



ELS CORRECTORS AUTOMÀTICS EN LÍNIA DE LLENGUA CATALANA COM A DIFUSORS DE LA NORMATIVA I LA SEVA CONTRIBUCIÓ EN LA CONFIGURACIÓ DE L'ESTÀNDARD

Helena Borrell Carreras*

Resum

Aquest article aborda el paper dels correctors automàtics en línia de llengua catalana com a agents estandarditzadors i difusors de la normativa. Davant l'auge d'aquests sistemes, sobretot en el món educatiu, és necessari esbrinar l'impacte que tenen en la llengua escrita. Aquest estudi, doncs, investiga la difusió de la normativa gramatical i d'un estil determinat a partir de les dades proporcionades pel corrector LanguageTool i una experiència amb estudiants universitaris del grau de Traducció i Interpretació de la Universitat Autònoma de Barcelona (UAB). L'objectiu és avaluar el grau i el motiu d'acceptació de les propostes de correcció que fan els correctors automàtics per tal de reflexionar sobre quin model de llengua es promou i l'impacte que té en la configuració de l'estàndard.

Paraules clau: corrector automàtic; llengua estàndard; normativa; gramàtica; estil.

ONLINE CATALAN GRAMMAR CHECKERS AS AGENTS OF LINGUISTIC NORM DISSEMINATION AND THEIR CONTRIBUTION TO SHAPING STANDARD LANGUAGE

Abstract

This article explores the role of online Catalan grammar checkers as standardising agents and disseminators of linguistic norms. With the increasing use of these systems, particularly in education, it is necessary to assess their impact on written language. The study investigates how grammatical norms and specific stylistic choices are disseminated, drawing on data provided by the grammar checker LanguageTool and an experience with university students on the Bachelor's Degree in Translation and Interpreting at the Universitat Autònoma de Barcelona (UAB). The aim is to evaluate the level of acceptance of the suggestions made by grammar checkers and the reasons behind accepting them in order to reflect on the language model promoted by the tool and its impact on shaping standard Catalan.

Keywords: grammar checker; standard Catalan; linguistic norms; grammar; style.

* Helena Borrell Carreras, professora associada del Departament de Filologia Catalana de la Universitat Autònoma de Barcelona. helena.borrell@uab.cat  [0001-8346-4369](https://orcid.org/0001-8346-4369)

Citació recomanada: Borrell Carreras, Helena. (2025). Els correctors automàtics en línia de llengua catalana com a difusors de la normativa i la seva contribució en la configuració de l'estàndard. *Revista de Llengua i Dret, Journal of Language and Law*, 84, 277-290. <https://doi.org/10.58992/rld.i84.2025.4432>

Sumari

- 1 Introducció
- 2 Implantació de la normativa i llengua estàndard
 - 2.1 La normativa i l'estàndard
 - 2.2 La difusió de la normativa
- 3 Els correctors automàtics: evolució i funcionament
- 4 Metodologia
 - 4.1 Aspectes gramaticals i d'estil
- 5 Resultats
 - 5.1 Dades d'acceptació de LanguageTool
 - 5.2 Motius pels quals s'accepten els canvis
- 6 Conclusions
- 7 Agraïments
- 8 Referències

1 Introducció

Els correctors automàtics han esdevingut una eina clau en la correcció i millora de textos escrits. La dependència creixent d'aquestes eines ha impulsat la investigació i n'ha millorat el desenvolupament en els últims anys, sobretot arran de la incorporació de la intel·ligència artificial (IA) en aquesta tecnologia. Si bé l'ús dels correctors automàtics en català és molt elevat en l'àmbit de l'ensenyament,¹ encara hi ha poca recerca sobre com es fan servir i quins efectes tenen en l'aplicació de la normativa i la configuració de l'estàndard. Aquest estudi, doncs, vol emplenar aquest buit analitzant l'impacte dels correctors automàtics en línia de llengua catalana en els textos escrits en català.

Els correctors en línia gratuïts i en català són LanguageTool i Softcatalà. Segons el web de Softcatalà, aquest corrector rep quaranta milions de visites l'any.² Davant l'alt nombre de persones que utilitzen aquestes eines, cal preguntar-se quin model de llengua difonen. Si bé en l'àmbit de l'ortografia els correctors automàtics recullen les disposicions de l'*Ortografia catalana* (Institut d'Estudis Catalans, 2017) i les seves propostes de correcció les segueixen fil per randa, les recomanacions pel que fa a la gramàtica i l'estil fan que el text resultant pugui adoptar certes formes en detriment d'altres que també són acceptables. En el cas de la gramàtica, els correctors, com veurem en les dades d'aquest estudi, solen fer propostes de correcció per a qüestions que la *Gramàtica de la llengua catalana* (GIEC) (Institut d'Estudis Catalans, 2016) no condemna, malgrat que no són les formes preferibles en registres formals. En alguns casos aquests suggeriments de correcció es presenten com a opcions preferibles i, en altres, com a única solució correcta. Pel que fa a l'estil, hi ha propostes de correcció que donen preferència a un mot sobre un altre,³ de manera que hi ha formes —algunes recollides per la GIEC i altres, no— que es descarten per prioritzar-ne d'altres que recomanen alguns llibres d'estil i que, en alguns casos, es consideren més genuïnes. També hi ha casos en què el corrector detecta que hi pot haver un abús de determinats mots o construccions i recomana canviar-los (per exemple, *realitzar* es marca per advertir d'un possible ús abusiu i es proposa utilitzar *fer*).

Mitjançant aquest estudi, doncs, també volem saber si els correctors automàtics són elements que contribueixen a difondre un determinat estàndard, com també fan els mitjans de comunicació i l'escola. Una de les preguntes que es plantegen en el debat sobre la qualitat del català en el marc del Pacte Nacional per la Llengua és amb quins instruments es pot incidir sobre la qualitat lingüística més enllà del sistema escolar (Generalitat de Catalunya, 2021, p. 4). Amb el fi de contribuir al debat, suggerirem la necessitat que es tinguin en compte aquestes eines per avaluar la implantació de la normativa i la qualitat de la llengua.

2 Implantació de la normativa i llengua estàndard

2.1 La normativa i l'estàndard

Malgrat que no tothom està d'acord que siguin dos conceptes diferents (Ginebra, 2018), distingim entre estàndard i normativa per entendre que en la fase de planificació del corpus d'una llengua hi ha la selecció del corpus i codificació (que esdevé la normativa) i la difusió i implantació (l'estàndard que resulta de

1 L'editor, traductor i enginyer Jaume Ortola (comunicació personal, 9 de gener de 2025) disposa de dades que avalen que els correctors es fan servir sobretot en l'àmbit de l'ensenyament, perquè en el període de vacances escolars en disminueix l'ús dràsticament.

2 Les dades han estat proporcionades per Jaume Ortola, que forma part de l'equip de tots dos sistemes. No disposem de dades d'usuaris de LanguageTool.

3 Els correctors en línia de llengua catalana encara no ofereixen recomanacions per reformular frases que poden ser poc clares, cosa que ja fan correctors en altres llengües. Aquesta funció s'incorpora mitjançant l'entrenament automàtic amb IA, i actualment s'està treballant perquè es pugui implementar, segons Ortola (comunicació personal, 9 de gener de 2025).

l'aprenentatge i acceptació de les normes per part dels usuaris). La llengua estàndard és l'àmbit principal d'aplicació de la normativa (Institut d'Estudis Catalans, 2019a), però els agents estandarditzadors no difonen unes normes homogènies: en primer lloc, perquè hi ha molts aspectes gramaticals de la normativa que presenten opcionalitat i jerarquització (unes formes es prefereixen a d'altres) i, en segon lloc, perquè hi ha agents estandarditzadors que tenen un criteri propi, sovint reflectit en un llibre d'estil que recull formes no normatives que poden arribar a esdevenir part de la llengua estàndard. Un exemple n'és el portal *ésAdir*, de la Corporació Catalana de Mitjans Audiovisuals, que ha esdevingut un recurs de referència per a molts usuaris i que recull peces lèxiques i algunes construccions gramaticals que la normativa no accepta en registres formals.

L'Institut d'Estudis Catalans, a l'hora de mesurar la implantació de la normativa, té en compte la llengua estàndard. Amb la finalitat d'avaluar si la normativa s'assimila, la Secció Filològica va engegar l'any 2020 l'Observatori de la Qualitat Lingüística, que avalua la qualitat de la llengua en tots els àmbits de la comunicació pública, tenint en compte si aquesta llengua segueix les disposicions normatives. Tot i que els informes que ha publicat fins ara provenen de mostres extretes de programes ràdio, televisió i campanyes electorals, el seu objectiu és arribar a totes les capes de la societat (Salvanyà, 2023).

Si la recerca en normativa s'ha d'ocupar dels problemes de comunicació que tenen els usuaris amb aquestes normes (Costa Carreras, 2013, p. 278), és essencial que s'investiguin els mitjans pels quals es difonen, ja que, en el procés d'estandardització d'una llengua, en condicions sociolingüístiques favorables, la difusió de la normativa hauria d'afavorir que l'estàndard tingués la qualitat desitjada. Com que el fet que la normativa s'implanti en la llengua estàndard es considera sinònim de qualitat lingüística en aquesta varietat (Costa Carreras, 2023), cal determinar fins a quin punt els agents que difonen la normativa són eficaços en la seva implantació.

2.2 La difusió de la normativa

Tradicionalment, la planificació lingüística ha descuidat l'avaluació de les mesures adoptades en la codificació d'una llengua i ha cregut que aspectes com l'estandardització o l'enriquiment estilístic eren qüestions tècniques que es resolien fàcilment sense necessitat de revisió posterior (Vila, 2016). Els models tradicionals de planificació lingüística, com el de Haugen (1983), consideraven que la intervenció s'havia de fer de dalt a baix i no tenien en compte el retorn després de la difusió de la varietat codificada. A partir dels anys vuitanta es va comprendre que hi havia més agents que prenen part en el procés perquè, entre altres raons, les propostes no sempre s'adoptaven tal com s'havia previst (Liddicoat i Taylor-Leech, 2020).

Davant la necessitat d'avaluar la reestandardització de la llengua catalana i de monitoritzar com s'ha implementat, especialment des de la publicació de la GIEC (Costa Carreras, 2020), cal avaluar-ne l'ús, incloent-hi el que en fan els agents encarregats de fer complir la normativa. En aquest context, Vila i Sendra (2018) han identificat diferents vies de difusió, que combinen les formes tradicionals i les digitals i que sovint estan interconnectades. Aquests mitjans són llibres, articles de premsa, revistes, programes de ràdio i televisió, pàgines web, blogs i xarxes socials. Amb aquest estudi volem contribuir a aquesta llista amb un nou agent de difusió de la normativa: els correctors automàtics, que han esdevingut peces de suport clau en l'escriptura de textos.

3 Els correctors automàtics: evolució i funcionament

Sovint hi ha la creença que la IA suplantarà la feina de les persones en l'àmbit lingüístic i, de fet, el debat sobre l'automatització de les tasques de traducció és ben viu (Vieira, 2018; Moneus i Sahari, 2024; Zanaty, 2024). En el cas dels correctors automàtics, però, encara presenten moltes limitacions que fan que no siguin

al cent per cent fiables i que calgui, ara com ara, la revisió humana dels textos. Malgrat això, continuen sent eines extremament útils, tant per als aprenents d'una llengua com per als usuaris avançats. Els inicis de la correcció automàtica els trobem als anys vuitanta, però aquests sistemes han evolucionat molt des de llavors. L'enfocament clàssic de la correcció automàtica ha estat dissenyar manualment regles gramaticals. Per això, les eines de correcció de text integrades als processadors de textos dels anys vuitanta i noranta eren relativament simples i se centraven en la detecció de faltes d'ortografia més que en errors gramaticals o estilístics complexos. Els correctors basats en regles són eficients per a textos breus i ben estructurats, però els seus models són poc flexibles quan es tracten textos més complexos o amb ambigüitat semàntica (Soni i Thakur, 2018).

Els avenços en processament de llenguatge natural (PLN) han fet que es desenvolupin correctors entrenats amb motors d'aprenentatge automàtic que milloren molts aspectes. Aquests mètodes fan servir un corpus per fer anàlisis estadístiques sobre els textos i detectar-hi i corregir errors automàticament. Tot i això, a diferència dels sistemes basats en regles, és difícil explicar els errors que resulten d'aquests sistemes, controlar-ne el resultat i fer-hi canvis si és necessari. Per això es creu que els models híbrids, que combinen les regles i l'aprenentatge automàtic, són els més adequats per millorar el rendiment del sistema (Soni i Thakur, 2018). Fa temps que s'ha demostrat que es pot utilitzar un corpus de textos per entrenar el sistema a identificar patrons de frases correctes, i els resultats es poden filtrar aplicant regles manualment (Chodorow et al., 2007).

En el cas del català, els dos correctors automàtics gratuïts i en línia, LanguageTool i Softcatalà, ara com ara estan basats en regles. Tot i això, s'estan fent proves per introduir-hi la IA per tal que detectin problemes de puntuació i millorin l'estil oferint propostes de correcció sobre algunes qüestions com ara les frases massa llargues, l'abús de nominalitzacions o els gerundis inadequats (Ortolà, comunicació personal, 9 de gener de 2025). LanguageTool és una eina que, segons els seus creadors, és molt popular perquè, instal·lada com a extensió al navegador, fa la correcció en qualsevol document en línia. A més, té un diccionari personalitzat al qual els usuaris registrats amb un compte poden afegir termes. L'eina detecta errors ortogràfics, gramaticals i d'estil (marcats amb colors diferents) i fa recomanacions de correcció amb un missatge i un enllaç a un recurs lingüístic en línia. Softcatalà és una associació sense ànim de lucre dedicada al foment de les tecnologies en català que ha rebut diversos premis com a reconeixement per la seva tasca i que ofereix un corrector al web que funciona de manera molt similar a LanguageTool.

D'altra banda, la correcció lingüística és molt complexa perquè conté moltíssimes variables. No només cal tenir en compte les regles gramaticals, sinó també el registre, la varietat geogràfica i l'estil. Els programadors han de prendre moltes decisions i fer canvis, ja que els usuaris aporten informació, amb l'objectiu de minimitzar els falsos positius i falsos negatius (Ortolà, comunicació personal, 9 de gener de 2025). Quan aquestes eines incloguin plenament la IA hi podria haver conseqüències importants per a la llengua estàndard, perquè quan s'entrena un sistema i es posa a disposició dels usuaris hi ha poc control dels resultats. Si bé les regles dels correctors es poden desconnectar si no funcionen, és difícil saber si el motor aprendrà algun aspecte de la llengua que després no es podrà modificar fàcilment.

De fet, ja hi ha estudis que mesuren l'impacte de la IA en la llengua i les seves conseqüències en certes varietats dialectals (Goldstein et al., 2024). En una línia semblant, Smith et al. (2024) exploren com els models de llengua generada per IA reforcen la ideologia de la llengua estàndard i perpetuen les jerarquies entre llengües, cosa que afavoreix la discriminació de les varietats no estàndards. Si ens volem anticipar a això, caldrà entrenar models perquè tinguin en compte la variació funcional i geogràfica del català i preveure el possible biaix de la IA, ja que podrien contribuir a crear un estàndard que encara homogeneïtzi més la llengua i amb el qual els parlants de la perifèria lingüística no se sentissin identificats.

4 Metodologia

Aquest estudi combina els mètodes quantitatius i qualitatius per investigar el grau d'acceptació de les propostes de correcció i els motius que hi ha darrere de l'acceptació o no de les recomanacions dels correctors automàtics. Ens hem centrat exclusivament en LanguageTool perquè és l'eina de la qual disposem de dades d'ús actualment (recollides el mes de desembre de 2024).

Per portar a terme aquest estudi hem comprovat, primer de tot, si el corrector detectava tots els problemes gramaticals i d'estil d'un conjunt de frases (vegeu 4.1) i si detectava qüestions que no tenien cap problema. És a dir, hem identificat els falsos positius —casos en què el corrector detecta un error quan no n'hi ha cap— i els falsos negatius —casos en què el corrector no detecta cap error quan n'hi ha un— d'un conjunt de frases per saber el grau de fiabilitat de l'eina en aquest context particular. A partir de dades aportades pels gestors de LanguageTool, hem identificat el grau d'acceptació de les propostes de correcció.

Per conèixer els motius que hi ha darrere de l'acceptació o no d'una forma determinada, hem fet servir el mateix conjunt de frases. Hem seleccionat un grup de nou alumnes de quart curs del grau de Traducció i Interpretació de la Universitat Autònoma de Barcelona (UAB) amb l'objectiu que fessin les correccions que creguessin convenients amb l'ajut del corrector automàtic i amb accés a fonts disponibles en línia (enllaç a *Optimot* i a altres recursos que recomana el corrector per ampliar la informació i prendre una decisió fonamentada). Els estudiants estaven matriculats a l'assignatura Llengua d'especialitat per a traductors i intèrprets (català), i abans havien superat l'assignatura Llengua catalana per a traductors i intèrprets, fet que els situava en un nivell aproximat al C2 del Marc europeu comú de referència per a les llengües (MECR). Com que un dels objectius del curs era que produïssin textos especialitzats amb correcció lingüística, hem partit de la hipòtesi que acceptarien la majoria de recomanacions que els fes el corrector sempre que els semblessin adequades tenint en compte els seus coneixements i les consultes que poguessin fer a altres recursos. Amb la finalitat de saber si les havien acceptades, hem demanat als estudiants quins canvis havien fet en el conjunt de frases seleccionat per a l'estudi i per què (si havien llegit la recomanació i per quin motiu exactament n'havien fet cas o no). Això ens permetrà conèixer quines són les motivacions davant l'acceptació d'un suggeriment de correcció i mesurar la força perlocutiva del missatge que dona el corrector automàtic.

4.1 Aspectes gramaticals i d'estil

Les frases que hem seleccionat per a aquest estudi tracten aspectes gramaticals i d'estil. Pel que fa a les qüestions gramaticals, hem triat aspectes de la gramàtica que presenten opcionalitat des de la publicació de la GIEC (2016), perquè o bé hi ha formes preferibles en registres formals o bé hi ha més d'una forma vàlida però cap de preferent. Són les següents:

G1.⁴ Ús del demostratiu sense nom explícit en posició de subjecte

Segons la GIEC (§ 16.4.4), l'ús anafòric del demostratiu sense el nom en posició de subjecte “fa molt carregós el text” i és “preferible que el subjecte sigui el-líptic”. Hi ha una opció preferible, però totes dues són normatives.

Frase per corregir: *No m'agrada la casa que s'han comprat, ja que aquesta és massa fosca.*

G2. Construcció “al + infinitiu” amb valor temporal

4 Hem fet servir la lletra G i un número per indicar que es tracta d'un aspecte gramatical; més avall hem usat E i un número per indicar que es tracta d'un problema d'estil. Analitzarem les dades obtingudes amb aquestes referències.

La GIEC (§ 31.4.2.1) diu: “Al costat de la construcció ‘en + infinitiu’, a la qual s’ha donat preferència en els registres formals, hi ha també la construcció ‘al + infinitiu’”. Per tant, totes dues solucions són vàlides des d’un punt de vista normatiu.

Frase per corregir: *A l’arribar a casa, va fer el sopar.*

G3. Construccions de relatiu amb infinitiu

La GIEC (§ 29.3.5) diu: “En certs casos en què el sintagma nominal té un caràcter no específic i, per tant, no denota cap entitat concreta, les construccions amb preposició poden alternar amb una oració de relatiu d’infinitiu, com en *Busquen un hotel on passar la nit* i *Busquen un hotel per a passar la nit*. De les dues solucions, la construcció amb la preposició és el recurs més corrent i el preferible”. Per tant, totes dues solucions són normatives, però hi ha una forma preferible.

Frase per corregir: *Què fer a Barcelona a l’estiu.*

G4. Perífrasi *portar* + SN [temps] + gerundi/participi/SP

Per a aquest fenomen, la GIEC (§ 31.5.4) també jerarquitzava les solucions: “Hi ha una construcció temporal relativament recent que equival semànticament a l’oració impersonal amb el verb *fer*, que és la construcció més genuïna”.

Frase per corregir: *Porto molts anys esperant aquesta oportunitat.*

G5. Ús de possessiu en objectes inanimats

La GIEC (§ 16.5.3) diu que “resulta forçat” fer servir el possessiu en objectes inanimats i dona exemples de casos en què “prescindim del possessiu si queda a prop”. En aquest punt, Ginebra (2017, p. 32) opina que la GIEC “no és concloent”.

Frase per corregir: *Van comprar la casa i van iniciar la seva restauració al cap d’un any.*

G6. Relatiu *on* amb valor temporal

La GIEC (§ 27.6.2) diu: “L’adverbi relatiu *on* no té en català un ús tan extens com en altres llengües romàniques, de manera que no és gaire habitual amb valor temporal”. Malgrat indicar-ne la baixa freqüència d’ús, no es diu que sigui inacceptable. Tanmateix, si consultem la *Gramàtica essencial de la llengua catalana* (GEIEC) (Institut d’Estudis Catalans, 2018, § 2.4.2) i la *Gramàtica bàsica i d’ús de la llengua catalana* (GBU) (Institut d’Estudis Catalans, 2019b, § 38.5), trobem una contradicció amb les disposicions de la GIEC, ja que tant la GEIEC com la GBU afirmen que el relatiu *on* no té un ús temporal. Per tant, no podem extreure cap conclusió de si aquest ús és normatiu.

Frase per corregir: *Era d’una època on sempre s’havia de callar.*

G7. L’expressió de les hores en els dialectes central i nord-occidental

Sobre aquesta qüestió, la GIEC (§ 31.2.5.1) diu: “Les fraccions horàries s’expressen en la major part del central i el nord-occidental amb el sistema de campanar, que és específic de la llengua catalana. En la resta del domini lingüístic és d’ús general el sistema de rellotge”. Com que no podem saber quina part del central no fa servir el sistema de campanar, podem considerar que totes dues expressions de les hores són vàlides en aquest dialecte.

Frase per corregir: *Són les vuit i mitja del matí.* (Dit per un parlant del dialecte central)

G8. Concordança de participi amb el pronom acusatiu

La GIEC (§ 13.5.2.4) diu que “la concordança entre el pronom d’acusatiu i el participi se sol mantenir en els registres formals”. Això vol dir que tant fer la concordança com no fer-la són dues possibilitats normatives.

Frase per corregir: *S’han fet una casa i l’han pagat al comptat.*

G9. Forma reforçada del pronom *es* davant d’un verb començat per [s]

La GIEC (§ 8.3.2) diu que “se sol usar la forma plena *se* davant de verb començat per [s] per a evitar el contacte de sibilants”, però és “igualmente acceptable l’ús de *es*”. Per tant, totes dues solucions són vàlides.

Frase per corregir: *Des de casa, es sent el soroll de les obres.*

G10. Manteniment de la preposició *en* en una subordinada substantiva d’infinitiu

Segons la GIEC (§ 26.5.2.1), “el manteniment de *en* és acceptable, però la preposició preferible en els registres formals és *a*”. Per tant, hi ha una opció preferible, però l’altra també és normativa.

Frase per corregir: *Estava interessat en comprar la casa.*

G11. Ús de la preposició *a* en la construcció “*d’aquí/ací (a) + espai de temps*”

La GIEC (§ 19.3.4.3b) diu que “és habitual” suprimir la preposició, però “se sol mantenir en els registres formals”. Totes dues opcions són acceptables.

Frase per corregir: *Comprarem la casa d’aquí un any.*

En relació amb els aspectes d’estil, hem seleccionat paraules i construccions que la normativa accepta en registres formals o a les quals no fa esment (excepte en el cas E4, en què condemna la construcció explícitament), però que el corrector suggereix corregir.

E1. *No obstant* (en lloc de *no obstant això*)

Aquesta construcció està recollida a la GIEC (§ 25.3.2d), però alguns llibres d’estil la condemnen, com el de la Universitat de Barcelona o del *Diari de Barcelona*.

E2. *Realitzar*

De vegades s’utilitza *realitzar* com un verb semànticament buit, malgrat que té un significat precís, i alguns llibres d’estil recomanen evitar-lo en aquest sentit, com el de la Universitat de Barcelona i ésAdir, així com l’Optimot.

E3. *Creuar* (en el sentit de ‘travessar’)

S’ha considerat un calc del castellà, tot i que el *Diccionari de la llengua catalana* (DIEC2) (Institut d’Estudis Catalans, 2007) recull el verb en aquest sentit.

E4. *A nivell (de)* en sentit figurat

La GIEC (§ 19.9.3.1a) afirma que aquest ús no és acceptable en registres formals, però s’utilitza col·loquialment.

E5. *Tipus* (en lloc de ‘mena’)

Se’n fa un ús abusiu, segons algunes fonts com ésAdir, i alguns autors defensen que *tipus* i *mena* no són sinònims (Marquet, 2003).

E6. *Així com també*

La normativa no condemna explícitament aquesta expressió, però hi ha fonts que afirmen que és redundant, com ara l'Optimot o el llibre d'estil de la Diputació de Barcelona.

E7. *Al meu parer*

El DIEC2 recull aquesta expressió, però alguns lingüistes la veuen com un calc del castellà pel fet de col·locar el possessiu abans del substantiu (Ruaix, 2003).

E8. *Ja sigui*

La GIEC no recull específicament aquesta construcció i hi ha llibres d'estil que la condemnin (com el de la Diputació de Barcelona), però sí que es recull com a part del text en materials normatius.

E9. *De ben segur*

L'expressió no apareix al DIEC2 ni està llistada entre els adverbis modals de la GIEC, però sí que està recollida al *Diccionari essencial de la llengua catalana* (DEIEC) (Institut d'Estudis Catalans, 2023). A més, l'Optimot diu sobre aquesta qüestió: “Tot i que la construcció *de ben segur* no es documenta en el diccionari normatiu, està ben formada i es fa servir per expressar una seguretat absoluta”.

5 Resultats

5.1 Dades d'acceptació de LanguageTool

Dels vint aspectes que hem passat pel corrector automàtic LanguageTool (onze de gramaticals i nou d'estil), s'han detectat errors en totes les frases menys a G1, G3, G5, G7 i G8. Si tenim en compte el que hem vist més amunt, veiem que la normativa accepta aquestes cinc frases tal com són, malgrat que es tracta de construccions no preferibles o menys habituals en registres formals. També hem de tenir en compte que la normativa no condemna explícitament cap dels errors detectats en les altres frases, excepte a E4 (*a nivell de*). En el cas de G6 (relatiu *on*), com ja hem explicat, no podem afirmar que la normativa l'accepti o el condemni.

Pel que fa al grau d'acceptació de les propostes de correcció dels aspectes que s'han detectat, segons les dades de desembre de 2024, observem que els suggeriments s'accepten de manera majoritària, però alguns tenen un grau més alt d'acceptació que d'altres:

Taula 1

Percentatge d'acceptació de suggeriments de correcció de LanguageTool de desembre de 2024 (llista en ordre descendent)

Aspectes gramaticals i d'estil	Grau d'acceptació
G11: preposició <i>a</i>	84,60 %
G10: preposició <i>en</i>	79,80 %
G2: <i>al</i> + infinitiu	79,50 %
E9: <i>de ben segur</i>	77,10 %
G6: relatiu <i>on</i>	76,40 %
E6: <i>així com també</i>	76,00 %
G9: pronom <i>es</i>	75,20 %
G4: <i>portar</i> + gerundi	74,70 %
E1: <i>no obstant</i>	73,10 %
E5: <i>tipus</i>	72,80 %
E8: <i>ja sigui</i>	68,80 %

Aspectes gramaticals i d'estil	Grau d'acceptació
E7: <i>al meu parer</i>	68,70 %
E3: <i>creuar</i>	63,10 %
E2: <i>realitzar</i>	57,80 %
E4: <i>a nivell (de)</i>	56,20 %

Font. Elaboració pròpia a partir de les dades de LanguageTool.

Disposem, també, de dades que ens permeten saber quins són els aspectes detectats més sovint. A continuació, presentem el nombre de casos oberts per a cada qüestió durant tot el mes de desembre de 2024:

Taula 2

Llista d'aspectes lingüístics segons el nombre de casos oberts durant el desembre de 2024 (en ordre descendent)

Aspectes gramaticals i d'estil	Nombre de casos oberts
E2: <i>realitzar</i>	76.367
G10: preposició <i>en</i>	75.454
E4: <i>a nivell (de)</i>	70.526
G9: pronom <i>es</i>	69.982
G2: <i>al</i> + infinitiu	60.993
E5: <i>tipus</i>	18.457
E8: <i>ja sigui</i>	17.693
G6: relatiu <i>on</i>	15.860
E1: <i>no obstant</i>	10.668
G4: <i>portar</i> + gerundi	6.124
E6: <i>així com també</i>	5.238
E3: <i>creuar</i>	3.009
E9: <i>de ben segur</i>	2.741
G11: preposició <i>a</i>	1.994
E7: <i>al meu parer</i>	1.615

Font. Elaboració pròpia a partir de les dades de LanguageTool.

Altres casos que s'obren molt sovint —és a dir, el corrector detecta un error i el marca com a tal— i que tenen un grau molt alt d'acceptació són els ortogràfics, que aquí no analitzem, però que demostren que els usuaris prenen decisions que podríem qualificar de “raonables”, ja que les qüestions ortogràfiques no presenten opcionalitat i, per tant, solen corregir la paraula quan se'ls ho recomana. Per exemple, hi ha el cas de l'accent a la tercera persona del singular del present del verb *ser* (*és*), que el desembre de 2024 va registrar 769.926 casos, dels quals se'n va acceptar el 88,30 %.

Pel que fa a la freqüència en la qual s'obren casos, no hi ha cap patró que ajudi a identificar si els aspectes més problemàtics són relatius a la gramàtica o a l'estil, però sí que s'observen tendències pel que fa a l'assimilació de la normativa, per exemple, sobre el canvi de preposició davant d'infinitiu (G10). En aquest sentit, una avaluació sobre la implantació de la normativa ja advertia que aquest aspecte sintàctic és de difícil implantació (Ginebra, 2015), tot i que avui dia la GIEC admet opcionalitat i el manteniment de la preposició original del verb és acceptable. Pel que fa a l'estil, com es desprèn de la taula, l'abús de *realitzar* (E2) i l'ús de l'expressió *a nivell (de)* (E4), considerada col·loquial per la GIEC, apareixen molt sovint. No podem extreure conclusions sobre quins casos són més problemàtics, perquè les dades poden variar d'un mes a l'altre, però sí que podem afirmar que hi ha certes qüestions d'estil que són força recurrents. En el

cas d'E4, també podem inferir que la normativa no s'assimila —per desconeixement, potser—, ja que la GIEC no accepta aquesta forma en registres formals però té un alt nombre de casos oberts.

Pel que fa a l'extrapolabilitat de les dades, hi ha alguns elements que cal tenir en compte. Algunes regles poden tenir variacions; per exemple, en *al meu parer* també s'inclouen altres construccions amb el possessiu anteposat, com ara *al meu càrrec*. D'altra banda, cal considerar la finalitat del text, ja que el corrector actua en qualsevol text que estigui en línia, que pot ser un missatge informal o una nota personal. En aquest sentit, podria ser que no s'acceptessin els suggeriments perquè no interessa corregir el text i no pas perquè s'interpretin com a erronis o poc raonables.

Per tot això, les dades no poden mostrar amb exactitud el total d'usuaris que han fet el canvi ni per què l'han acceptat o no, però sí que ens permeten tenir una idea general de quins aspectes del conjunt de frases seleccionat generen més positius i quins suggeriments se solen acceptar més sovint. Per investigar els motius pels quals s'accepten algunes propostes més que d'altres, hem portat a terme l'experiència que veurem a continuació.

5.2 Motius pels quals s'accepten els canvis

Una hipòtesi sobre el motiu pel qual s'accepten els suggeriments de canvi és la força perlocutiva del missatge que dona el corrector automàtic. Així, en la majoria de casos estudiats el corrector afirma que és preferible fer el canvi que recomana. Els errors marcats com a qüestions d'estil solen presentar les fórmules “és preferible”, “és recomanable” o “val més”, entre altres. També és rellevant destacar que només el canvi de preposició *en* (G10) i l'ús de la preposició *a* en *d'aquí a* (G11) es marquen com un problema gramatical. La resta d'aspectes es consideren d'estil, cosa que fa pensar que les regles s'han fet tenint en compte l'opcionalitat de la GIEC. Per al cas G10 apareix el missatge següent: “Canvi de preposició. Probablement cal canviar la preposició”, i en el cas de G11 el corrector diu: “Expressió incorrecta. Cal escriure 'd'aquí a'”. Recordem que aquests dos aspectes també presenten opcionalitat segons la GIEC.

Si ens fixem en les dades del grau d'acceptació, veiem que els dos darrers casos són els que tenen el percentatge més alt (tot i que no hi ha una gran diferència entre aquests i el següent). Per això, podem conjecturar que el motiu és la força perlocutiva del missatge, ja que són els dos únics casos en què la voluntat prescriptiva és evident. Tot i això, Ortolà afirma que no ha observat mai un canvi significatiu en l'acceptació dels usuaris si es modifica el missatge, per la qual cosa inferim que no hi presten gaire atenció. Segons Ortolà, hi ha pocs usuaris que facin clic a l'enllaç que inclouen alguns suggeriments de correcció (normalment l'Optimot i en alguns casos és Adir) per entendre per què hi ha un error, la qual cosa fa pensar que molts usuaris probablement accepten o no el suggeriment simplement perquè els sembla més o menys raonable. Com que aquesta informació contradiu la nostra hipòtesi, hem volgut investigar la qüestió més a fons amb l'experiment que hem dut a terme amb el grup d'estudiants.

Hem subministrat un exercici de correcció a través de Google Forms que contenia tots els aspectes gramaticals i d'estil d'aquesta investigació, després que els estudiants s'haguessin instal·lat LanguageTool com a extensió del navegador. Els participants han corregit les frases individualment segons les indicacions del corrector automàtic i d'acord amb els seus coneixements lingüístics. El formulari també incloïa dues preguntes sobre el motiu pel qual havien acceptat o no la recomanació del corrector i si havien consultat els recursos enllaçats al suggeriment.

Les dades mostren que els participants han pres decisions ponderades, ja que han fet canvis a les frases en què el corrector no identificava cap error —en totes menys a G3 (relatiu amb infinitiu), i la meitat dels participants també a G8 (concordança de participi), en ambdós casos per desconeixement que existís una opció preferible en registres formals. Pel que fa a les qüestions gramaticals que sí que detectava el corrector, els participants han fet canvis de manera unànime en tots els casos (tot i que algunes propostes no són

exactament com suggereix el corrector). En relació amb les qüestions d'estil, també han acceptat la majoria de propostes, tot i que amb alguna petita excepció —el 78 % ha aplicat el canvi a E3 (*creuar*) i el 89 %, a E7 (*al meu parer*). Els participants han afirmat que havien acceptat les propostes perquè estaven d'acord amb el canvi suggerit o perquè desconeixien que hi havia una opció més adequada i, en rebre la proposta, l'havien acceptada perquè els havia semblat raonable. Això suggereix que hi ha un alt grau de confiança en el corrector, tot i que la meitat de participants ha consultat altres fonts per confirmar que la tria era adequada (alguns per iniciativa pròpia i altres seguint l'enllaç proporcionat pel corrector).

6 Conclusions

Després d'analitzar les dades d'acceptació dels suggeriments de correcció en els casos estudiats, podem afirmar que la gran majoria de propostes són acceptades pels usuaris i, per tant, que els correctors automàtics són agents de difusió de la normativa que tenen un impacte important en la manera com s'escriuen els textos i en la configuració de la llengua estàndard. Tot i això, no podem obviar que els correctors automàtics no són totalment fiables, ja que no detecten tots els problemes lingüístics que un text pot presentar. Per tant, malgrat l'impacte evident dels correctors en la difusió de la normativa i com a eina per a la millora de la llengua escrita en un registre formal, no podem afirmar que aquesta difusió sigui exhaustiva ni totalment eficaç. Malgrat aquesta constatació, tampoc no podem afirmar que hi hagi cap altre agent difusor de la normativa que sigui eficaç al cent per cent per assegurar la implantació de la normativa ni la qualitat lingüística.

En relació amb els elements estudiats, no hem pogut identificar un patró concret pel que fa a la freqüència d'aparició de certs aspectes lingüístics, però sabem que hi ha qüestions que la normativa no admet i que continuen apareixent en els textos que escriuen els usuaris —per exemple, E4 (*a nivell de*)—, i casos en què la normativa estableix una preferència en registres formals, però els usuaris han utilitzat l'opció no preferible —i sovint condemnada per la normativa abans de la publicació de la GIEC— abans de passar el text pel corrector. En aquests casos, cal destacar que la majoria de propostes de correcció són acceptades i, per tant, podem afirmar que el corrector difon la normativa més “tradicional” o els trets considerats més “genuïns”.

Pel que fa al grau d'acceptació de les propostes, tal com hem apuntat més amunt, els casos que s'accepten en un grau més alt presenten un missatge més obertament prescriptiu (i es tipifiquen com a qüestions gramaticals, no pas d'estil). Inicialment, això ens ha fet pensar que la força perlocutiva del missatge era la causa de l'acceptació més alta del suggeriment, però després de l'experiència duta a terme amb el grup d'estudiants hem constatat que els canvis s'apliquen quan l'usuari els troba raonables, a vegades simplement perquè confia en el corrector i a vegades perquè confirma la tria cercant en una altra font.

Si bé aquestes dades ens permeten entendre l'impacte dels correctors com a difusors de la normativa i agents d'estandardització, aquest estudi té limitacions, com ja hem indicat anteriorment. En primer lloc, l'acceptació de les propostes de correcció és relativa perquè no sabem en quin context es fa servir LanguageTool. En aquest sentit, és possible que l'eina estigui sempre activada però que l'usuari decideixi no fer cas del suggeriment perquè no té la correcció com a prioritat o perquè el registre és informal.

En segon lloc, l'experiència duta a terme amb el grup d'estudiants no és representativa de tots els usuaris, ja que molts poden no tenir els mateixos coneixements lingüístics que l'alumnat del grau de Traducció i Interpretació. Tot i això, l'experiment s'ha fet amb persones que, suposadament, tenen més capacitat de prendre decisions fonamentades atesa la seva formació i, per tant, la tendència podria haver estat de no acceptar les propostes en un grau tan elevat, en contra del que faria un usuari amb menys criteri (pel fet de tenir menys coneixements lingüístics i confiar més en l'eina).

Així doncs, malgrat que les dades puguin presentar un cert biaix, podem afirmar que els correctors automàtics són elements que contribueixen a difondre un model de llengua determinat i, per tant, és important que es tinguin en compte a l'hora d'avaluar la implantació de la normativa i la qualitat de la llengua estàndard. Quan aquestes eines incorporin plenament la IA, caldrà estudiar de nou com és la normativa que difonen i si els usuaris accepten igualment la majoria de canvis proposats.

7 Agraïments

Volem agrair la col·laboració i bona disposició de Jaume Ortolà, ja que la informació sobre el funcionament del corrector automàtic i les dades d'ús aportades han estat essencials per dur a terme aquest estudi. Volem donar les gràcies, també, als revisors d'aquest article per les recomanacions que ens han fet per millorar-lo.

8 Referències

- Chodorow, Martin, Tetreault, Joel i Han, Na-Rae. (2007). [Detection of grammatical errors involving prepositions](#). Dins Fintan Costello, John Kelleher i Martin Volk, *Proceedings of the Fourth ACL-SIGSEM Workshop on Prepositions* (p. 25-30). Association for Computational Linguistics.
- Corporació Catalana de Mitjans Audiovisuals. (2025). [ésAdir. El portal lingüístic de 3cat](#).
- Costa Carreras, Joan. (2023). [Reflexions i respostes sobre la qualitat lingüística del català](#), *Treballs de sociolingüística catalana*, 33, 159-175.
- Costa Carreras, Joan. (2020). Are terminology planning evaluation and language policy and planning evaluation applicable to the evaluation of standardisation? *Current Issues in Language Planning*, 21(1), 1-21. <https://doi.org/10.1080/14664208.2018.1553913>
- Costa Carreras, Joan. (2013). [La recerca sobre normativa: proposta general](#). *Treballs de sociolingüística catalana*, 23, 269-286.
- Diputació de Barcelona. (2021). [Llibre d'estil de la Diputació de Barcelona](#). Unitat Lingüística de la Diputació de Barcelona.
- Generalitat de Catalunya. (2025). [Optimot. Consultes lingüístiques](#).
- Generalitat de Catalunya. (2021). [Pacte Nacional per la Llengua. Qualitat de la llengua](#).
- Ginebra, Jordi. (2018). [Planificació, codificació i estandardització de les llengües](#). [Material docent de la UOC]
- Ginebra, Jordi. (2017). [La nova normativa de l'Institut d'Estudis Catalans. Guia pràctica](#). Publicacions URV.
- Ginebra, Jordi. (2015). [¿Sabem si la normativa sintàctica s'assimila?](#). *Treballs de sociolingüística catalana*, 25, 95-111.
- Goldshtein, Maria, Ocumpaugh, Jaclyn, Potter, Andrew i Roscoe, Rod D. (2024). The social consequences of language technologies and their underlying language ideologies. Dins Margherita Antona i Constantine Stephanidis (ed.), *Universal Access in Human-Computer Interaction - 18th International Conference, UAHCI 2024* (p. 271-290). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-031-60875-9_18

- Haugen, Einar. (1983). The implementation of corpus planning: Theory and practice. Dins Juan Cobarrubias i Joshua A. Fishman (ed.), *Progress in language planning: International perspectives* (p. 269-290). De Gruyter Mouton. <https://doi.org/10.1515/9783110820584.269>
- Institut d'Estudis Catalans. (2023). [*Diccionari essencial de la llengua catalana*](#).
- Institut d'Estudis Catalans. (2019a). [*Criteris generals sobre la llengua normativa i la llengua estàndard. Acord de la Secció Filològica del 15 de novembre de 2019*](#).
- Institut d'Estudis Catalans. (2019b). [*Gramàtica bàsica i d'ús de la llengua catalana*](#).
- Institut d'Estudis Catalans. (2018). [*Gramàtica essencial de la llengua catalana*](#).
- Institut d'Estudis Catalans. (2017). [*Ortografia catalana*](#).
- Institut d'Estudis Catalans. (2016). [*Gramàtica de la llengua catalana*](#).
- Institut d'Estudis Catalans. (2007). [*Diccionari de la llengua catalana*](#). [2a ed.]
- Liddicoat, Anthony i Taylor-Leech, Kerry. (2020). Agency in language planning and policy. *Current Issues in Language Planning*, 22, 1-18. <https://doi.org/10.1080/14664208.2020.1791533>
- Marquet, Lluís. (2003). [*"Mena" i "tipus", mots no sinònims*](#). *Llengua Nacional*, 43, 20.
- Moneus, Ahmed M. i Sahari, Yousef. (2024). Artificial intelligence and human translation: A contrastive study based on legal texts. *Heliyon*, 10(6), e28106. <https://doi.org/10.1016/j.heliyon.2024.e28106>
- Ruaix, Josep. (2003). [*Calcs en els adjectius o pronoms determinatius*](#). *Llengua Nacional*, 43, 15-16.
- Salvanyà, Jaume. (2023, 2 de juliol). [*Per a què serveix l'Observatori de la Qualitat Lingüística?*](#), *Núvol*.
- Sendra, Montserrat, i Vila, Francesc Xavier. (2018). [*La difusió explícita de la normativa: vies i estratègies en la societat de la informació*](#). A Miquel Àngel Pradilla (ed.), *Comunitat lingüística i norma: actes del IV Col·loqui Internacional "La lingüística de Pompeu Fabra"* (Tarragona, 18, 19 i 20 de novembre de 2013), 57-68.
- Smith, Genevieve, Fleisig, Eve, Bossi, Madeline, Rustagi, Ishita i Yin, Xavier. (2024). [*Standard language ideology in AI-generated language*](#). *arXiv:2406.08726*.
- Soni, Madhvi i Thakur, Jitendra S. (2018). A systematic review of automated grammar checking in English language. *ArXiv*, abs/1804.00540. <https://doi.org/10.48550/arXiv.1804.00540>
- Universitat de Barcelona. (2023). [*Llibre d'estil de la Universitat de Barcelona*](#). Serveis Lingüístics de la Universitat de Barcelona.
- Universitat Pompeu Fabra i Ajuntament de Barcelona. (2020). [*Diari de Barcelona \(DdB\): llibre d'estil*](#).
- Vieira, Lucas Nunes. (2018). Automation anxiety and translators. *Translation Studies*, 13(1), 1-21. <https://doi.org/10.1080/14781700.2018.1543613>
- Vila, Francesc Xavier. (2016). [*L'avaluació de la implantació de la normativa: La perspectiva sociolingüística i glotopolítica*](#). *Treballs de sociolingüística catalana*, 25, 129-140.
- Zanaty, Dina G. (2024). The future of human translation in the artificial intelligence era. *Delta University Scientific Journal*, 7(2), 257-274. <https://doi.org/10.21608/DUSJ.2024.320340.1089>