



## Guia sobre la traducció automàtica: usos, preedició, postedició i problemes habituals

La traducció automàtica és el procés de traducció de textos mitjançant un sistema informàtic sense intervenció de l'usuari. Com a procés automàtic no supervisat, el resultat no pot tenir les mateixes garanties que un procés en què hi ha hagut una supervisió humana. Encara que l'evolució de la tecnologia fa que cada vegada els resultats de la traducció automàtica siguin més satisfactoris, no es pot considerar el text resultat d'una traducció automàtica com un text definitiu amb garanties i qualitat.

Això no impedeix, però, que per a determinades finalitats sigui acceptable fer servir una traducció automàtica no supervisada.

### Usos de la traducció automàtica

Es poden distingir dos usos bàsics per a la traducció automàtica: comprensió i publicació.

- **Comprensió** (assimilació). La finalitat d'aquest tipus de traducció és senzillament tenir una idea aproximada del contingut del text. És aplicable en casos en què el document traduït només té la finalitat de fer conèixer de manera aproximada el contingut del document original. Atès que és una traducció sense cap tipus de supervisió lingüística, no se'n pot difondre de manera directa o indirecta el contingut sense que hi hagi una revisió i una comparació del text traduït amb l'original. Si es vol fer una traducció amb aquesta finalitat purament informativa del contingut aproximat del document original, és convenient acompanyar la traducció d'un advertiment que indiqui que es tracta d'una traducció automàtica sense revisar (vegeu, per exemple, les [plantilles del Word](#) publicades pels Serveis Lingüístics).
- **Publicació** (disseminació). La finalitat d'aquest tipus de traducció, per contra, és la de generar un document destinat a la difusió i que, per tant, ha de tenir les garanties de qualitat que assegurin que el contingut de la traducció correspon al del text original. Avui dia, aquesta garantia només es pot obtenir amb la intervenció humana. El text resultant de la traducció automàtica s'ha de revisar (tenint present el text original) i corregir en cas que es detectin errades induïdes pel procés de traducció.

### Problemes de la traducció automàtica

Malgrat l'avenç de la tecnologia relacionada amb la traducció automàtica, la naturalesa pròpia del llenguatge fa difícil veure proper el moment en què els ordinadors tinguin prou fiabilitat per fer traduccions acurades al 100 %. Els problemes de la traducció automàtica es poden classificar en quatre grups:

- **Ambigüitat del llenguatge.** Els fenòmens com ara la polisèmia (quan una paraula té més d'un significat) o l'homografia (quan dues paraules s'escriuen igual) poden crear textos que es poden interpretar de maneres molt diferents.
- **Diversitat formal.** Les maneres d'expressar un mateix contingut poden ser molt diverses i no sempre totes les solucions són adequades als diferents tipus de text.
- **Diversitat estructural entre les llengües.** L'estructura corresponent al mateix contingut en cada llengua pot arribar a ser molt diferent. En aquest cas la dificultat s'incrementa com més allunyades són les llengües entre les quals es vol traduir.
- **Context de producció del text.** Sovint la interpretació del text depèn de factors externs (coneixements culturals, socials i científics, per exemple) que poden donar les claus per interpretar correctament el missatge.



## Optimització de la traducció automàtica

Amb la finalitat de minimitzar els errors de la traducció automàtica és convenient fer un tractament del text abans de la traducció (*preedició*) i un altre després de la traducció (*postedició*), que en el cas de les traduccions automàtiques per a la comprensió pot ser més superficial o, fins i tot, inexistent.

### Preedició

La preedició consisteix a preparar el text que s'ha de traduir automàticament per evitar o minimitzar els possibles errors produïts en la traducció automàtica.

La tasca bàsica que cal aplicar en qualsevol preedició és la revisió ortogràfica del text. Aquesta tasca es pot fer mitjançant l'ajuda d'eines de revisió ortogràfica o gramatical (des de les eines de verificació pròpies dels programes d'edició fins a altres aplicacions de suport extern com ara el [LanguageTool](#)).

En el cas de les traduccions automàtiques amb la finalitat de publicar el document final, pot ser interessant fer una actuació més profunda i, mitjançant macros o cerques automàtiques, localitzar i marcar els fragments que potencialment poden provocar errors en la traducció automàtica.<sup>1</sup> De tota manera, aquesta actuació sempre ha de tenir com a base la tipificació dels errors de traducció automàtica detectats de manera sistemàtica, i això depèn sempre, d'una banda, del parell de llengües implicat en la traducció i, de l'altra, del sistema de traducció automàtica usat.<sup>2</sup>

### Postedició

Un cop s'ha traduït el document és indispensable que hi hagi una nova intervenció: la postedició. En la postedició cal fer una revisió del document abans de donar-lo per definitiu. El grau d'intervenció, com en el cas de la preedició, depèn de la finalitat de la traducció.

En el cas de la traducció usada com a eina de comprensió del text, la postedició mínima consisteix en la revisió ràpida del document per assegurar-se que tot el document ha quedat traduït i per detectar que no s'ha produït cap desajustament de la composició del text que en pugui dificultar la lectura. D'altra banda, si es vol aprofitar la traducció resultant per adjuntar-la a l'original amb vista a facilitar la comprensió del text a altres usuaris, és convenient, com ja s'ha dit, deixar constància que no s'ha fet cap revisió de la traducció automàtica.

Quan la finalitat de la traducció és la publicació de la traducció, cal que la postedició sigui més profunda. A banda de la revisió formal del text, cal assegurar-se que el contingut del text traduït correspon al del document original i, per tant, encara que es facin servir eines electròniques per detectar possibles errors sistemàtics, sempre és necessària una lectura atenta del document traduït acarant-lo amb l'original.<sup>3</sup>

---

1. Per exemple, en el cas que el text que s'ha de traduir contingui molts noms propis, mitjançant una substitució automàtica o una macro, es pot fer que tots els mots començats amb majúscula inicial no precedits de punt quedin marcats, perquè en la revisió es comprovi que cap nom no ha quedat traduït per error.

2. Per exemple, si en traduccions de l'espanyol al català es detecta que el traductor automàtic no resol correctament la traducció del relatiu *cuyo*, mitjançant una substitució automàtica o una macro es poden marcar sistemàticament tots els fragments de l'original amb aquest pronom per assegurar una revisió exhaustiva del text.

3. Pot semblar paradoxal, però com més es perfeccionen els processos de traducció automàtica, més atenta ha de ser la tasca de postedició; perquè la traducció generada automàticament pot ser un text perfectament intel·ligible però amb errors subtils més complicats de veure.

## Errors habituals

Tot i que les tipologies d'errors poden variar força depenent del sistema de traducció automàtica emprat i del parell de llengües de la traducció, es pot fer una classificació d'errors habituals.

- **Errors induïts per incapacitat tècnica de processar correctament l'original.** En aquest apartat es distingeixen diversos casos que demanen actuacions diferents i que, de vegades, es poden solucionar en la fase de preedició:
  - *Errors ortogràfics de l'original.* En general, els errors ortogràfics o gramaticals del text original tenen com a conseqüència una traducció errònia. Tot i que alguns sistemes de traducció ja són capaços d'esquivar determinats errors ortogràfics de l'original, la millor manera de minimitzar aquests errors és la revisió prèvia del text feta en la preedició.
  - *Mala segmentació de l'original.* En alguns casos, depenent tant del sistema de traducció com del format del document que s'ha de traduir, hi pot haver errors provocats per una mala segmentació del text del document original.

Per exemple, en la traducció de documents en format PDF, hi pot haver problemes amb la interpretació dels salts de línia que poden produir segmentacions anòmales del text que quedaran reflectides en la traducció. En altres ocasions —per exemple, marques de format o altres informacions no visibles en les interfícies d'edició de determinats programes— poden provocar segmentacions errònies de frases o fins i tot de mots.

De vegades, el simple canvi de format del text (per exemple, de PDF a format del Word o de format del Word a text net) pot evitar errors d'aquest tipus, tot i que sempre cal tenir en compte les pèrdues gràfiques o de format que poden implicar aquests canvis.
  - *Grau d'especialització del text.* Quan el text conté molta terminologia que el sistema de traducció no reconeix, pot passar que o bé tradueixi malament el terme perquè en fa una traducció no adequada al camp d'especialització o bé que el deixi sense traduir en la llengua original. En aquest cas el suport de glossaris especialitzats o de bases de dades terminològiques pot ser d'ajuda en la resolució dels errors terminològics, com també l'ús d'eines de verificació ortogràfica complementàries que permetin verificar l'ortografia de la terminologia.<sup>4</sup>
- **Errors de traducció de noms propis.** Tot i que cada vegada és un dels aspectes que es controlen més en els sistemes de traducció automàtica, encara pot ser font d'errors la traducció de determinats noms propis. Com en altres casos, una postedició acurada és el remei fonamental per localitzar aquests errors.
- **Errors d'interpretació semàntica de l'original.** Potser és l'error més característic de la traducció automàtica, en què la incapacitat del sistema per interpretar el significat del text original fa que produeixi traduccions totalment esbiaixades.
  - *Errors provocats per l'ambigüitat.* Fenòmens lingüístics com ara la polisèmia i l'homografia són els principals causants de traduccions errònies. Només la lectura atenta en el procés de postedició permet localitzar aquests errors.
  - *Errors provocats per l'ús de fraseologia.* L'ús de locucions, frases fetes o d'altres expressions en sentit figurat també poden provocar una mala interpretació per part del sistema de traducció automàtica i donar com a resultat traduccions errònies o amb solucions inadequades.

---

4. En casos com aquest cal valorar la importància d'alimentar els diccionaris de verificació ortogràfica amb terminologia especialitzada, atès que això pot facilitar la revisió d'errors de traducció d'aquest tipus.

- **Errorls induïts per diferències morfosintàctiques entre les llengües.** Les diferències entre les estructures sintàctiques i els diferents recursos morfològics entre les llengües de traducció també són una de les causes principals d'error.
  - *Errorls de concordança de gènere.* Els canvis de gènere entre els noms equivalents de les dues llengües de traducció pot provocar errors en la selecció adequada d'elements relacionats amb aquests mots com ara adjectius, determinants o pronoms.
  - *Errorls de col·locació.* En el procés de transformació del text d'una llengua a l'altra, el traductor automàtic pot traslladar a la llengua d'arribada estructures corresponents a la de partida. El cas de la col·locació de l'adjectiu en relació amb el substantiu seria un exemple habitual en el cas de la traducció entre el català i l'espanyol.
  - *Errorls de selecció de constituents i estructures sintàctiques.* També hi pot haver errors en la selecció de determinats tipus de determinants, adjectius o pronoms a l'hora de construir el text traduït. Com a exemples típics en la traducció entre català i espanyol hi hauria els diferents usos dels possessius, alguns determinants, els pronoms de relatiu o els pronoms febles adverbials.
  - *Errorls de canvi de règim preposicional.* Les diferències dels règims sintàctics entre les dues llengües també poden introduir errors freqüents en les traduccions. Com a exemple típic hi hauria les diferències en l'ús de la preposició *a* en complements directes en català i espanyol, que de retruc poden afectar altres elements com ara la selecció de pronoms febles.
- **Errorls induïts per diferències estilístiques o per excés de literalitat.** Les diferències estilístiques entre les llengües de traducció fan que sovint es traslladin directament determinades estructures i usos habituals de la llengua de partida a la d'arribada. Com a exemples habituals en les traduccions entre el català i l'espanyol es poden esmentar les diferències en l'ús del futur en el sentit d'obligació, les perífrasis d'obligacions o les formes de tractament personal. A banda dels casos en què es pot provocar una estructura incorrecta en el text en la llengua d'arribada, en aquest apartat també s'inclouen els casos en què el text resultant presenta una traducció tan literal del text original que la redacció en la llengua d'arribada queda poc natural.

## Recomanacions finals

En les diferents tasques relacionades amb el procés de traducció automàtica cal fer servir totes les eines i recursos que poden facilitar la feina.

- En les fases de preedició i postedició cal fer servir les eines de verificació ortogràfica i de correcció gramatical que es tinguin a l'abast, com ara els diccionaris complementaris del Word o el LanguageTool (vegeu un recull de recursos publicats al *Retrucs* amb l'etiqueta [corrector](#)).
- Per a la tasca de traducció hi ha disponibles diferents eines i recursos ben útils i adequats per a les diferents necessitats. En el cas de traduccions habituals de textos amb estructures molt repetitives o molt fixades pot ser molt útil també combinar la traducció automàtica amb sistemes de traducció assistida (vegeu un recull de recursos publicats al *Retrucs* amb l'etiqueta [traducció](#)).
- En la fase de postedició també pot ser útil fer servir el sistema de [marques de revisió del Word](#) per detectar errors sistemàtics per si es volen implementar cerques i substitucions en futures traduccions en les fases tant de preedició com de postedició. És convenient també conèixer l'[eina de substitució del Word](#) per aprofitar-la bé en aquesta tasca.



## Bibliografia i més informació

DOMÈNECH BAGARIA, Ona. «[La traducció entre el castellà i el català](#)». Barcelona: Universitat Oberta de Catalunya, 2012.

FORCADA, Mikel L. «[Making sense of neural machine translation](#)». [Publicació preliminar de l'autor d'un article publicat a *Translation Spaces*, 6:2 (2017)].

GENERALITAT DE CATALUNYA. DEPARTAMENT DE LA VICEPRESIDÈNCIA. SECRETARIA DE POLÍTICA LINGÜÍSTICA. [Criteris de traducció de textos normatius del castellà al català](#). Barcelona: Biblioteca Tècnica de Política Lingüística, 7, novembre de 2010.

GIMENO, Ramon. «[Els sistemes de traducció automàtica: 8 errors habituals](#)». A: Lletres i TICs: Llengua, edició, disseny i tecnologies de la paraula, juliol de 2013.

GINESTÍ ROSELL, Mireia; FORCADA, Mikel L. «[La traducció automàtica en la pràctica: aplicacions, dificultats i estratègies de desenvolupament](#)». A: *Caplletra*, 46, primavera 2009.

PÉREZ, Rosa; RODÀ, Judit. «[La traducció automàtica i l'experiència del BOE en català](#)». A: *Llengua i ús: revista tècnica de política lingüística*, 21, 2001.

SANTANACH I SABATÉS, Laia. [La traducció automàtica amb postedició en una UE multilingüe: el cas del català](#). Barcelona: Universitat Oberta de Catalunya, febrer de 2018.

TORRES-HOSTENCH, OLGA et al. (coord.) [L'ús de traducció automàtica i postedició en les empreses de serveis lingüístics espanyoles: Informe de recerca ProjectA 2015](#). Bellaterra: Universitat Autònoma de Barcelona, 2016.

ZARETSKAYA, Ana. «[Machine Translation Post-Editing at TransPerfect - the 'Human' Side of the Process](#)». A: *Revista Tradumàtica. Tecnologies de la Traducció*, 15, 2017.