



■ Euskaraz idatzitako testuetan akats ortografikoak ez izatea posible izango da gaurtik aurrera Xuxen zuzentzaile ortografiko berriari esker. Ez da hori, ordea, testu honen kasua.

# Xuxen

Azken hiru urteotan garatutako programa euskararako zuzentzaile ortografikoa da

INAKI ALEGRIA / USURBIL

Euskaraz idatzitako testuetarako zuzentzaile ortografikoa da Xuxen. Testuetako hitzen analisi morfologikoan dago oinarrituta. eta hitz batek analisi posiblerik ez duenean abisu bat azaltzen du Xuxenek. Balizko erroreen aurrean erabiltzaileak aukera desberdinak ditu, besteak beste, ontzat ematea, berak zuzentzea, zuzenketarako proposamenak eskatzea edo, berriro lema horretatik deribatutako forma guztiak ezagun ditzan, hiztegi partikularrean sarrera ematea.

Programaren prestakuntza lan handia izan da nagusiki bi arazoengatik: euskarak duen berezko konplexutasun morfologikoa batetik eta euskararen morfologiari buruzko deskripzio sistematikoaren falta bestetik. Azken hiru urteotan garatutako programa hau UPV / EHUko Informatika Fakultatea, UZEI eta Baionako Hizkia enpresaren arteko elkarlana izan da. Informatika Fakultateko taldeak prototipoaren garapena eta koordinazio-lana egin du, UZEIk berme linguistikoa eman dio eta Hizkiak produktu komertzialaren ardura hartu du. Macintosh-erako bertsioa asteburu honetatik aurrera liburudendetan salgai izango den bitartean, PCrako azarorako izango da prest. Lan honetan ezin da aipatu gabe utzi honako Erakundeen laguntza: IVAP / HAEE, Gipuzkoako Foru Aldundiko Ekonomia Saila eta Euskadi / Akitania lankidetzako programa.

**XUXENEN EZAUGARRI FUNTZIONALAK** H o n a k o hauek dira X u x e n e n

ezaugarri funtzionalak:

Oinarri-makina:

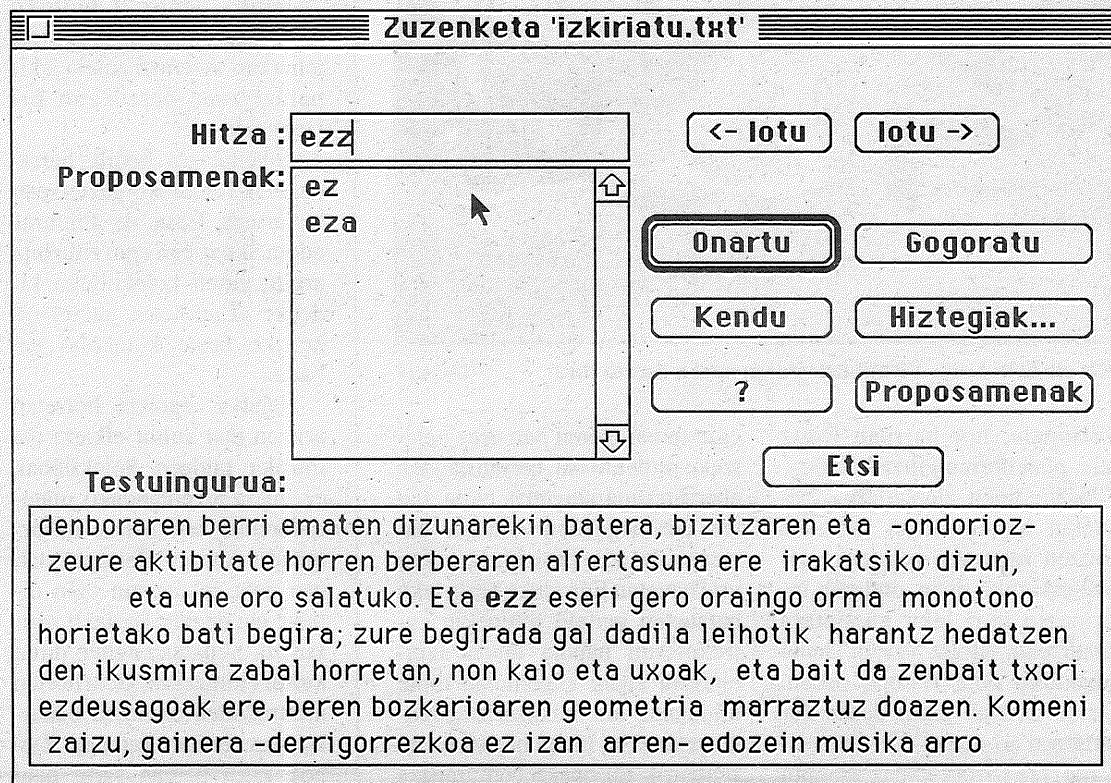
- Macintosh, 7.0 Sistema, 4 M memoria.

Testu prozesadore bateragarriak (testuaren formatea mantentzeko):

- Microsoft Word 4.0, 5.0, 5.1
- Word Perfect 2.0

Aukera nagusiak:

- Testu bat zuzentzea. Dokumentu editatu oso bat egiaztatuta eta zuzentzea.
- Testu sail bat egiaztatzea eta zuzentzeko prestatzea. Ari den bitartean ez da pantaila aurrean izan behar. Bukaeran zuzenketa azkarrerako aukera.
- Zalantzen kontsulta. Editatu ahala hitz solteak kontsultatzea arbelaren bidez.
- Hiztegi pertsonalak osatzea lexiko bereziturako.



Hiztegi orokorra:

- Zabala eta eguneratua: 60.000 sarrera baino gehiago.
- Euskaltzaindiaren azken erabakiak kontutan hartuz.

**XUXENEN DISEINU EZAUGARRIAK** Aipatutako konplexutasun morfologikoa dela eta, hitz bat ondo dagoen ala ez erabakitzeke ezin da beste hizkuntza batzuetarako egiten den bezala hitz-zerrenda bat kontsultatu eta kito; lema batetik abiatuta sor daitezkeen forma zilegiak asko direnez, zerrenda izugarria izango bailitzateke.

Adibidez, izen batetik abiatzen bagara, deklinabide-atzizki bakar bat erantsiz 135 legezko forma lor daitezke (elipsiak kontuan hartzen baditugu zenbaki hau izugarri igotzen da). Gainera, horrela jokatu izan bagenu, era-

Fitxategia	Edizioa
Ireki...	%0
Zuzenketarako prestatu...	%P
Zuzendu arbelekoa...	
Hiztegia ireki...	%H
Laguntza	%?
Irten	%Q

biltzailearen hiztegiaren lema bakarrik sartuz bere deklinabide osoa ulertu beharrean, Xuxenean gertatzen den legez, lema horri dagozkion forma guztiak sartu beharko lituzke banan-banan erabiltzaileak. Aurreko guztia-rengatik hitz zilegiak ondo identifikatzeko analisi morfologikoa ezinbestekoa zen.

Analisi morfologikoa 1983an Helsinkiko Unibertsitateko Koskeniemi irakasleak proposatutako bi mailatako formalismoan dago oinarrituta. Formalismo hau hasiera batean suomierarako proposatu bazen ere,

oso arrakastatsua izan da beste edozein hizkuntzatarako eta euskara bezalako hizkuntza eranskietarako batez ere.

Lexikoak 60.000 sarrera baino gehiago du, datu-base batean gordeta eta 120 azpilexikotan banatuta. Sarrera bakoitzari jarraitze-klase bat egokitzen zaio bere atzean etor daitezkeen atzizkien multzoa definitzen duena. Morfemak biltzean gertatzen diren azaleko aldaketak adierazteko hogeita lau erregela morfologikokoak daude. Hauetako erregela bakoitzak karaktere baten txertaketa, ezabapena edo aldaketa bat noiz gertatzen den adierazten du. Adibidez, zortzigarren erregelan ondoko aldaketa deskribatzen da: lexikoko k letra azaleko g bihurtzen da baldin k letra 'ko' atzizkikoa bada, eta aurretik dagoena n letraz bukatutako lema edo l, m edo n letraz bukatutako toki-izena baldin bada. Esate baterako izan eta ko morfemak biltzean izango forma sortzen da.

Abiadura handitzeko gehien erabiltzen diren hitzekin zerrenda bat osatu da, berorien analisi morfologikoa ekiditzearen.

**ZUZENKETAK EGITEKO PROPOSAMENAK** Errore baten aurrean erabiltzailearen esku dago programari proposamenak eskatzea. Lan horretan errore tipografikoak eta ortografikoak tratamendu desberdina dute. Tipografikoetan karaktere baten galera, txertaketa zein aldaera edo bi karaktere jarraien arteko trukaketa hartzen da errorearen iturritzat, eta alde-erantzizkoa eginez bilatzen dira proposatzeko hitz egokiak.

Euskararen ezagumendu eskasak, batasunaren azken aldaketan berri ez izateak edo erabilpen dialektal eragindako errore ortografikoak edo tipikoak deitzen dira. Horrelakoak detektatu eta zuzendu ahal izateko Xuxenek azpilexiko eta erregela bereziak ditu. Adibidez, haundi forma lexiko berezi batean dago handi forma hobetsiarekin lotuta; 'haundiaren' formaren analisia egitean haundi+aren lortzen da, baina haundi errore bezala markaturik egotean handi bihurtzen da eta, ondorioz, handi+aren sorkuntzaz handiaren sortuko da proposamen gisa. Erregela berezien artean haren galera eta x-s-z aldaketa deskribatzen dutenak daude. Horrela zuaiteko analizatzean automatikoki zuaitez+ko lortzen da eta sorkuntzaz zuaitezko proposamena.

## Xuxen, erabiltzailearen begitik

ANTTON GURRUTXAGA / USURBIL

Euskarazko testuak zuzentzen diharduten pertsonen begiek Xuxen argitara eman dela dakitenen, lasaitu ederra hartuko dute. Hain zuzen ere, testuak denbora luzean arreta handiz irakurri beharrik sortzen duen nekeak eraginda akatsez ez oharrezko arrisku handiena dakarren lana gaizki idatzitako hitzei erreparatzeko da.

Xuxenek ez ditu sintaxi akatsak harrapatzen, hori ere egingo ahal du egunen batean, baina horrelakoak zuzentzaile trebatuari nekez itzurtzen zaizkio, atentzia ematen dutelako. Aldiz, hitz baten barruan, bereziki luze bada, letraren bat oker egotea, bi letraren arteko ordena aldatua egotea edo letraren bat falta edo sobera izatea, lanean denbora pixka bat eman ondoan oharka-

bean ontzat ematea ez da mirari.

Horiez gain, Xuxen Euskaltzaindiaren erabakiekin bat ez datozen idazkeraz eta aditz-formez ere jabetzen da (oraindik ere erabakigai diren hainbat arazotan, Xuxenen egileek aukera jakin batzuk onetsi dituzte, gaur egun onartuen diren joeren ildotik jotzeko borondatez hautatuak). Xuxenek horrelako hutsegiteen bilaketa eta zuzenketa ordenadorean automatikoki egiteko modua eskaintzen du. Euskara hizkuntza eranskaria izateak (alegia, deklinabide-atzizkia hitzari erantsirik agertzeak) sortutako arazoa konponduko zuen sistema (analizatzaile morfologikoa) sortzeko lana ikerkuntza-alorrekoa izan da, ez aplikazio hutsekoa, eta goresgarria oso.

Esandakoaz gain, Xuxenen bertute aipagarrienak erabiltzailearen hiztegi pertsonalak sortu

ahal izatea (Xuxenen hiztegi finkoaren hedadura handitzeko aukera ematen duena) eta testua aztertu ahal zuzentzeko edo alde aurretik egiaztatu eta ondoren akatsak detektaturik daudela segidan zuzentzeko aukera eskaintzea dira. Lehen prototipetatik orain merkaturatu den bertsiora abiadura bikoiztu egin da.

Informatikaren munduan, hala ere, azkartasuna aplikazioek egunetik egunera hobetzen duten alderdia da, Xuxenen etorkizuneko bertsioek areago landu beharko dutena. Euskal Herriaren egindako lehen CD-ROMaren ondoren (Elhuyarren 'Ezagutu Gipuzkoa'), Xuxen zuzentzaile ortografikoa dugu kalean eta euskara hizkuntza industriaren munduan barneratzeko behar den teknologia garatzen ari dela ikusita poza hartu behar luke euskaldunak.