

■ Els límits de la traducció automàtica

Ricard Fité Labaila

Com tot en la vida, un traductor automàtic té avantatges i inconvenients. Dit d'una altra manera, és una gran ajuda i alhora un gran perill. Partirem d'aquesta premissa. És una gran ajuda perquè permet traduir més de pressa, i per tant abordar empreses inimaginables abans de la seva aparició. Per exemple, al nostre país ha fet possible l'existència de diaris d'informació general en doble edició, com és el cas del *Segre* i *El Periódico*, i ha facilitat la versió bilingüe de diverses publicacions, com ara el *Diari oficial de la Generalitat de Catalunya*. I és un gran perill perquè, si se'n fa un mal ús, pot generar textos degradants per a la llengua que en teoria pretén servir.

Hi ha una sèrie de factors que condicionen l'eficàcia d'un traductor automàtic. Aquests factors poden ser de dues menes: poden tenir a veure amb les característiques de les llengües amb què es treballa, o bé amb la tipologia dels textos a traduir. En el primer cas, el de les llengües, és obvi que no és el mateix traduir entre dues llengües similars, pertanyents a una mateixa branca (castellà i català, o italià i català, posem per cas), que fer-ho entre dues llengües allunyades (anglès i català, o basc i català). Més enllà de com estigui configurat internament el traductor, els resultats obtinguts seran més bons i fiables com més s'assemblin estructuralment les llengües posades en relació. En el segon cas, el dels textos, també és evident que un traductor automàtic donarà més bons resultats amb textos que accepten més bé la literalitat, com ara una publicació de temàtica econòmica, financera o borsària, que no pas amb textos que l'accepten més malament, i aquí podríem parlar des de



La torre de Babel,
Peter Brueghel el Vell

la creació literària, amb la poesia com a màxim exponent, fins a tots aquells gèneres que requereixen una interpretació, ja sigui perquè s'allunyen del registre estàndard, com ara expressions d'argot, llenguatge col·loquial, modismes, frases fetes, refranys, etcètera, ja sigui perquè presenten unes característiques específiques que fan inviable aplicar-hi la traducció literal, com passa amb la majoria d'acudits basats en jocs de paraules.

Traducció automàtica equival, doncs, a traducció literal. Per tant, els límits d'un traductor automàtic són els mateixos que els d'una traducció literal. Si la frase admet una traducció literal, el traductor automàtic fa la feina ben feta. Si la frase no l'admet, el

traductor automàtic falla estrepitosament. I com que mai es pot assegurar que un text, sigui de la mena que sigui, admetrà sempre una traducció literal, la conclusió és que un traductor automàtic no pot fer la feina tot sol, que no se'l pot deixar sol. Que pot ajudar-nos molt, sí. Però que ho pot fer tot, no. Vet aquí el gran límit. I el respecte que ens mereixi aquest límit determinarà el bon o mal ús que se'n faci, és a dir que jugui a favor o en contra de la llengua que en resulti.

Des del punt de vista de l'usuari, en el nostre cas de la persona que ha de fer una traducció del castellà al català o viceversa, el traductor automàtic en part li facilita la feina i en part l'hi complica. L'hi facilita perquè guanya temps. Està estadísticament comprovat que traduint amb una eina d'aquesta mena la feina s'enllesteix més de pressa que si es tradueix de la manera tradicional, manualment, havent de teclejar tot el text de cap i de nou. Podríem dir que un traductor automàtic resol ben satisfactòriament la part no creativa de la traducció, la part més mecànica, la més literal. Però també l'hi complica perquè l'eina tradueix cegament, automàticament, com no podia ser d'una altra manera. I en aquest apartat, els estralls derivats dels automatismes poden ser molt variats, en funció de com estigui configurat el traductor.

Segui com sigui, el rendiment que es pot treure d'un traductor automàtic té molt a veure amb el coneixement del seu funcionament per part de l'usuari. És fonamental, doncs, que l'usuari pugui tenir un coneixement previ d'aquest funcionament. Si, a més, té la sort de poder accedir a la seva configuració per modificar-ne el comportament, els resultats encara seran millors. Per tant, com que per a la majoria dels usuaris de traductors automàtics, sobretot si són ocasionals, d'entrada és difícil saber com estan configurats internament, només l'experiència continuada amb la mateixa eina permetrà obtenir-ne un més bon rendiment.

Entre els avantatges que presenta un traductor automàtic, a més del temps que es guanya, hi ha la correcció ortogràfica. El traductor automàtic no fa faltes d'ortografia. Comet altres errors, però d'ortografia no. I no solament no en comet, sinó que impedeix que aparegui el dubte. Per tant, l'usuari veurà de seguida que no s'haurà de preocupar mai de si una determinada paraula s'escriu amb *ela* geminada o no (*protocol·lari* serà sempre la traducció de *protocolario*), si porta una *essa* o dues (*decomisos, postissos*), si va amb *q*

o amb *c* (*iniqua, conspícua*), si va amb una ena o sense (*trasvase = transvasament, transtorno = trastorn*), si porta accent o dièresi (*veí, veïns*), etcètera.

Un altre dels avantatges del traductor automàtic és la regularitat estilística dels textos que serveix. En aquest sentit, fomenta l'homogeneïtat de les variants morfològiques adoptades, i aquesta característica el fa especialment útil per a textos de registre estàndard. Per exemple, un text traduït automàticament no alternarà formes diferents de plural, d'acord amb la seva configuració: si té l'ordre de traduir *deseos* per *desigs*, no traduirà ara *desigs*, ara *desitjos*; els *boscós* no es barrejaran amb els *boscs*, i formes verbals com *van dir* no es combinaran amb *digueren*. Es pot afirmar, doncs, que en aquests casos un traductor automàtic garanteix un índex superior de coherència textual. Ho comprovaríem comparant una traducció manual feta entre dues o més persones, sense haver-los donat criteris previs, amb una traducció automàtica del mateix text revisada per aquestes mateixes persones.

Però les persones que revisen una traducció automàtica han de ser molt conscients de les limitacions de l'eina utilitzada, és a dir dels inconvenients que presenta un text obtingut per aquest procediment. Per exemple, han de preveure que el traductor no distingeixi si una paraula està escrita amb majúscula inicial o no. Que el traductor la distingeixi o no, dependrà de com estigui configurat. Però si no la distingeix, pot passar que un senyor que en el text traduït es diu David Fonts, en el text original es digui David Fuentes, i que aquest greu error passi per alt en la revisió i acabi sortint publicat. Aquest és un dels perills de la traducció automàtica: que no distingeixi entre noms comuns i noms propis.

Prevenir un perill com aquest no és una feina fàcil, i a més l'antídote perfecte no existeix. Com que en el món de l'enginyeria informàtica tot és possible, és a dir que tot es pot fer, podem donar al traductor automàtic l'ordre de no traduir les paraules que es trobi escrites amb majúscula inicial, i així resoldrem per sempre que el senyor David Fuentes no aparegui mai més amb el cognom adulterat. Però aquesta prevenció té un dany col·lateral, i és que la paraula *fuentes* no la traduirà mai més quan se la trobi amb majúscula inicial, tant si és un cognom com si no ho és, de manera que també deixarà de traduir-la quan sigui un nom comú que aparegui al principi d'un paràgraf o bé després de punt: «Fuentes de la

directiva del Barça han desmentit la notícia.»

Si l'usuari pot intervenir en la configuració del traductor, i la pot modificar, haurà de triar entre renunciar a la traducció de *fuentes* amb majúscula en qualsevol circumstància, aplicant-hi l'ordre esmentada, o bé córrer el risc de divulgar cognoms falsos. Si no hi pot intervenir, s'haurà d'aprendre quin és l'error possible del traductor i estar-hi atent. En qualsevol cas, no es podrà estalviar la revisió del text.

Aquest exemple de majúscules i minúscules resumeix la problemàtica que presenten els traductors automàtics. Una problemàtica derivada del fet que no tenen accés al coneixement del món que té el cervell humà, i per tant no poden interpretar les paraules i les frases com ho fem les persones. Si és cert que la capacitat del llenguatge humà per generar frases és infinita, i sembla que ho és, perquè la lingüística teòrica així ho afirma, i per tant també és infinita la variació lingüística amb tota la seva complexitat semàntica, caldria donar infinites ordres a un traductor automàtic perquè pogués donar compte de totes les frases en tots els contextos possibles, però això és impossible. És impossible, per exemple, donar una llista finita de noms i cognoms a un traductor automàtic: qualsevol nom comú és susceptible de convertir-se en un nom propi. És impossible que un traductor automàtic identifiqui tots els títols de llibres, pel·lícules, cançons, programes de televisió, etcètera, no traduïbles, i els deixi en la llengua original. És impossible també que distingeixi si en el text a traduir hi ha paraules o frases de llengües diferents, amb el risc consegüent de cometre errors ridículs: *come on!* = *menja on!* Fins i tot és impossible reduir les paraules (és a dir unes formes, uns significants) a un sol significat, i donar-los una correspondència única en una altra llengua. Això de vegades és possible: en català i castellà es pot entendre el mateix per les correspondències *casa* = *casa* i *autorización* = *autorització*. Però hi ha una gran quantitat de paraules polisèmiques en què aquesta identificació no es pot fer (un *hueso* pot ser un *os* o un *pinyol*; un *sueño* pot ser un *son*, una *son* o un *somni*; *alcanzar* pot voler dir *aconseguir* o *arribar*, etcètera). Com que tot això és impossible, el traductor automàtic no pot funcionar sol, necessita la revisió humana.

Totes aquestes limitacions ho són en termes absoluts. Perquè s'ha de dir que, en termes relatius, el traductor automàtic és una eina de gran ajuda. Només insistim que cal ser conscients de la

impossibilitat que pugui funcionar autònomament, sense intervenció humana *a posteriori*. Continuant amb l'exemple d'abans del problema de les majúscules i les minúscules, un traductor pot tenir una configuració que li permeti traduir una paraula d'una manera quan estigui escrita amb majúscula, i d'una altra manera quan estigui escrita amb minúscula. Per exemple, *China* = *Xina*, però *china* = *xinesa*; o bé *Irán* = *Iran*, però *irán* = *aniran*. Això sí, sense oblidar que aleshores quan es trobi un *Irán* amb majúscula que no sigui topònim, sinó la forma de futur del verb *ir* (per exemple, a principi de paràgraf o després de punt), el traduirà com si fos el topònim: *Iran* (és a dir, a la gironina, si se'ns permet la broma).

En el cas concret de la doble edició d'*El Periódico de Catalunya*, hem pogut constatar una evolució positiva dels traductors automàtics en qüestió de pocs anys. El diari en català va sortir al carrer el 1997 amb l'ajuda d'un traductor automàtic molt simple que, resumint-ho molt, es limitava a fer correspondre una paraula amb una altra, sense classificar gramaticalment, i una seqüència de diverses paraules amb una altra, també sense informació gramatical. La versió actual, que data de l'any 2005, supera l'anterior amb la incorporació d'informació gramatical (substantiu, adjectiu, verb, etc.) a qualsevol paraula o seqüència, informació imprescindible per poder dotar el traductor de regles que han permès millorar-ne el rendiment.

Per tant, la millora tècnica d'un traductor automàtic pot paliar en part els inconvenients lingüístics que presenten els textos resultants. Com més complex i perfeccionat sigui el traductor, més capacitat tindrà d'admetre regles, ja sigui sobre qüestions convencionals, com ara la distribució esmentada de majúscules i minúscules, ja sigui sobre aspectes morfosintàctics, com ara problemes de gènere i concordança, distribucions prepositives, etcètera. Per exemple, un traductor automàtic dotat només d'equivalències entre paraules (*triste* = *trist*) cometrà uns errors que en canvi podrà evitar un traductor convenientment dotat de regles de concordança, unes regles que li permetran convertir correctament un adjectiu invariable en castellà (*mi, tu su, triste, verde, marroquí*) en un adjectiu variable en català (*el meu, la meva; el teu, la teva; el seu, la seva, trist, trista; verd, verda; marroquí, marroquina*) segons que sigui masculí o femení el substantiu que acompanyi: *mi padre* = *el meu pare, mi madre* = *la meva mare*, etcètera.

Anàlogament, regles d'aquesta mena permetran traduir seqüències com ara *la calle per el carrer* (i evitar l'agramatical *la carrer*), *datos facilitados per dades facilitades* (en lloc de l'incoherent *dades facilitats*) i *buenas costumbres per bons costums* (en lloc del castellanisme *bones costums*). Pel que fa a una qüestió sempre polèmica entre lingüistes com l'ús de les preposicions *per* i *per a*, a un traductor automàtic també se li pot dir, mitjançant una regla, que les equivalències *por = per* i *para = per a* es neutralitzin en *per* si a continuació hi ha un infinitiu: *por el niño = pel nen*, *para el niño = per al nen*, *para dormir = per dormir*. I així tindrem un traductor automàtic que en aquest punt controvertit de la normativa del català aplicarà sense error la teoria de Coromines-Solà, perquè la norma és regular (a diferència de la de l'Institut d'Estudis Catalans, que fa de més mal interpretar, tant per als humans com per a les màquines).

Per acabar aquest apartat de regles, també cal advertir que de vegades generen problemes nous no previstos, que criden molt l'atenció de l'usuari, en la mesura que s'apliquen a paraules en contacte tant si pertanyen al mateix sintagma gramatical com si no. No és inusual, per tant, comprovar com una regla postulada per resoldre un problema concret, en provoqui un de pitjor. Es tracta, en cada cas, d'avaluar quina és la instrucció que més convé que tingui un traductor amb caràcter general.

Però tan important com el rendiment d'un traductor automàtic segons el nombre d'equivalències i regles que el conformen, són les sensacions que pot tenir el seu usuari en relació amb la qualitat de la llengua que li subministra, és a dir, de la llengua traduïda. Si l'usuari és ocasional i no es dedica professionalment a la traducció, li pot semblar tot plegat un acte de màgia que li farà exclamar, com al castís de la sarsuela, que «las ciencias adelantan que es una barbaridad». Però si es tracta d'un usuari que es dedica professionalment a la traducció, l'enlluernament de la immediatesa i de la feina mecànica estalviada no li ha d'amagar en cap cas els perills subtils que entranya una traducció automàtica en aspectes que van més enllà de l'encert objectiu de l'eina. Aquests perills es resumeixen en una frase utilitzada sovint per la gent que revisa traduccions automàtiques de castellà a català: «Tot sembla català, aparentment ho és, però això no és català. Són paraules catalanes, però en català no ho hauríem escrit així.» Aquest és un dels factors més importants a tenir en compte a l'hora d'avaluar els límits de la

traducció automàtica. Perquè es tracta d'una qüestió que pot passar desapercebuda a ulls poc crítics, molt acostumats al contacte quotidià entre castellà i català, i amb criteris laxos o poc escrupolosos respecte a què pertany només a una llengua, què pertany només a l'altra, què pertany a totes dues alhora, o que hi pot arribar a pertànyer de manera acceptable en virtut de l'evolució respectiva i de l'obligada convivència.

No estem parlant de construccions del català objectivament incorrectes o que no formen part de la llengua normativa, com ho serien l'absència indeguda de pronoms febles o l'ús de castellanismes. No. Ens referim a aspectes més subtils, que afecten l'ordre de la frase, el dring, la sonoritat, la tria d'un lèxic determinat, la tradició («ho hem dit així tota la vida», «no ho hem dit mai així»), en definitiva tots aquells intangibles que fan reconèixer un text com a català, sense interferències de cap mena.

Un exemple d'això que diem es veu tot sovint en l'ordre de les preguntes. Algunes llengües inverteixen l'anomenat ordre lògic de subjecte i verb de la frase declarativa quan la fan interrogativa. En el cas del francès, aquesta inversió és fixa i obligatòria, es tracta d'una norma clara, rígida, inflexible. En el cas del castellà, aquesta inversió es pot fer o no, però més aviat s'acostuma a fer: «¿Cree usted que ya hemos superado la crisis?» En català, aquesta inversió no se sent com a natural, i per tant espontàniament no s'acostuma a fer: «¿Vostè creu que ja hem superat la crisi?» Un traductor automàtic no donarà compte d'aquesta inversió: «¿Creu vostè que ja hem superat la crisi?» A això ens referim: aparentment la frase és catalana, potser ningú gosaria dir que és incorrecta, però la veritat és que segurament mai hauríem formulat la pregunta en aquest ordre si no estiguéssim traduïnt.

Revisar traduccions automàtiques i reflexionar-hi és un bon exercici per constatar l'existència de dubtes com aquest que hem exposat. És cert que aquests límits es poden experimentar revisant o llegint una traducció feta manualment, perquè el contacte de llengües hi és igualment present. Però és evident que l'ús del traductor automàtic fa augmentar el perill de generar frases amb paraules de la llengua d'arribada, considerades una per una, però que en realitat no són sinó traduccions literals sense ànima si un s'hi fixa una mica.

Es tracta, doncs, d'evitar que a l'usuari d'un traductor automàtic li passi el mateix que a aquell famós estudiant de música de l'acudit, que li deia al seu professor: «Mestre, jo poso les notes i vostè ja posarà la música.» Tots estarem d'acord que l'objectiu final d'una bona traducció és que es pugui llegir sense noses, que no es noti que és una traducció, que sembli un text escrit directament en la llengua traduïda. En definitiva, que cap lector lamenti no poder-ne llegir la versió original. A això aspirem, al capdavant, tots els professionals de la traducció, amb l'ajuda de traductors automàtics o sense.

RICARD FITÉ LABAILA és coordinador lingüístic de l'edició en català d'*El Periódico de Catalunya*