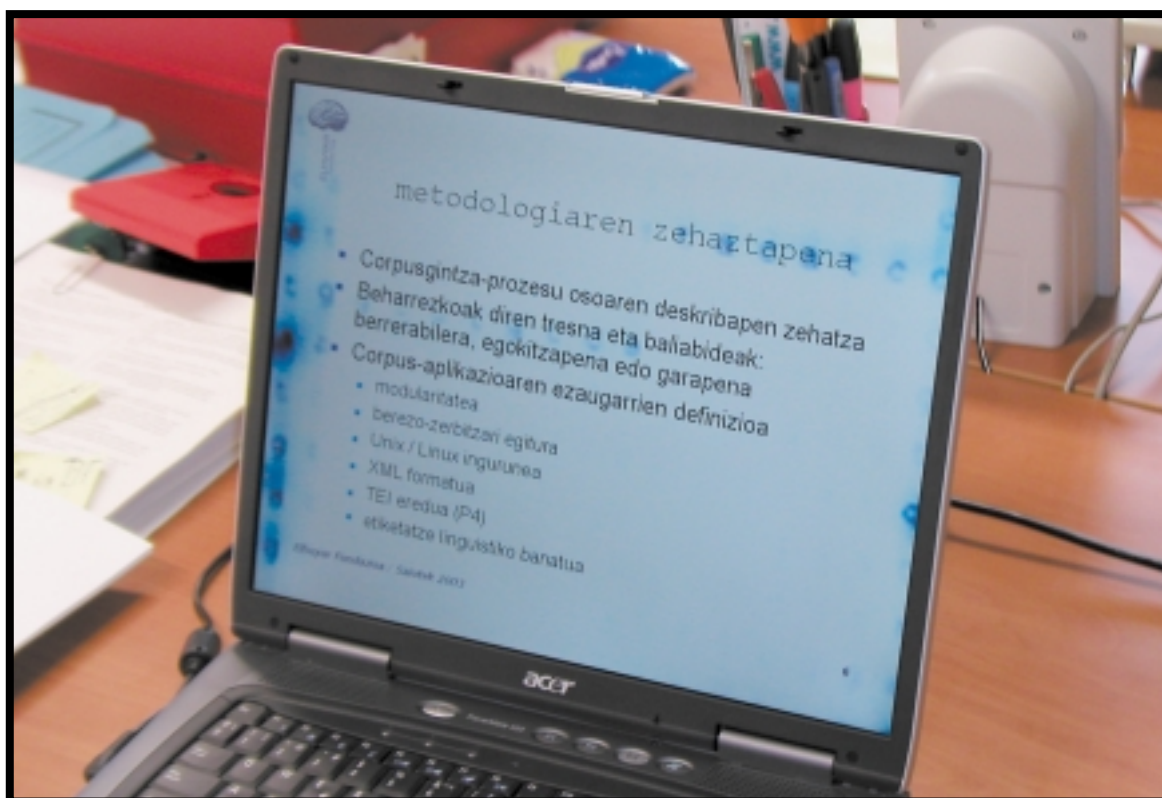


Hizking21

hizkuntz ingeniari-tza XXI. mendearen atean

Rafa Saiz Elizondo
Elhuyar



Testuen prozesutik haratago, konputagailuek hizkuntzaren alorrean egin dezaketena jorratzen du hizkuntz ingeniari-tzak. Makinekin komunikatzeko bidea gero eta naturalagoa izatea da, besteak beste, helburu orokorra: erabiltzailea ohi duen bezala mintzatuko da, edo idatziko du, eta makinek ulertu egingo diote, eta, hala programatuz gero, obeditu ere bai. Lehendik grabatutako mezuak errepikatu beharrean, mezu komunikatiboak sortuko dituzte makinek, idatzizkoak zein hitzezkoak. Horretarako, ordea, bide luzea eta oinarritzko lan handia egin behar da.

Hizking21 izenaz ekin zaion proiektuaren helburua hau da: 2005 urterako, gaur egun ingeleserako dau den tresna infolinguistikoak euskaraz ere izatea. Horietako asko euskararako propio sortuak izango dira, beste batzuk beste hizkuntzetarako daudenak egokituz lortuak. Euskararen morfologia eta sintaxi berezia direla eta, lehenago inork jorratu gabeko arazoak topatuko dira, eta horiek gainditzeko garatzen den teknologia Euskal Herria mundu mailako erreferentzia bilaka lezake alor honetan.

● Makinari 'entzun' ● duena nola ezagutu eta nola 'idatzi' behar duen irakatsi behar zaio.

Zer dagoen egun

Gaur egun euskarazko hizkuntz teknologia informatikoak aipatzekotan, ezinbestekoa da Euskal Herriko Unibertsitateko IXA eta Aholab taldeak aipatzea. Hizkuntzaren tratamendurako hainbat tresna informatiko garatu dituzte: zuzentzaile ortografikoa, lematizatzailea, desanbiguatzailea, etab.; eta horiek izango dira, hein handi batean, proiektuaren abiapuntua. Horiek lan egingo badute, ordea, erreferentziak behar dituzte, lexikoak; eta horiek osatu eta eguneratzeko berriz, ezinbesteko tresna bilakatu dira corpusak, hizkuntza errealaren isla diren testu-biltegi sailkatu, etiketatu eta antolatutak.

Proiektuaren xehetasunak

Hizking21 proiektuak 7.600.000 -ko aurrekontua du. Eusko Jaurlaritzaren Industria, Merkataritza eta Turismo Sailak Infoingeniaritza Linguistikoa interes estrategikoko ikerketa-lerro izendatu zuen, eta Etorrek programaren bidez lagundu dio.

Hizking21en bost partzuer bildu dira: Elhuyar Fundazioa, Euskal Herriko Unibertsitateko IXA eta Aholab taldeak, Vicomtech elkarte eta Robotiker Fundazioa. Laguntzaile moduan, Eleka S.L. enpresak ere badihardu proiektu honetan, IXAren eta Ehuyarren artean sortua. Zein bere aldetik, hizkuntz gaitasuna duten sistematik diseinatzeko behar den ezagutzaren eta gaitasunaren jabe dira. Partzuer guztien lanari esker, posible izango da hemendik gutxira hainbat tresna informatiko eskura izatea eguneroko aplikazioetan txertatzeko moduan.

Elhuyarrek urteetan zehar hiztegitan egindako lanak, bai hizkuntz hiztegitan bai hiztegi teknikoetan, baliabide aproposa izango dira aipatu lexikoi horiek osatu eta elikatzeke. Zientzia eta Teknikaren alorrean osatu eta bildu duen materiala ere baliotsua izango da oso corpus bereziak sortze bidean.

Aholabek ahotsaren analisisa eta tratamendua du bere jardun nagusia. Ahots-formatuko fitxategi batetik testu idatzira igarotzeko tresnak baditu, eta alderantzizkoak ere bai. Horretarako ere ezinbestekoa da erreferentziak izatea: makinari 'entzun' duena nola ezagutu eta nola 'idatzi' behar duen irakatsi behar zaio.

Zenbat eta gehiago garatu hizkuntza naturalaren prozesamendua, orduan eta aiseago erabili ahal izango dira lan-tresna informatikoak.



G. ANDONEGI

Horra bitarteko tresna eta baliabideak euskarari estu lotutakoak dira. Beraz, gehienak propio garatutako teknologiaz egindako programak dira. Interfazeei dagokienez, berriz, beste hizkuntzetan dagoeneko

- **Euskararen morfologia**
- **eta sintaxi berezia**
- direla eta, lehenago**
- inork jorratu gabeko**
- arazoak topatuko dira.**

aurreratutako lanaren zati bat baliagarria den arren —avatarren diseinua oso aurreratuta dago—, euskaraz hitz eginarazi behar zaie. Horretan ere Aholabek eta Vicomtehek beren bidea egina dute eta aurrera segituko dute.



G. ANDONEGI

Makinekiko komunikazioa ahalik eta modu naturalenean gauzatu nahi da.

Robotiker Fundazioa erreferentzia da ekipoen arteko konektagarritasunaren arloan Euskal Herrian, eta oinarrizko teknologiaren ardura izango du Hizking21en. ➔