

Autor

Albert Vila i Lusilla
Servei de Normalització Lingüística

Una evolució ràpida

En els darrers temps els correctors informàtics de textos escrits en català han evolucionat ràpidament. Avui som ja força lluny d'aquells primers verificadors que només distingien entre grafies reconegudes i no reconegudes, i entre conjunts tipogràfics normals o anòmals, i que no ens suggerien cap possible solució a les errades. Una millora crucial fou la introducció del suggeriment d'alternatives.

Tot i això, les empreses multinacionals no han anat gaire més enllà de la verificació i correcció purament ortogràfiques; potser destacaríem el corrector del WordPerfect. Però una consideració a part mereixen diverses iniciatives sorgides a Catalunya, amb incursions més enllà de la simple ortografia de les paraules aïllades del context.

Encara en l'entorn DOS, el Pla d'informàtica educativa del Departament d'Ensenyament va posar a punt l'ADHOC, i l'empresa DGEC va crear el Printex. L'aparició dels entorns Windows va fer possible un pas important per a l'usuari: el treball directe sobre els editors dels programaris de més ús en el mercat, mentre simultàniament s'ampliava la capacitat de correcció d'errades. Entre aquests nous productes es poden esmentar el WordVer 1.0 i posteriorment el WordCorrect 2.0 (per a PC i per a Apple) i l'Escriu (per a Apple).

Podem preguntar-nos ara quins seran i quins haurien d'ésser els passos immediats. La distinció no és gratuïta, perquè la informàtica, com qualsevol altra tècnica, té la seva pròpia dinàmica interna, que no necessàriament condueix dreturerament a la solució dels problemes més bàsics.

Això es pot entendre fàcilment amb un exemple. Una de les errades més freqüents en català és la construcció del complement directe animat amb la preposició *a*. El caràcter elemental de l'errada no implica, però, que l'estrategia informàtica necessària per detectar-la sigui senzilla, ja que exigiria al programa la realització d'una veritable anàlisi sintàctica de les oracions.

Per acostar-nos més al problema farem en primer lloc una formulació simple —o més aviat simplista— que més endavant

anirà rebent una severa crítica implícita. Comencem doncs dient que el corrector perfecte és aquell que:

- a) ens avisa sempre que ens equivoquem, sigui quina sigui la causa de l'error, i
- b) ens indica en cada cas l'alternativa correcta.

L'expressió *sigui quina sigui la causa de l'error* ens encara amb l'existència de diversos tipus d'error, d'una importància gramatical, una freqüència estadística, unes dificultats estructurals i unes vies de solució específiques de cada un.

Cal que l'usuari tingui ben present que hi ha diverses causes d'incorrecció, i abandoni la idea dissortadament massa estesa que correcció és només correcció ortogràfica. Un escrit pot ésser incorrecte també per causes morfològiques, sintàctiques i semàntiques.

Deixarem aquí de banda les errades tipogràfiques. Aquestes responen a la transgressió de convencions com ara que darre un punt cal deixar un espai en blanc i a continuació començar amb majúscula, i són en general força ben resoltes.

Estat actual i suggeriments

Ara farem una descripció de la situació actual en els àmbits ortogràfic, morfològic, sintàctic i semàntic, i proposarem algunes fites, d'acord amb les necessitats específiques de l'usuari que escriu en català —que no són pas les mateixes que les de qui escriu en un altre idioma.

Per a aquesta part descriptiva hem comptat amb la col·laboració de l'empresa DGEC, autora del WordCorrect 2.0, producte que va tenir el suport de la Generalitat, i d'Inèdit, fabricant de l'Escriu. Atesa la finalitat d'aquest escrit, convé tenir present que el fet que insistim més en les mancances que en els aspectes ben resolts no implica cap demèrit dels productes. Qui els usa, o qui els recomana, ha de saber en quin punt de l'evolució es troben, què els pot demanar i què no.

Ortografia

Davant una seqüència de lletres no reconegudes com una forma d'una paraula de

la llengua, el corrector elabora alternatives afegint, traient o canviant lletres, segons un patró preestablert, i si les combinacions resultants són formes correctes, ens les suggereix com a alternativa.

D'aquesta manera es detecten sense problemes les típiques errades mecanogràfiques, consistents en la inversió d'una seqüència de lletres, o la supressió o duplicació indeguda d'una lletra, i ens n'és suggerida la forma correcta. Així es troben sense dificultat les formes correctes corresponents a *condicionla, estalbir, rtesor, i espeesa*. A partir de *tostrar* un dels correctors ens remet a *torbar* i a *trobar*, però l'altre només a *torbar*.

S'obté també el suggeriment de les alternatives correctes a errades ortogràfiques usuals com ara *intermiten, maravella o almíbar*. Una forma com *colocar* és considerada incorrecta, però només un dels dos correctors ens suggereix la forma bona; el mateix passa en un cas de dues errades acumulades, com ara *barnís*.

Pel que fa a l'ús dels accents diacrítics, els dos correctors segueixen estratègies oposades. L'un accepta sense avís les frases següents:

Em va dir que si.

Val mes que ens afanyem.

M'ha quedat molt be.

Encara no he molt el cafè.

És una dóna fantàstica.

L'altre, en canvi, detecta qualsevol forma susceptible de dur un diacrític i s'hi atura.

Es fa difícil de triar entre els dos mètodes: el que es guanya en fiabilitat es perd, òbviament, en temps.

En canvi, tots dos productes accepten sense advertiment errades corresponents a casos d'alternança entre una paraula i un grup de paraules tals que les alternatives siguin homòfones o quasi homòfones:

Si va entretenir una bona estona.

Aquesta tarda té vist.

No tel donaré pas.

Mi hauràs d'ajudar.

Per una qüestió de coherència, convindria que un corrector que avisa quan hi ha el dubte entre *més* i *mes* avisés també quan el dubte és entre *mi* i *m'hi*.

I, encara, tots dos correctors ignoren les diferències morfològiques i ortogràfiques entre elements afins fonèticament i alhora semànticament:

No puc perdre ni un minut més.

Temen ésser descobert, va estirar-se a terra.

Li semblava que aviat tornaria.

Ja no podria saltar més.

Corren com corria, no va veure res.

Les errades corresponents al mal ús de les variants d'una paraula a causa del context (article determinat, preposició *de* i pronoms febles) són tractades amb una certa vacil·lació. Tots dos correctors coincideixen a bandejar

Anar de hora al llit.

i a suggerir la forma oportuna. En una frase com ara

Me sembla que sí.

tots dos correctors avisen; pel que fa al suggeriment, l'un adopta el criteri rigorista i l'altre el lax. Tots dos correctors assenyalen que

l'universitat

és incorrecte, però només un indica sense vacil·lació l'article correcte; l'altre suggereix per igual *el* o *la*. A la frase

S'el emporta.

un corrector ho dóna per bo; l'altre no, però no dóna el suggeriment correcte. I, finalment, amb

Te'n ha donat tres.

tots dos correctors queden satisfets malgrat la incorrecció.

Es fa imprescindible d'homogeneïtzar el tractament de les combinacions de verbs i pronoms febles: un tractament tan divers és pitjor que una absència total d'anàlisi, atesa la confusió que es pot crear, i més en un punt tan delicat com aquest.

Morfologia

Les errades estrictament morfològiques són detectades pel mateix mètode que les ortogràfiques. Tots dos correctors suggereixen la substitució correcta de *perderé*. Però les formes que, segons la lògica del corrector, es troben més allunyades de la correcta, són detectades com a errades, però no ens n'és suggerida l'alternativa correcta. Així, només un remet *poguent* a *podent*; davant una forma com *diguessi*, l'un remet a *digués* i a *diguessis*, però l'altre només a *diguessis*. I ni l'un ni l'altre no ens donen les formes bones de *coneixo* o de *beveïssis*.

La capacitat d'anàlisi combinatòria, que pot ésser considerada força bona per a l'ortografia estricta, hauria d'ésser ampliada en les terminacions, amb la qual cosa la remissió a les formes correctes a partir de les incorrectes seria més completa.

Sintaxi

La lingüística ens diu que és impossible generar i analitzar frases d'una manera solament seqüencial. Cal, si més no, el recurs al concepte de domini, que dóna pas als típics diagrames arborescents. De moment, però, els correctors fan servir una pseudogramàtica, consistent en un conjunt de regles de restricció de formulació seqüencial, derivades de les verament gramaticals. Així, per exemple, estableixen regles de l'estil següent:

«Un article no pot precedir un demostratiu.»

Amb aquesta tècnica són capaços de detectar elements que es troben en un context inadequat però deixen passar moltes errades que no poden tenir una traducció seqüencial, com la construcció defectuosa del complement directe animat, abans al·ludida.

Un dels dos correctors té previstes més de 300 d'aquestes pseudoregles; en aquest cas la pantalla en subministra una petita explicació, ampliable si es vol a una de més extensa.

Només un dels dos correctors és capaç de detectar les errades contingudes a

unes anàlisi de sang
una faldilla negra

i suggerir la forma correcta. Davant una frase com ara

Estic content de que hagi vingut.

un dels correctors detecta l'errada, però no fa el suggeriment correcte; l'altre ens fa confirmar primer si és *què* o *que*, i en ésser ratificat el *que* suggereix l'eliminació del *de*. Però són molts els casos indetectables per anàlisi seqüencial, i que per consegüent són acceptats sense reserves:

Invitarem a les teves amigues.

Els hi van posar el llibre davant.

Vocabulari i semàntica

Arribem a l'àmbit més difícil, tot i que molta gent no l'hi considera. Diu Badia i Margarit al pròleg del *Diccionari de la llengua catalana* de l'Institut d'Estudis Catalans (DIEC) que «molta gent, que es mostra preocupada per saber si un mot és o no és al diccionari, sovint sembla indiferent, en canvi, davant allò que el diccionari li diu que un mot significa». Pel que portem dit, ja veiem com és de perillosa la simplificació habitual: saber si una paraula existeix o no, i en el primer cas, saber com s'escriu.

Potser serà possible algun dia que, mitjançant una anàlisi semàntica del context, un programa informàtic sigui capaç d'advertir-nos de la possible inadequació lèxica quan escrivim, posem per cas, *esterilitzat* en comptes d'*estilitzat*, però de moment aquest no és del cas.

Encara més: cal tenir present la diferència entre oració gramaticalment correcta i enunciat amb sentit. Valgui l'exemple clàssic de Chomsky: «Colorless green ideas sleep furiously.» En la mesura que exigim al nostre corrector la capacitat de detectar combinacions semànticament anòmales, no solament fem molt més complicades les anàlisis que aquest ha de fer, sinó que toquem amb el caràcter obert que té l'ús del codi lingüístic. En darrer terme, no hi ha —ni hi pot haver— un giny capaç de discernir amb certesa que allò que efectivament hem dit, per estrafolari que sembli, no és justament allò que volíem dir.

De moment, hem de cenyir-nos a qüestions més elementals, lligades al vocabulari entès com a catàleg.

Tots dos correctors reconeixen com a vàlides les formes contingudes en les edicions actualitzades del *Diccionari general de la llengua catalana*, de Pompeu Fabra, i del *Diccionari de la llengua catalana* d'Enciclopèdia Catalana (DLC).

L'aparició del DIEC ha comportat la incorporació de termes nous, o la modificació ortogràfica d'alguns dels antics. Un dels correctors els ha incorporats ja; l'altre, no encara. En aquest darrer, els mots que ja es trobaven en el DLC hi estan incorporats (és el cas de *maquis*). Però termes com ara *bandolerisme*, *urbanisme* o *oceanologia*, absents del DLC i presents al DIEC, no hi han estat incorporats encara. El mateix passa amb el terme *rivet*, que en el DIEC ha canviat d'ortografia. Sembla clar que caldria actualitzar el repertori de les paraules vàlides.

Els correctors comparteixen un problema amb els diccionaris, que és el de l'amplitud d'incorporació explícita de derivats i compostos. Així, un terme com *declinable* s'ha incorporat a tots dos correctors. *Ritualment* ha estat inclòs en l'un, però no en l'altre. I termes com ara *desprogramar* i els augmentatius i diminutius, com *paràs* o *caminet*, no estan inclosos en cap dels dos.

Els barbarismes com *gordo*, *concluir* o *lograr* són detectats. Però només quan la forma correcta és molt semblant a la incorrecta, el corrector és capaç de suggerir-nos la forma adient: és el cas d'*acondiconar*, que

remet a *condicionar*. Novament la tècnica emprada és la permuta de lletres, i d'aquesta manera, és clar, mai no obtindrem *aconseguir* a partir de *lograr*.

Abans s'ha suggerit una major atenció a les terminacions; si això fos possible, també *concluir* remetria a *concloure*.

La detecció de barbarismes té, en el cas del català, una importància especial que potser no té en altres idiomes, per la seva abundor i freqüència; l'usuari pot esperar que el corrector no solament detectarà l'errada sinó que farà el suggeriment adient.

Convindria, doncs, a partir dels bons repertoris de barbarismes que hi ha al mercat, dotar els correctors d'una eina suplementària que contingués les equivalències d'aquells que encara són de presència corrent en els escrits, l'eliminació dels quals no pot confiar-se al simple mètode habitual de combinatòria de lletres.

Consideració final

En la mesura que la tècnica informàtica ho permetrà, es podran aconseguir algunes millores importants dels correctors, com les

indicades. Però aquest suggeriment adreçat als fabricants comporta una reflexió per als usuaris i especialment per als qui tenen per funció d'assessorar-los. Un corrector, per perfecte que sigui, mai no podrà suplir un coneixement insuficient de la llengua: amb el seu ús s'ha de poder arribar a desfer amb fiabilitat dubtes aïllats, però no es pot resoldre mai una insuficiència radical d'un escrit; per consegüent la competència lingüística bàsica de l'usuari és imprescindible.

