, com a mínim en part. En tot cas, la feina de millora del programa s'ha de fer sempre prenent en consideració la tipologia dels textos que es tradueixen (en el nostre cas, el registre periodístic en sentit ampli), i que determinaran la solució última per la qual s'opta. Solució que, si en la pràctica diària es revela ineficaç, es pot substituir per una altra que s'intueixi com a més productiva.

De totes maneres, els inconvenients de la traducció automàtica en el cas que comentem no s'acaben en la falta de regles. Fins i tot des del punt de vista estrictament lèxic, que constitueix el seu punt fort, sovint no surt a compte introduir en el programa segons quines solucions, perquè es corre un risc d'error encara més gran, per poc que l'editor del text badi, i s'arriba a la conclusió que és preferible que en certs casos el programa *no ajudi*. Així s'esdevé amb totes aquelles paraules en espanyol la traducció catalana de les quals comporta una pluralitat de significats respecte al significat de referència. Seria el cas de «le mandó un telegrama», que es converteix

en «li va manar un telegrama» en lloc de «li va enviar un telegrama», o de «la llamada operación retorno», traduïda per «la trucada operació tornada» en lloc de «l'anomenada operació tornada». No sembla, doncs, probable que arribi a existir un programa d'ajuda a la traducció capaç de discernir si, davant un sintagma aparentment simple en castellà com «el fallo del tribunal», en català pertoca escriure «la decisió del tribunal» o bé «l'error del tribunal», tenint en compte que les dues seqüències serien solucions possibles, i la segona, per inhabitual, encara més notícia.

Davant de casos com aquest, molt més freqüents que no sembla en qualsevol registre lingüístic, i que tenen a veure amb la semàntica i el context, ni tan sols un programa dotat de regles gramaticals, és a dir d'una certa intel·ligència artificial, ens podria resoldre el problema satisfactòriament, atès que cap sistema de traducció automàtica posseeix, ni posseirà mai, l'anomenat *coneixement del món*, és a dir la informació sobre les coses i sobre la vida en general de què només podem disposar els éssers humans més enllà de l'estrictament lingüística. D'aquí que la fiabilitat de qualsevol programa d'ajuda a la traducció no s'hagi de mesurar mai pel nombre de paraules que és capaç de traduir, sinó pel nombre de frases que és capaç de traduir bé.

No és aquest el lloc adequat per donar compte de tota la casuística meixidora de comentari que genera diàriament un programa d'aquestes característiques. Tanmateix, en general es pot afirmar, a més de tot el que hem ressenyat, que els aspectes que requereixen més atenció per part de l'equip editor són els relatius a les especificitats sintàctiques del català: els pronoms relatius i febles —»lo» pot ser en català «el», «lo» i «ho»—, les preposicions —»para» a vegades ha de ser «per» i a vegades «per a»—, l'ordre de la frase, els modismes, les frases fetes i els sentits figurats en el terreny semàntic, la genuïtat de la clàusula considerada globalment davant el risc constant de calc, i les convencions típiques de qualsevol publicació, entre les quals ocupa un lloc destacat saber quins són els noms propis que s'han de traduir i quins no —que el programa adapta, en principi, de manera sistemàtica i indiscriminada.

Per acabar, voldria esmentar tres aspectes més de l'apassionant feina que hi ha darrere del programa d'ajuda. En primer lloc, la tasca inacabable de combatre la implacable i imparabile regularitat de la màquina, atès que, a diferència d'un traductor humà, davant d'un mateix estímul el programa d'ajuda ofereix sempre la mateixa solució, del tot aliè a la influència del context. En segon lloc, la tasca no menys interminable de resistir a la falsa frustració que produeix traduir sempre en una mateixa direcció, en el nostre cas del castellà al català, cosa que sovint ens porta a pensar, erròniament per descomptat, que la llengua d'arribada sempre és més pobra, limitada i desgraciada que la d'origen. I en tercer lloc —i aquí arriba el premi—, la satis-

facció que produeix aconseguir un titular del mateix cos i amb un impacte equivalent al de l'original sense cap altra ajuda que la pròpia inventiva i una certa intuïció: «Sin filtro o emboquillado», titular d'un reportatge sobre el tabac, va acabar sent «Amb filtre o sense», molt millor que el maquinal «Sense filtre o amb broquet»; i una intencionada i treballada portada amb el títol «El medicamento levanta ampollas» va tenir un digne acompanyant als quioscos un diumenge al matí gràcies a l'expressió «fa bullir la sang», que manté el doble sentit original i per això supera amb escreix el corresponent genèric «aixeca polseguera».

En el cas d'*El Periódico*, doncs, la resposta a la pregunta d'aquesta taula rodona, què fan els mitjans de comunicació amb la llengua?, és ben senzilla: fer arribar el seu producte al quiosc en castellà i en català, perquè el lector pugui triar. Però tampoc es pot ignorar que, com a mínim en la primera època d'aquesta doble versió, la informació és concebuda en una llengua i traduïda a l'altra, amb el difícil objectiu —gairebé sempre complert, afortunadament— que el lector tingui tothora la sensació que està llegint en versió original. Si fa no fa, com al cinema.