

# Hizkuntzaren Teknologiarako adituak prestatzen

Kepa Sarasola

Ixa kidea eta UEU kidea

*Unibertsitate mailako euskarazko masterrak ez dira asko. Hamaika ahalegin eta ilusioaren fruituak dira. Zenbat irakasle eta ikasle, hor, lanean elkarrekin! Euskal Herriko unibertsitateetako curriculum ofizialak betetzen ez dituzten zuloak betetzera datoz, proiektu berriak zabalduz: euskarazko kazetaritza, hizkuntzaren teknologiak, herri-erakundeen kudeaketa, soziolinguistika. HIZTEK masterrak giza talde baten ilusioa biltzen du, eta hirugarren aldiz bere ateak zabaltzen ditu. Ikus dezagun non kokatzen den eta zer sortu nahi duen.*

Hizkuntza-Teknologiak funtsezkoak dira Informazio eta Komunikazioaren Gizartea esaten dugun horretan. Epe ertainean pertsona eta makinaren arteko komunikazioa, hainbat aplikazio informatikotan, geure hizkuntzan egin ahal izango dugu, ez makinaren hizkuntzan. Tresna mugatuak izango dira, eta beti errore-maila batekin, baina, hala ere, laguntza ederra emango digute. Gaur egun eskura ditugu horrelako zenbait hizkuntza-aplikazio, halanola: ortografia-zuzentzaileak eta estilo-zuzentzaileak, hiztegi-kontsultak on-line, itzulpen-laguntzak, Internetarako bilatzaileak, hizketa testua bihurtzen duten sistemak, testua irakurtzen dutenak edo bigarren hizkuntza ikasteko sistemak. Baina horrelako sistema gehienak ingeleserako balio dute, ez beste hizkuntzetarako. Beste hizkuntzetako hitzunok ahalegin handia egin behar dugu atzean ez gelditzeko, are gehiago euskara bezalako hizkuntza txikikoek.

## **Arazoa onura bihurtuta**

Betiko leloa da hori; ematen du beti ibili behar garela negarrez gure txikitasunarengatik eta gure ezintasunarengatik. Baina ezetz! Gure alde ere joka dezake egoera horrek. Nola? Ikus dezagun bada. Argi dago hizkuntzaren industria honetan ingeleserako produktuen merkatua oso handia dela, baina alor honetan gabiltzan zenbait ikertalde eta enpresak uste dugu produktu horiek ez direla zabaldu modu egokian beste hizkuntzetarako, badagoela espazio eta zeregin, ikerketan eta produktuen mailan ere bai, gure ekarpenak bideratu ahal izateko. Guk euskaldunok, eta oro har europarrok ere, ohituta gaude eleaniztasunean bizi izaten, eta horren ondorioz, gaitasun handiagoa lortu

dugu eleaniztasuna lantzeko. Gure kasuan, gainera, euskara oso bestelako ezaugarriak dituenetz, probaleku ezin hobea dugu geure elea hizkuntza-produktu informatikoen moldagarritasuna frogatu ahal izateko. Beraz, alde batetik hainbat ahalegin, gastu eta buruhauste ekartzen digun eleaniztasunak, beste alde batetik, arazo horiei aurre egiten ikasi eta gero, kokapen oso aurreratuan jartzen gaitu nazioarteko mailan ekarpen esanguratsuak egiteko. Beraz, hemen sor dezakegun Hizkuntza-Teknologia esportagarria eta nazioartekora ateratzeko modukoa izango dela uste dugu.

### **Euskararen egoera hizkuntza teknologian orain**

Aztertzen badugu Euskararen Softwarearen Katalogoa (<http://www.ueu.org/softkat>) hizkuntzaren prozesamendurekin lotuta dauden 28 aplikazioak aurkituko ditugu honela sailkatuta: ediziorako laguntzak (Xuxen eta Elhuyar-Word), hizketaren tratamendua (Bizkaieraren Fonoteka eta AhoTTS Testu-Ahots Bihurgailua), Euskara ikasteko metodoak (Bai & Bye, BOGA eta HEZINET), Lematizatzailea eta informazioa bilatzeko tresna (Euslem), datu-base dokumentala (Kapsula), corpus (XX. mendekoa), eta 16 baliabide lexikal (hiztegiak, esamoldeak, ...). Esan dezakegu, beraz, ez gaudela basamortuan, hori ez dela hutsaren hurrengoa, baina bai oso gutxi gaztelaniarako, frantseserako edo, batez ere, ingeleserako eskuragarri dauden ehunka programa eta baliabiderek.

### **Eginkizuna**

IXA taldearen 15 urteko ibilbidean urrats-kate bat definitu dugu hizkuntzaren teknologiarri estrategia baten arabera ekiteko. Ideia nagusia oso sinplea da: *Hasieran oinarritzko baliabideak eta tresnak sortu behar dira, eta geroago sortu merkatu-aplikazioak. Alderantziz ez dela egin behar!* Produktu posibleen artean bereiztu izan ditugu zein diren hizkuntza-baliabide, zein tresna, eta zein merkatuko aplikazioa. Eta horren arabera baliabide, tresna edo aplikazio bakoitza noiz egin behar den aurreikusi izan dugu, ekoizpen prozesu hori optimizatu nahian. Produktu bakoitza produktu berrien garapenean ahalik eta modu zabalenean berrerabilia izatea da gure helburua.

Eskuragarri diren baliabide, tresna eta aplikazioen katalogo integratua behar da. Geure hizkuntzarentzat eta antzeko hizkuntzentzat. Talde desberdinetan lan bera errepikatzea dirua eta lana xahutzea da. Ezinbestekoa lanen koordinazioa. Emaitzak

publikoak izan beharko lirateke, beste ikertzaileek erabiltzeko moduan utzi beharko lirateke.

Euskararako baliabide linguistikoak sortu beharko dira: testu-bilketa (corpus orokorra, berezituak eta eleanizdunak, ahal bada 100 milioi hitzekoa). Hiztegi elektronikoak eta hitzen sailkapen orokorra adieraren arabera.

Aplikazio berriak plazaratu beharko dira: informazio-bilatzaileak (eguraldi-partreak, burtsa, kirolak, berriak, bideo-eskaerak, irudi-eskaerak, ...), domotika, itzulpen-laguntzak. irakaskuntza-sistemak (e-learning), elbarrientzako laguntzak, telebista digitala, elkarrizketa-sistemak, multimedia-sistemak, ...

Ingeniaritza linguistikoan I+G horretan (Ikerketan eta Garapenean) arituko den komunitate zabal bat sortu behar dugu. Gaur, 2003. urtean, estimatzen dugu 120-150 pertsona gabiltza lanean Euskal Herrian hizkuntzaren prozesamenduaren inguruan, zuzenki edo zeharka. HIZKING21 proiektu estrategikoaren barruan, gauzak ondo planifikatuz gero eta formazio-plan egokiak aurrera eramanez gero (masterrak eta doktorego-programak), urte gutxi barru zenbaki hori bikoiztu edo hirukoiztu egin daitekeela aurreikusten dugu. Komunitate horrek, noski, konpartitu beharko ditu metodologiak, teknologia, algoritmoak eta programak.

HIZKING21 proiektuan inguruan bildu garenok (Elhuyar fundazioa, VICOMTech, Robotiker eta EHUko Aholab eta IXA taldeak) etorkizun horretan sinisten dugunez, konpromisoa hartu dugu eragile-lanak egiteko gure gizartean. Gure lana esparru hauetan egingo dugu: formazioa, nazioarteko lankidetzak, azpiegiturak, oinarritzko baliabideak eta behatoki teknologikoa.

Formazio-plan horretan kokatzen da UEUren papera. Udako Euskal Unibertsitateak heziketa-behar hori bete ahal izateko, hainbat unibertsitate eta enpresatako adituak bilduz, 2001/2002 ikasturtean HIZKUNTZA-TEKNOLOGIAK masterra jarri zuen martxan. Ikasketa berri honek UPV/EHUren oniritzia jaso du, eta beraz, unibertsitateak bere titulu propio hau bere babesarekin zabaldu du. Udako Euskal Unibertsitatea eta EHUko Lengoaia eta Sistema Informatikoak sailaren arteko lankidetzak fruitu onak eman ditu. Orduan hasi ziren hogeitaz hamar ikasleetatik bederatzi ari dira gaur egun lanean Hizkuntzaren Teknologiaren alorrean.

Nazioartekoan puntako mailan mugituko den industria sendoa sortu dezakegu. Ikertalde, enpresa eta erakunde ofizialetatik ahaleginak egiten ari gara asmo horiek gauzatzeko.

Aldi berean Hizkuntzaren Teknologiarako adituak prestatu behar dira, horixe lortu nahi du UEUren HIZTEK masterrak. Informatikariak, ingeniariak, hizkuntzalari eta filologoak dira harrobi horretarako lehengaia.