

EUSKARAZKO ERROREEN ETA DESBIDERATZEEN ANALISIRAKO LAN-INGURUNEA. DETERMINATZAILE- ERROREEN AZTERKETA ETA PROZESAMENDUA

Tesiaren egilea: Larraitz Uriá Garin

Unibertsitatea: Euskal Herriko Unibertsitatea

Saila: Euskal Filologia

Tesi-zuzendaria: Igone Zabala Unzalu eta Montse Maritxalar Anglada

Tesiaren laburpena:

Erroreen Analisia psikolinguistikaren alorrean sortu zen 70eko hamarkadan. Ordenagailuaren erabilerak azken urteotan izan duen hedaduraren ondorioz, aukera berriak zabaldu dira, eta egun, jada, beste hainbat arlotan egiten diren ikerketen oinarri bilakatu da.

Guk euskarazko erroreak eta desbideratzeak aztertzen ditugu bi ikerketa-alorretan ekarpenak egiteko: erroreen tratamendu automatikoaren alorrean, euskararako ortografia-, gramatika- eta estilo-zuzentzailea edota aldaki dialektalen markatzailea bezalako tresnak garatzea dugu helburu; ordenagailuz lagundutako hizkuntzen i(ra)kaskuntzaren alorrean, berriz, hizkuntzaren i(ra)kaskuntza-prozesua bideratzen laguntzeko baliabideak prestatzea da gure xedeetako bat. Alor batean zein bestean, erabiltzaileen beharrei erantzuteko gai diren laguntza-tresna eragin-korrek garatu ahal izateko, ezinbestekoa da euskaraz egiten diren erroreen eta desbideratzeen analisi sakona egitea, gure hizkuntza-komunitateak bizi duen egoera soziolinguistikoa kontuan hartuta.

Testuinguru honetan, erroreak eta desbideratzeak bereizten ditugu. Erroretzat hartzen ditugu arau linguistikoak betetzen ez dituzten egiturak, edozein kasutan gaizki egongo direnak. Desbideratze esaten diegu, berriz, gramatikaren ikuspuntutik zuzenak izan arren, testuinguru edota egoera jakin batean egokiak ez diren egiturei: behin eta berriz errepikatzen diren hitzei/egiturei edota testuinguru komunikatibo jakin baterako desegokiak diren baliabideei (adibidez, baliabide dialektalak erabiltzea euskara batua erabili behar den testuinguruetan, edota alderantziz).

Tesi-lan honetan, euskarazko erroreen azterketa eta prozesamendua bideratzeko hainbat irizpide finkatu ditugu eta lan-ingurune oso bat eratu dugu erroreen eta desbideratzeen analisirako ezinbestekoak diren baliabideekin: corpusak (hizkuntza-ikasleen edota hiztun arrunten testu-bildumak), adibide erroredunak corpusetan etiketatzeko *EtikErro* deritzon editorea, erroreen eta desbideratzeen sailkapen bat eta etiketatutako adibideak gordetzeko bi datu-base: *erroreak* eta *ikasleak*. *Erroreak* datu-basean erroreen eta desbideratzeen adibideak biltegiratzeko ditugu, adibide bakoitza automatikoki tratatzeko beharrezkoa den informazio linguistikoarekin eta teknikoarekin batera; *Ikasleak* datu-basean, berriz, hizkuntza-

ikasleek egindako erroreen eta desbideratzeen adibideak gordetzen ditugu, adibideei dagokien informazio linguistikoarekin eta psikolinguistikoarekin batera. Informazio hori guztia da helburu ditugun tresna informatikoak garatzeko oinarria.

Errore mota guztien analisia aldi berean egitea ezinezkoa zaigunez, tesi-lan honetan determinatzaile-erroreak izan ditugu aztergai eta determinatzaile-errore batzuk automatikoki detektatzeko erregelak egin ditugu. Erregela horiek euskararako garatzen ari garen *XUXENg* gramatika-zuzentzailean integratuko dira.