

# Hizkuntzalari Euskaldunen I. Topaketa

## Eibarren, 2013ko uztailaren 4an

### Aurkibidea

- Hitzaldi gonbidatua: "Igarri, ikertu, jakin?" Iñaki Segurola
- Komunikazioak:
  1. *Irakaslegaiak ahozko gaitasunetan trebatzea*. Amaia Alvarez eta Edu Zelaieta
  2. *Ahozkotasanaren printzak Koldo Mitxelenaren idatzietan*. Agurtzane Azpeitia eta Elixabete Perez
  3. *Ahozko tradiziozko corpuserako lehenengo urratsak*. Asier Sarasua
  4. *Goi-nafarrera arkaiko eta zaharreko testuen edizio eta azterketa berrirantz*. Urtzi Reguero
  5. *Osagarrien markapen bereizgarria euskal hizkeretan*. Ane Odria
  6. *Ergatibotasuna, iragangaiztasuna eta aditz laguntzailearen hautaketa euskal hizkeretan*. Ane Berro
  7. *Informazio egitura eta kuantifikatzaileen interpretazioa*. Urtzi Etxeberria eta Aritz Irurtzun.
  8. *Euskal alokutibora joan-etorria*. Aitor Lizardi eta Manex Munduate
  9. *Hizkuntza-aldakortasuna Nafarroa ipar-mendebaldeko euskaran*. Oihana Lujanbio
  10. *Hizkuntz aldakortasuna: Euskara Baturanzko joera elkarrizketen bidez azterturik*. Orreaga Ibarra
  11. *Euskarazko testuen sinplifikazio automatikoa*. Itziar Gonzalez Dios
  12. *EPEC corpusa predikatu mailan etiketatzeko eta ustiatzeko oinarriak: BVI, EPEC-RoISem eta e-ROLDa*. Ainara Estarrona.
  13. *Ezagutza linguistikoa Ingelesa-Euskara itzulpen automatikoko sistemen garapenean*. Nora Aranberri
  14. *BASYQUE: hizkeren arteko aldakortasun sintaktikoa aztertzeko aplikazioa*. Gotzon Santander, Xan Aire, Larraitz Uria eta Ricardo Etxepare
  15. *Gazteak lagun artean idazten*. Elixabete Perez eta Esther Zulaika
  16. *Aldakortasun sintaktikoa euskarazko perpaus nominalizatuetan*. Maia Duguine

## Hitzaldia

### "Igarri, ikertu, jakin"

Iñaki Segurola, idazlea eta hiztegiilea.

Hizlariak aitortzen du ez dakiela zer ikertu eta zer jakin behar den. Eskuarki hori “goitik” erabakitzen dela.

Aurea eta atzea kontrajartzen: aurretik ez liguke inork agindu behar zer ikertu behar dugun. Atzetik harrapatu behar gintuzke ikertzeko, igartzeko gogo batek.

Jakitea eta igartze/antzemate/hautematea kontrajartzen dira. “Irizpide objetiboak” ere zalantzan jartzen, eta doktoletzako tesien kontra egiten.

Planck-en fedea eta Einstein-en *Einfühlung* (intropatia, barrensumatzea); azken hitz honek ikertzea eta igartzea lotzen laguntzen du. Gainera, hizlariaren ustez, *ikertu* eta *igerri/igarri*-k jatorri etimologiko bera lukete euskaraz.

Zer ikertu behar den jakin ezean, zer ez den ikertu behar: “corpus-mania”, euskararen jatorria, euskararen iraupena edo belarrimotz orojakileen laineza aitutzen adibide gisa.

Nolanahi ere, hizlariak baditu bere apetatxoak, eta halaxe botatzen ditu, apeta pertsonal moduan: bere profezia soziolinguistiko baten zuztarrak, euskara batuak hartu dituen bide okerrak, eta, bukatzeko, “erdara ere hemengoa da” esaldiaren gorabeherak.

1. Hizkuntzaren didaktika Gasteizko Irakasle Unibertsitate Eskolan duen erronka nagusi bat. Irakaslegaiak ahozko gaitasunetan trebatzea.

Amaia Alvarez eta Edu Zelaieta (UPV/EHU)

amaia.au@ehu.es edu.zelaieta@ehu.es

Komunikazioaren laburpena

Idatzizko gaitasunak jaun eta jabe gertatu dira tradizio akademikoan berriki arte. Erabateko nagusitasun horren ondorioz, ahozko hizkuntza eta haren ikas-irakaskuntza bigarren maila batean egon dira eskoletan zein unibertsitateetan. Azkenaldian, ordea, ikasgelako ahozko gaitasunak aztergai hartu dituzten hainbat ikerlan agertu dira, eta hizkuntzaren didaktikaren inguruko ikuspegi berri bat ekarri dute. Marko gaurkotu horrekin bat dator aurkeztu nahi dugun lana: Haur Hezkuntzako irakaslegaien ahozko gaitasunak neurtzeko eta hobetzeko ikertzen ari gara, geroko maisu-maistrak trebeagoak izan daitezen euren jardun profesionalean: hurrekin gelan, lankideekin bileretan eta gurasoekin elkarriketetan.

Hizkuntza eta hezkuntza uztartzen dituen ikerketa egitasmo hau gazte samarra bada ere, lehen argitalpen akademiko-zientifikoak emanak ditu dagoeneko. Hori dela eta, proiektuaren orain artekoaz zein hemendik aurrerakoez ere jardungo dugu.

## 2. Ahozketasunaren printzak Koldo Mitxelenaren idatzietan.

Agurtzane Azpeitia (Mondragon Unibertsitatea) eta Elixabete Perez (Deustuko Unibertsitatea)  
aazpeitia@mondragon.edu elixabete.perez@deusto.es

Komunikazio honen hausnargaia IKTek berpiztu duten ikerresparru zabal batean kokatzekoa da: bitarteko edo kanalen araberako bi kodeen —ahozko adierazpidearen eta idatziaren— arteko harremanak (besteak beste, erregistro-aukera, bat-batekotasuna...), eta zehazkiago, “ahozko” moldeek bitarteko idatzian duten isla.

Horretarako, euskal prosa modernoaren zutabe garrantzitsua den Koldo Mitxelenaren diskurtso idatziko testu hautatu batzuetan ahozketasunak duen presentziaren hari-muturretako bat joko dugu begiz.

Agerian jartzen ahaleginduko gara haren prosa zipriztintzen duten tartekiak, enuntziatu parentetikoak (Azpeitia 2012), baliabide indartsuak gertatzen zaizkiola Errenteriako idazleari ahozketasunaren printzak idatzian txertatzeko, irakurlearekin “solasteko”. Xede hori erdietsi nahian, ahozketasunaren ezaugarriak mugatu, definitu eta Mitxelenaren euskal testuotan itxurazko “hizketaldi” horiek utzitako lorratzak atzeman eta aztertuko ditugu.

Uste dugu gure proposamena gaurkotasun handiko bi arlotan lagungarri gerta daitekeela: Mitxelenaren prosaren oinarri eta osagaiak sakonago definitu eta ezagutzeko, eta euskal hiztunaren erabileraren, erregistroaren, kodeen arteko harremanen (ahozketasun eta idazketaren) ezaugarriak ikertzeko.

Azkenik, azpimarratu nahi dugu ikerketa honek topaketon helburuetako bat bete-betean betetzen duela, bi ikertzaileok euskal unibertsitate banakoak izan eta auzolanean ikertzearen ondorioa baita.

### 3. Ahozko Tradiziozko Corpuserako lehenengo urratsak.

Asier Sarasua (Badihardugu Euskara Elkarte)

asier@ahotsak.com

Komunikazio honen helburuak:

- Ahozko jardun librean oinarritutako corpus baten ezaugarriak (eta zailtasunak) azaltzea.
- Ahozko Tradizioaren Corpusaren bereizgarriak zedarritzea.
- Lematizazio prozesu semi-automatikoa azaltzea eta bidean aurkitutako oztopoen berri ematea.
- Lehenengo emaitzak aurkeztea, alde on eta txarrekin.

Laburpena:

Euskaraz hainbat corpus ditugu erabilgarri eta azkenaldian berriak ere azaldu dira (azkena, ETC, Egungo Testuen Corpus erraldoia). Hala ere, corpus gehienek izaerak nahiko antzekoak dira: gehienbat idatzizko dokumentuak dituzte oinarri eta hizkuntza erregistroa formala da, nagusiki.

Euskal Herriko Ahotsak proiektuan 5.000 ordu ditugu grabatuta. Grabazio horietan euskaldun zaharren jardun libre dago jasota. 1900-1940 bitartean jaiotako hiztunak, gehienak alfabetatu gabeak eta euren bizitza euskaraz bizi izan dutenak. Beti pentsatu izan dugu grabazio horietan jasotako eduki linguistikoak garrantzi berezia izan dezakeela beste Corpus batzuk osatzeko, beste inon jaso ez den ondarea biltzen duelako ahozko tradizioaren corpus honek. Hau da, Euskal Herriko Ahotsak proiektuan zehar egindako transkripzioetan oinarrituz, XX. mende hasieran jaiotako euskaldunen bat-bateko hizkera naturalaren ezaugarriak bilduko dituen Corpora izango genuke.

Ahotsak proiektuaren helburua bilketa eta katalogazioa da, eta hasieratik ikusi dugu gure gaitasunetik gora dagoela Corpus erraldoi hori behar bezala landu eta lematizatzea. Corpusaren jatorria ezagututa (ahozko ekoizpena, estandarretik urruti, hiztun nagusien ahotan, ahozkoaren aje guztiekin...), jakin bagenekien horrelako lan batek zailtasun handiak izango zituela, gaur egun ez baitago ia-ia lan automatikorik egiterik. Hala ere, emaitza batzuk lortze aldera, 2012. urtean Ahozko Tradiziozko Corpusaren proiektua abiatu genuen. Sistema semi-automatiko bat erabilita, lehenengo urratsak eman ditugu eta hasierako emaitzak ere lortu ditugu.

Saio esperimental baten emaitza gisa ulertu eta erabili behar da guztia, noski, baina lehenengo emaitzak lortu ditugu jada:

- 1) lema bakoitzaren aldaerak;
- 2) aldaeren eta lemen banaketa geografikoa;
- 3) aldaeren eta agerpen ezberdin guztien datu estatistikoak;
- 4) erabilera-adibideak. Kasu guztietan, bildutakoaren transkripzio-zatia eta bideoa kontsultatzeko aukera ere eskaintzen baita.

#### 4. Goi-nafarrera arkaiko eta zaharreko testuen edizio eta azterketa berrirantz.

Urtzi Reguero (UPV/EHU)

urtzi.reguero@ehu.es

Aurkezten dugun komunikazio proposamen honetan gure doktorego tesi izango denaren proiektua zertan den azalduko dugu. Oraindik asmoa besterik ez den lanari "Goi-nafarrera arkaiko eta zaharra: ikerketa eta testuak" izenburua eman diogu.

Euskalkien azterketa diakronikoak egitea ezinbesteko beharra da euskararen historiaren nondik norakoak ezagutzeko, eta goi-nafarrera arkaiko eta zaharrari dagokionez, ez da lan diakroniko orokorrik egin orain artean, alegia, ez da testu guztiak aztertu eta goi-nafarrera arkaiko eta zaharra osoki deskribatu duen ikerketarik. Beraz, ikusten dugun hutsunea betetzeko, gure lanak helburu argi eta zehatzak ditu; honatx:

- a) XV. mendetik hasi eta XVIII. mendearen erdialdera bitarteko goi-nafarrerazko testu ezagun guztien edizio filologikoak egitea, hizkera aztertzen hasi aurretik hauek finkatuta izateko asmoz. Hauen artean Beriain eta Elizalderen liburuak ditugu, baina baita laburragoak diren gutun, olerki eta bestelako testuak ere.
- b) Testu hauen azterketa egitea hizkuntzazko ezaugarri nagusiak kontuan izanda: grafia, fonologia, izen eta aditz morfologia, sintaxia eta lexikoa aztertuko genituzke. Dialektologia diakronikoan ohi denez, kontuan izango ditugu ezaugarriok arkaismoak, hautuak ala berrikuntzak diren.
- c) Goi-nafarreraren nondik norakoak ezagutu eta orduko dialektoekin alderatuz euskalki zaharren artean kokatzea.

Goi-nafarrera aztertzen hasterako, ezer baino lehen kontuan hartu behar genukeen eragileetako bat da Nafarroa euskararen eremuaren erdigunean dagoela, eta hain zuzen ere horregatik inguruan dialekto eta hizkera gehiago dituela. Nahiz eta denboraren poderioan atzerakada handien jasan duen euskal dialekto historikoa izan, erdigunean egoteak berritzaile izatea ekarri bide dio goi-nafarrerari. Uste dugu euskararen gainerako dialektoekin komunean dituen ezaugarriak zein goi-nafarrera berezi egiten dituenak diakronikoki aztertzeak eta hauen kronologia egiteak euskararen dialektizazioaren eta historiaren ezagutzari ekarpen garrantzitsua ekar diezaiokeela.

## 5. Osagarrien markapen bereizgarria euskal hizkeretan.

Ane Odria (UPV/EHU)

ane.odria@ehu.es

Nire doktore tesiaren xede nagusia euskal hizkera batzuetan gertatzen den Osagarrien Markapen Bereizgarria (Differential Object Marking, DOM) aztertzea da. Egitura iragankorretako zenbait osagarrik kanonikoki dagokien absolutiboaren orde (1), datibo markapena jasotzen dute hego-mendebaleko euskal hizkera batzuetan (2) (Fernández 2008, Mounole 2008, Fernández & Rezac 2010, Odria 2012), eta markapen hori munduko hainbat hizkuntzatan gertatzen den DOM fenomenoarekin lotuta dago (Aissen 2003). Datibo kasu eta komunztadura daramaten osagarri ez-estandar horien izaera sintaktikoa argitzea da, hain zuzen ere, nire tesiaren egiteko nagusia.

Euskararen DOM osagarriak gaztelaniaren DOMarekin –bai a markapenarekin (Torrego 1998) (3) eta bai leismoarekin (4) (Fernández-Ordoñez 1999, Ormazabal & Romero 2012)– ukipenean dagoen eremu geografikoan baino ez dira gertatzen –ez dago datiboz markatutako osagarririk ipar-ekialdeko hizkeretan–, eta horrek bi hizkuntzen DOM fenomenoak elkarreraginean daudela iradokitzen du. Horregatik, doktore tesi hau funtsean euskal hizkeretan oinarritzen bada ere, gaztelaniaren datuak ere aztertzen ari naiz, fenomeno beraren adierazle diren bi hizkuntzen arteko konparaketa egiteak bataren zein bestearen jokabide sintaktikoa zein den argitzen lagundu dezakeelakoan.

Euskal hizkeren arteko aldakortasuna eragiten duen DOMa da, beraz, tesi honen ikergai nagusia, eta fenomenoaren izaera deskribatzeaz gain, berau azalduko duen analisi teoriko bat proposatzea du helburu lan honek. Komunikazio honetan, DOM osagarrien deskribapenetik abiatuta, garatzen ari naizen tesiaren hipotesi teoriko nagusiak aurkeztea da nire asmoa.

### Adibideak

- Nik zu-ø ikusi z-aitut.
- Nik zu-ri ikusi di-zu-t.
- He visto a Juan.
- Le he visto a Juan.

### Erreferentzia bibliografikoak

- Aissen, Judith. 2003. "Differential object marking: iconicity Vs. economy." *Natural Language and Linguistic Theory* 21. 435-485.
- Fernández, Beatriz. 2008. "Quirky dative objects in Basque". *European Dialect Syntac Meeting III*, Venezia, irailak 18-21.
- Fernández, Beatriz & Milan Rezac. 2010. "Datibo osagarri bitxiak eta Datiboaren Lekualdatzea: ari nai diyot eta kanta egin nazu bidegurutzean." Beatriz Fernández, Pablo Albizu & Ricardo Etxepare

(arg.), Euskara eta euskarak: aldakortasun sintaktikoa aztergai. ASJUren gehigarriak LII. Bilbo: UPV/EHU. 113-149.

- Fernández-Ordoñez, Inés. 1999. "Leísmo, laísmo y loísmo". Ignacio Bosque & Violeta Demonte (arg.), Gramática descriptiva de la lengua española A. Madrid: Espasa Calpe. 1317-1395.
- Mounole, Celine. 2008. "Evolution of the transitive verbs in Basque and apparition of datively marked patients." G. Authier & K. Haude (arg.), Ergativity, Transitivity, and Voice. Berlin / New York: Mouton de Gruyter.
- Odria, Ane. 2012. What lies behind differential Object Marking: a Basque dialect survey. Master Tesia, UPV/EHU. Pello Salaburu, Patxi Goenaga & Ibon Sarasola (arg.), Sareko Euskal Gramatika [Interneten eskuragarri: [http://www.ehu.es/seg/\\_media/morf/5/2/1/9/sak/aneodriatesia.pdf](http://www.ehu.es/seg/_media/morf/5/2/1/9/sak/aneodriatesia.pdf)]
- Torrego, Esther. 1998. The Dependencies of Objects. Cambridge: MIT Press.
- Ormazabal, Javier & Juan Romero. 2012. "Object Clitics, Agreement and Dialectal Variation". Eskuizkribua, UPV-EHU.



## 6. Ergatibotasuna, iragangaiztasuna eta aditz laguntzailearen hautaketa euskal hizkeretan.

Ane Berro (UPV/EHU)

ane.berro@ehu.es

Nire doktorego tesiaren xede nagusia da subjektuaren ergatibo eta absolutibo kasuen arteko aldakortasuna eta aditz laguntzaileen hautaketa aztertzea euskaraz. Bi alderdi aldakor hauek bereziki euskal hizkeren barruan aztertzeko xedea dut.

Subjektuak erakusten duen kasu aldakortasuna aditz laguntzaileen hautaketarekin (izan/\*edin edo \*edun/\*ezan) bat dator eta normalean predikatuaren balentziak baldintzatzen du (1). Alabaina, predikatu eta egitura jakin batzuek aldakortasuna erakusten dute hizkera batetik bestera (2). Tesiaren egiteko nagusia egitura hauek ikertzea eta teoria bakarraren barruan azaltzea da.

Xede orokor hau gogoan dugularik, tesiak, zehatzago, ondorengo helburuak ditu:

- i. Zenbait eremutan ergatibo eta absolutibo kasuen artean antzematen den aldakortasuna eta aditz laguntzailearen hautaketa azaldu.
- ii. Kasuak eta aditz laguntzailearen hautaketak gertakari-egiturarekin eta aspektu lexikoarekin duten lotura aztertu. Helburu hau betetzeko, lexikoiaeren eta sintaxiaren zereginen arteko banaketa zehaztu dezakeen teoria bat proposatu.
- iii. Predikatu iragangaitzen artean sailkatu izan diren predikatu inakusatibo eta inergatiboek subjektuari dagokionez erakusten duten kasu aldakortasuna azaldu.
- iv. Giza hizkuntzaren barruko aldakortasunaren eremuaz eta bere mugez hausnartu, euskalki eta euskal hizkeren arteko aldakortasunari eta munduko hizkuntzen artekoari begiratu.

Hitzaldi honetan, aldakortasuna baldintzatzen duten aipatutako irizpideez arituko naiz, eta nire tesian proposatuko dudan teoriaren lehen urratsak azalduko ditut.

(1)

- a. Jonek Ane ikusi du
- b. Jone etorri da
- c. Langabezia igo da azken lau urteetan

(2)

- a. Jonek borrokatu du
- b. Jonek borrokatu da
- c. Kremlera honek ez du ondo igotzen
- d. Kremlera hau ez da ondo igotzen

## 7. Informazio egitura eta kuantifikatzaileen interpretazioa.

Urtzi Etxeberria eta Aritz Irurtzun (CNRS-IKER)

u.etxeberrria@iker.cnrs.fr

Kuantifikatzaileen gainera ikerketak tradizionalki jakin-min eta arreta handia sortu eta lortu du semantikari, logika ikertzaile, edo hizkuntzaren filosofia ikertzaileen artean ere. Jakin-min honen arrazoa kuantifikatzaileek erakusten dituzten propietate matematiko eta logikoei hertsiki lotua dago. Kuantifikazioaren gainera hizkuntzalaritzazko ikerketak berebiziko garrantzia izan dute hizkuntza naturalen ikerketa teorikoan. Kuantifikazioa hertsiki erlazioatuak dauden bi ikuspuntu desberdinetatik izan daiteke ikertua: (i) barne ikuspuntu batetik, kuantifikatzaileetan beraietan oinarrituaz, hala nola, kuantifikatzaileen barne egitura morfosintaktikoa eta beraien oinarritzko propietate semantikoak ikertuaz (ikus e.g. Etxeberria 2005, 2010, 2012, Etxeberria & Etxepare 2012); (ii) kanpo ikuspuntu batetik, kuantifikatzaileak agertzen diren testuinguruak kontuan hartuaz, hau da, kuantifikatzaileek perpausaren hartzen duten posizio sintaktikoa begiratuaz eta perpaus horretan agertu daitezkeen beste kuantifikatzaile batzuekin edo beste elementu batzuekin eduki dezaketen harreman semantikoak ikertuaz. Testuinguru honetan ulertu behar da beraz 'Informazio egitura, kuantifikatzaileak, eta beraien arteko eragina: Euskaratik ikerketa konparatibo batetarantz' izenburudun proiektuaren ekarpena. Bere helburu nagusia kuantifikazioa aipatu berri dugun kanpo ikuspuntu horretatik ikertzea izango da; zehazki, proiektu honen helburu nagusietako bat kuantifikatzaileek hainbat entonazio patroien (hau da, informazio egitura desberdinen) eraginpean zer nolako portaera duten ikertzea izango da.

Onartua izan da literaturan (ikus e.g. May 1977, 1985, Beghelli & Stowell 1997, Fox 2000) bi kuantifikatzailedun perpausak interpretazio anbiguoak bultzatzen dituztela. Euskaraz, argudiatu izan da modu honetako perpausak ez dutela anbiguotasunik sortzen (ikus Elordieta 2001, Etxeberria 2002, 2005), halere, badira perpaus batzuk irakurketa anbiguoak bultzatzen dituztenak. Har dezagun (1) adibidea, zeinak (gutxienez) bi interpretazio jaso ditzakeen: (i) distributiboa, non neskek indibidualki abesti bat abestu zuten, eta (ii) kolektiboa, non neskek koru bat osatuaz abesti bakar bat abestu zuten.

(1) Neska guztiek abesti bat abestu zuten.

Aldiz, subjektu posizioan agertzen den Kuantifikatzaile Sintagmako kuantifikatzailea fokalizatzen baldin badugu, (2)ko adibidean egiten dugun bezala, irakurketa distributiboa baino ez dezakegu lortu, (1b)n adierazten denez (fokoa hizki larriz adierazten dugu).

(2) Neska [GUZTIEK]F abestu zuten abesti bat.

Beraz, badirudi, kuantifikatzailearen fokalizatzeak perpausaren interpretazio posibleetan eragiten duela hauek aldatuaz: perpausa ezanbiguo egiten du. Bestalde, kuantifikatzaile ahul deituak (ikus Milsark 1977; ikus Etxeberria 2005, et seq. euskararentzat) anbiguoak dira irakurketa kardinal baten eta irakurketa

proportzional baten artean. Horrela, (3) bezalako perpaus batek batzuk-ek lortu ditzakeen bi irakurketak adierazten ditu.

(3) Ikasle batzuk saskibaloiean jolasten ari dira.

Irakurketa proportzionalean batzuk-ek lortzen duen interpretazioa adierazi daiteke: batzuk, besteek ez (dute jolasten saskibaloiean), eta ikasleetatik batzuk kuantifikatzaile partitiboaren pareko bihurtzen da. Irakurketa hau ikasleen multzoa ezaguna denean da bakarrik posible. Irakurketa kardinalan berriz, ikasle batzuk ezin daiteke ikasleetatik batzuk bezala adierazi, eta kuantifikatzaile ahulak kantitate bati baino ez dio egiten erreferentzia. Proiektu honentzako oso interesgarria dena zera da, batzuk kuantifikatzaile ahula fokalizatzen

duzun momentuan, (4)n adierazten dugun bezala, irakurketa kardinala desagertu egiten dela eta ikasle batzuk modu proportzionalean (batzuk, besteek ez) baino ezin daitekeela interpretatu da.

(4) Ikasle [BATZUK]F ari dira saskibaloiean jolasten.

Berriro ere, (2) adibidean gertatzen zen bezalaxe, kuantifikatzailearen fokalizazioak anbiguotasuna ezabatzen du.

Aurkezpenean, datu hauentzako analisi bat proposatuko dugu: batetik Sintagma Kuantifikatzaileen osaketa semantikoan (Schwarzschild 1996, Brisson 1998), eta bestetik Rooth (1985, 1992)ren fokoaren 'Alternatiben Semantika' deituan oinarrituaz.

## 8. Euskal alokutibora joan-etorria.

Aitor Lizardi eta Manex Munduate (UPV/EHU)

alizardi002@ikasle.ehu.es mmunduate001@ikasle.ehu.es

Lan honetan euskal alokutiboaren gaineko azterketa orokorra egingo dugu Lehenik ezaugarri morfologiko zein sintaktikoak aurkeztuko ditugu. Jarraian hizkeren artean sortzen duen aldakortasunaren berri emango da. Azkenik, beste hizkuntza batzuetako fenomeno zenbaitekin ere izan lezakeen erlazioa aztertuko dugu.

Morfologiari dagokionez, alokutibo morfemaren forma eta kokalekua landuko dira. Horrez gain, aditz jokatuan morfema alokutiboa agertzeak eragiten dituen aldaketak aipatuko dira: aditz laguntzailearen txandakatzeak (1), aldaketarik eza (2) edo -i-ren agerpena (3) (Rebuschi 1984).

Azken horretan jarriko dugu fokua batez ere, eta morfema horren azalpena ematen saiatuko gara. Zenbait hizkuntzalarik diotenaren aurka, aurredatibo bezala ez hartzeko argudioak daudela aipatuko dugu, eta aurrealokutibo izena emango diogu. Horrela, aurrealokutiboaren eta aurredatiboaren arteko ezberdintasunak azaleratzen saiatuko gara.

Sintaxia aztertzean, Oihartzabalen 1993ko artikuluari jarraituko diogu eta haren argudio zenbait baliatuko dugu alokutiboa argumentuetatik bereizteko.

Bestalde, aldakortasunaren atalean aldakortasun morfologikoa eta sintaktikoa bereiziko ditugu. Morfologikoan aurrealokutiboaren bilakaerak euskalkietan izan duen askotariko emaitza emango dugu aditzera (Zuazo 2008). Sintaktikoan, murriztapen sintaktikoen oinarri gramatikala finkatu eta faktore sozial nahiz geografikoen arabera nola aldatzen diren aztertuko dugu (Alberdi 1994).

Azkenik, euskarako alokutiboa galizierako cherekin (Longa & Lorenzo 2001) eta gaztelarazko datibo etikoekin (Jaeggli 1986) erkatuko dugu eta lehengo hurbilketa batean alokutiboa galizierako chetik gaztelarazko datibo etikoetarik baino hurbilago egon daitekeela adierazi.

1. Ni etorri naiz → Ni etorri nauk/n
2. Erori zaio → Erori zaiok/n
3. Ikusi dut → Ikusi dia/nat

### Aipamenak

- Alberdi, Xabier. 1994. *Euskararen tratamenduak: erabilera*. Iker 9. Bilbo: Euskaltzaindia.
- Jaeggli, Osvaldo A. 1986. "Three issues in the theory of clitics: case, doubled NPs and extraction". In Hagit Bover (eds.) *Syntax and Semantics 19*. New York: Accademic Press, 15-42.

- Longa, Víctor M. & Guillermo Lorenzo. 2001. "Universal constraints on 'superfluous elements': The case of Galician 'Arb Che'". In J. Gutierrez-Rexach & Luis Silva-Villar (eds.) *Current Issues in Spanish Syntax and Semantics. Studies in Generative Grammar 53*. Berlin, New York: Mouton De Gruyter. 175-191.
- Oihartzabal, Beñat. 1993. "Verb Agreement with Non Arguments: On Allocutive Agreement". In Jose Ignacio Hualde & Jon Ortiz de Urbina (eds.) *Generative Studies in Basque Syntax*, Amsterdam, John Benjamins. 89-114.
- Rebuschi, Georgés. 1984. *Structure de l'énoncé en basque*. Paris: SELAF.
- Zuazo, Koldo. 2008. *Euskalkiak euskararen dialektoak*. Donostia: Elkar.

## 9. Hizkuntza-aldakortasuna Nafarroa ipar-mendebaldeko euskaran.

Oihana Lujanbio (UPV/EHU)

oihamaio@hotmail.com

Zerk eragiten du hizkuntzaren aldaketa? Zergatik eusten diote leku batzuetan gehiago bertako ezaugarri linguistikoei eta beste batzuetan gutxiago? Horrelako galderari erantzun nahian hasi dugu tesi-lan hau.

Munduko gainerako hizkuntzak bezala, euskara etengabeko doitze eta eraldatze prozesuan bizi izan da, izan hizkuntzaren barrenge indarren eraginez, edo izan hizkuntzaren inguruko indarren eraginez. XX. mendean eraldatze-prozesu horretan elementu garrantzitsu bat sartu da: euskara batuaren sorrera, ofizialki 1968an. Euskara batua sortu eta erabili egin dugu, eta onura nabarmenak ekarri ditu horrek. Alde batetik, batua euskara irakasteko eta euskara esparru formaletan erabiltzeko ezinbesteko tresna bihurtu da. Horrela, euskaldun berrien kopurua asko hazi da eta euskara esparru guztietan erraz erabiltzeko modua egin da: irakaskuntzan, administrazioan, hedabideetan, literatura idatzian, Interneten... Eta beste aldetik, euskaldunok elkar hobeki ulertzeko bide bat dugu, erdarara aldatu behar izan gabe. Azken batean, euskarak egunerokoan behar dituen hainbat funtzioarako tresna beharrezkoa dugu euskara batua.

Batua esparru horietan denetan erabiltzeak orain arte erabili izan ditugun hizkeretan eragina du. Euskal hizkerak baturantz aldatzen ari direla dirudi, baina euskal hizkera guztiek ez dituzte joera berberak erakusten. Hori da gure hipotesia, bederen. Batetik, joera diferente horiek aztertu nahi ditugu, aldaketaren ulermenean pauso bat eman nahian. Horretarako, Nafarroa iparraldeko herri batzuetako hizkerak hartu ditugu aztergai. Herri euskaldunak dira, euskararen ezagutza altukoak, eta ez oso handiak. Hizkera horiek zaharrendatik gazteengana nola ari diren transmititzen aztertuko dugu, hain zuzen.

Euskararen aldaketa aztertzerako garaian, badago beste elementu garrantzitsu bat: hizkuntza bat baino gehiago bizi dira elkarrekin Euskal Herrian, oso maila diferentetan erabiliak herriaren eta erabilera-arloaren arabera. Nafarroaren kasuan, gaztelania dugu beste hizkuntza nagusia eta bere eragina ere kontuan hartu nahi dugu lan honetan.

Hizkuntza-aldakortasuna gutxi landutako arloa da Euskal Herrian, eta oraingoa oso une aproposa iruditzen zaigu horretarako. Izan ere, euskara batua 1968an sortu zen eta harrezkero gizarteratze-prozesua bizi du. Prozesu horren gorabeherak ezagutu dituzten belaunaldiak eskura ditugu eta horietako batzuk izan dira ikerketa honen subjektuak; gazteak eta adinekoak. Aldakortasuna aztertzen duen soziolinguistikaren metodologia baliatu dugu eta hizkera horietan aldakortasunean diren hizkuntza-ezaugarri batzuk gizarte-aldagaien arabera aztertu. Estatistika izan dugu laguntzaile, zenbakien bidez dena azaltzea posible dela uste ez badugu ere. Zenbakiek joerak erakusteko balio dute eta joerak bilatzeko erabiliko dugu estatistika.

## 10. Hizkuntz aldakortasuna: Euskara Baturanzko joera elkarrizketen bidez azterturik.

Orreaga Ibarra (NUP)

orreaga@unavarra.es

Komunikazio honen helburua da euskara batuaren eragina norainokoa den ezagutzea, eta nola gertatzen den adin eta mota guztietako hiztunengan. Xede honetan oinarrituz, adin desberdinetako hiztunak hartuko ditugu eta, haien arteko elkarrizketak aztertuko. 20-30 urte bitarteko hiru elkarrizketa ekarri ditugu, eta bertze hiru 35-65 bitartekoak, azkenekoak 75-85 urte bitartekoak dira.

Jakina denez, Euskara batua ezarri da azken hamarkadetan, eta hori dela eta, aldaera estandarren ezaugarriak agertzen hasiak dira gazteen hizkeran eta baita helduen hizkeran ere. Honi bateratze 'bertikala' deitu izan zaio (Auer 1998), aldiz, dialektoen arteko bateratzeari 'bateratze horizontala' esaten zaio. Bestalde, prozesu hori bera indartsua edo ahula izan daiteke (Bellmann 1998: 25). Dударik gabe, gazteenek eskoletan ikasi dute, eta eskolako ereduak eragin du tokian tokiko hizkerarekin. Honez gain, komunikabideak, eskola eta hainbat eragin faktore berdintzaileak dira aldaera estandarra hedatu dutenak.

Thomasonek (2001) dioenez, hizkuntzen arteko ukipena dagoenean, aldaketa batzuk bertzeak baino lehenago gertatzen dira. Eta fenomeno bat baino gehiago suertatzen dira: mailakatzea edo berdinketa (ukipen dialektala dagoenean, aldaerak murrizten dira), murriztapena (ukipen dialektalen bidezko sistema osoen murriztapena), hiperdialektalismoa (honen bidez erkidego edo hizkera baten aldaera batzuk hobesten dira, besteak baztertuz)<sup>1</sup> eta aldaeren berregokitzea, (aldiera geografikoak, aldaera sozial edo estilistiko bihurtzea).

Jakina denez, eta oinarri gisa erran dezakegu, aldaketa guztiak elkarrizketaren bidez zabaltzen direla. Gisa honetan, pertsona bat bere hizketaren alde batzuk egokitzen ditu bestearekiko, eta gero, batzuetan bertze bat kutsatzen du.

Hori dela eta, interesatzen zaigu elkar eragite hau nola geratzen den ikustea, adin guztietan eta Euskal Herriko toki guztietan. Xede hau dela eta, guk elkarrizketa arrunteetan oinarritu gara, hiztunak bi aldaeren erabiltzaileak direlako (herriko hizkera eta batua), gradu desberdinetan. Izan ere, ezinezkoa da ikustea noiz gertatzen den lehenbizikoz aldaketa linguistikoa, noiz hasten den hiztun hori aldaera berri hori erabiltzen. Horretarako, hiztuna hartu beharko genuke, eta 10 edo 20 urtetan zehar segitu eta ikusi ea aldaera "berri" hori mantentzen den urte horietan guztietan. Nolanahi dela ere, gure helburua izanen da alkaera horien hasiera nola gertatzen den aztertzea eta zein diren eragileak ikustea.

### Bibliografia

- Apalauza, A., 2010, Nafarroako ipar-mendebaleko hizkeren egitura geolinguistikoa, Gasteiz. Doktore-tesia (argitaratu gabe).
- Ariztimuño, B., 2010, "Tolosako eta Ataungo hizkuntz bariazioa eta konbergentzia-joerak", Uztaro

72, 79-96.

- Auer, J. C., 1984, *Bilingual Conversation*. Amsterdam-Philadelphia: John Benjamin Publishing Company.
- Bellmann, G. 1998, "Between Base Dialect and Standard Language", *Folia Linguistica*, XXXII/1-2, 23-34.
- Silva-Corvalán, C. 2008, "The limits of convergence in language contact", *Journal of Language contact*, Thema 2, 213-224.
- Thomason S., 2001. *Language contact*, Edinburg University Press.

#### 11. Euskarazko testuen sinplifikazio automatikoa.

Itziar Gonzalez-Dios (UPV/EHU)

itziar.gonzalezd@ehu.es

Ikerketa-lan honetan euskararen tratamendu automatikorako baliagarria izango den idatzizko testuen sinplifikazio automatikoa aztertzen ari gara. Lan hau Ixa taldean<sup>2</sup> egiten da, hizkuntzaren tratamendu automatikoan lan egiten duen Euskal Herriko Unibertsitateko ikerkuntza-taldea.

Testuak sinplifikatzea edo erraztea, ingelesez Text Simplification, testu zail bat errazago edo ulergarriago bihurtzea da, jatorrizko testuaren esanahia mantenduz. Sinplifikazioa hainbat mailatan egin daiteke: lexikoa, sintaxia...; esaterako, sintaxian egitura sintaktiko konplexuak (perpau koordinatuak, mendeko perpauak, aposizioak...) erraztea da helburu, lexikoan, berriz, hitz zailak edo maiztasun gutxikoak ezagunagoengatik edo maiztasun handiagoengatik ordezkatzeta da, etab.

Testu errazek abantaila ugari eskaintzen dizkiete, bai gizakiei, baita makinei ere. Gizakien kasuan, hainbat azpildetara bidera daiteke sinplifikazioa: 1) Atzerriko hizkuntzak ikasten ari direnak, 2) Analfabeto edo gutxi alfabetatuak, 3) Haurrak, 4) Afasikoak, 5) Gorrak, 6) Gaixotasun kognitiboak dituztenak eta abar. Makinei dagokienez, sinplifikazioa aurreprozesu moduan erabil daiteke HPko zenbait sistematan: analizatzaile sintaktikoak (parser), itzulpen automatikoa, galde-erantzun sistemak, laburpen automatikoak egiteko sistemak; beste aparatuek (mugikorak, PDAk, tabletak...) dituzten pantaila txikietan informazioa modu erosoan agertzeko ere baliagarria izan daiteke.

Testuen sinplifikazio automatikoan lan gehienak ingeleserako egin diren arren, azken urteotan beste hizkuntzetarako ere sistemak sortzen hasi dira, hala nola Brasilgo portugaserako, arabierarako, suedierarako, gaztelanierarako, frantseserako, danierarako, euskararako eta italierarako. Gainera, azken urte

---

2 <http://ixa.si.ehu.es/ixa>



hauean ikerketa-lerro hau garrantzitsua bihurtu da eta uneoro testuak sinplifikatzeko metodo berriak argitaratzen ari dira.

Euskarari dagokionez, gure lana bi atazatan bana daiteke: alde batetik, euskarazko egitura sintaktiko konplexuen azterketa egiten ari gara eta, bestetik, testuen sinplifikazioa automatikoki egingo duten tresnen egokitzapenean edo garapenean dihardugu. Zeregin horretan perpausen mugak identifikatzen dituen gramatika eta aposizioak detektatu eta sailkatzen dituen gramatika osatu ditugu.

## 12. EPEC corpora predikatu mailan etiketatzeko eta ustiatzeko oinarriak: BVI, EPEC-RoISem eta e-ROLda

Ainara Estarrona (UPV/EHU)

ainara.estarrona@ehu.es

Gure lana hizkuntza naturalaren prozesamenduaren alorrean kokatzen da eta zehazki IXA Taldearen lan-ildoan. IXA Taldean euskararen prozesamendurako erreferentziazko corpus bat (EPEC) (Aduriz et al., 2006) eraikitzen ari gara. Corpus osoa, eskuz, hiru mailatan etiketatu dugu oraindano: segmentazio mailan; sintaxi mailan, dependentizen eredu jarraituz (Aranzabe, 2008; Aldezabal et al., 2009); eta semantika mailan, Euskal Wordneteko adierak jarraituz eta soilik izen kategoriakoei gagozkiela (Pociello et al., 2010).

Tesi-lan honetan EPEC corpora predikatu mailan etiketatea dugu helburu, hots, adiera, balentzia, rol semantikoak eta zenbaitetan, hautapen murriztapenak. Horretarako, beharrezkoa dugu eredu bat eta guk PropBank-VerbNet (Kipper, 2000; Palmer et al., 2005a) eredu hautatu dugu nagusiki ondoko arrazoiengatik (Agirre et al., 2006):

- Abiapuntuzat sintaktikoki etiketatutako corpora dute: PennTreebank2 (Marcus, 1994).
- Beste hizkuntza batzuetarako ere masiboki erabili izan da (hindia, txinera, koreera, arabiera, gaztelera, katalana) (Bhatt et al., 2009; Palmer et al., 2005b; Palmer et al., 2006; Palmer et al., 2008; Xue, 2008; Xue & Palmer, 2009; Civit et al., 2005).
- Taldean alde aurretik egindako lanen antzeko ildoak proposatzen dira aditzak tratatzeko (Euskal Aditzen Datu Basea – EADB) (Aldezabal, 2004).

Predikatu mailan etiketatutako baliabide honek hizkuntzaren prozesamenduko hainbat atazetan lagunduko luke, hala nola, itzulpen automatikoan, informazioaren erauzketan, adiera desanbiguazioan, ikasketa automatikoan... eta era berean, hainbat fenomeno linguistiko aztertzeko eta euskal aditzen bilduma osatzeko bidea emango liguke.

Corpora etiketatzeko ezinbestekoa da metodologia bat zehaztea eta hau izango da tesi-lan honen beste funtsezko ekarpena. Metodologia 3 fasetan banatu dugu:

- Lehen hurbilpena: Eredua egokia dela berretsi eta eskuliburuaren lehen bertsioa
- Metodologiaren oinarriak finkatzen: eskuzko etiketatea laguntzeko bide erdi-automatikoak ere jorratu ditugu.
- Behin betiko metodologia: corpora modu erdi-automatikoan etiketatzeko metodologia proposatzen dugu.

Etiketate-prozesu honetan bi gauza aztertuko ditugu nagusiki:

- Proposatutako ereduak eta metodologiak corpusaren kalitatezko etiketatea bermatzen ote duten
- Hizkuntzen arteko aldeak eta parekotasunak

Eta ondorio bezala, ondoko emaitza edo baliabideak izango ditugu aditzen inguruan ikerketa egin nahi duen edozeinen eskura:

- Euskal aditzen lexikoia: Basque Verb Index (BVI)
- EPEC corpora predikatu mailan etiketatua: EPEC-RoLSem

Ikerketa errazteko asmoz, bi baliabide hauetan dagoen informazioa ikusgai jarri eta bilaketak egiteko aukera ematen duen tresna ere sortu dugu: e-ROLda.

## Bibliografia

- Aduriz, I., M. J. Aranzabe, J. M. Arriola, A. Atutxa, A. Díaz de Ilarraza, N. Ezeiza, K. Gojenola, M. Oronoz, A. Soroa and R. Urizar. (2006). Methodology and steps towards the construction of EPEC, a corpus of written Basque tagged at morphological and syntactic levels for automatic processing. In Andrew Wilson, Paul Rayson and Dawn Archer (eds.), *Corpus Linguistics Around the World*. Book series: *Language and Computers*. Vol. 56, 1-15. Rodopi (Netherlands).
- Agirre, E., Aldezabal, I., Etxeberria, J., Pociello, E. (2006). A Preliminary Study for Building the Basque PropBank. *Proceedings of the 5th International Conference on Language Resources and Evaluations (LREC)*. Genoa, Italy.
- Aldezabal, I. (2004). Aditz-azpikategorizazioaren azterketa. 100 aditzen azterketa zehatza, Levin (1993) oinarri harturik eta metodo automatikoak baliatuz. Leioa (Bilbao), Euskal Herriko Unibertsitatea, tesi-txostena.
- Aldezabal, I., Aranzabe, M.J., Arriola, J.M., Díaz de Ilarraza, A. (2009). Syntactic annotation in the Reference Corpus for the Processing of Basque (EPEC): Theoretical and practical issues. *Corpus Linguistics and Linguistic Theory* 5-2, 241-269. Mouton de Gruyter. Berlin-New York.
- Aranzabe M.J. (2008). Dependentsia-ereduan oinarritutako baliabide sintaktikoak: zuhaitz-bankua eta gramatika konputazionala [Syntactic resources based on Dependency Grammar: Treebank and Computer Grammar]. Basque Philology Department (UPV/EHU). PhD.
- Bhatt, R., Narasimhan, B., Palmer, M., Rambow, O., Sharma, D., Xia, F. (2009). A Multi-Representational and Multi-Layered Treebank for Hindi/Urdu, In the *Proceedings of the Third Linguistic Annotation Workshop*, held in conjunction with ACL-IJCNLP 2009, Singapore.
- Civit, M., Aldezabal, I., Pociello, E., Taulé, M., Aparicio, J., Márquez, L. (2005). 3LBLEX: léxico verbal con frames sintáctico-semánticos. In *XXI Congreso de la SEPLN*. Granada, Spain.
- Kipper, L., Dang, H.T., Palmer, M. (2000). Class-Based Construction of a Verb Lexicon. *AAAI/IAAI*: 691-696.
- Marcus, M. (1994). The Penn Treebank: A revised corpus design for extracting predicate argument structure. *Proceedings of the ARPA Human Language Technology Workshop*. Princeton, NJ.
- Palmer, M., Gildea, D., Kingsbury, P. (2005a). The Proposition Bank: A Corpus Annotated with Semantic Roles. *Computational Linguistics Journal*, 31:1.
- Palmer, M., Nianwen, X., Babko-Malaya, O., Chen, J., Snyder, B. (2005b). A Parallel Proposition Bank II for Chinese and English. *Frontiers in Corpus Annotation, Workshop in conjunction with ACL-05*. Ann Arbor, MI: 2005.
- Palmer, M., Ryu, S., Choi, J., Yoon, S., Jeon, Y. (2006). Korean PropBank. *Linguistic Data Consortium, Philadelphia*. LDC2006T03. ISBN: 1-58563-374-7.
- Palmer, M., Babko-Malaya, O., Bies, A., Diab, M., Maamouri, M., Mansouri, A., Zaghouani, W. (2008). A Pilot Arabic PropBank. *Proceedings of LREC 2008*. Marrakech, Morocco.
- Pociello, E., E. Agirre and I. Aldezabal. (2010). Methodology and Construction of the Basque WordNet. *Language Resources and Evaluation (LRE) Journal*.
- Xue, N. (2008). Labeling Chinese predicates with semantic roles. *Computational Linguistics*, 34(2): 225-255
- Xue, N., Palmer, M. (2009). Adding semantic roles to the Chinese Treebank. *Natural Language Engineering* 15, 1 (Jan.), 143-172.

### 13. Ezagutza linguistikoa Ingelesa-Euskara itzulpen automatiko sistemen garapenean.

Nora Aranberri (UPV/EHU)

nora.aranberri@ehu.es

Bada denbora IXA taldea Matxin espainieratik euskararako itzultzailea garatzen ari dela. Matxin erregeletan oinarritutako sistema da, eta beraz, ezagutza linguistikoan oinarritzen da itzulpenak egiteko. Bada, bi espezialitate bateratzen ditu proiektu honek: informatika eta hizkuntzalaritza.

Itzultzailearen garapenean, helburuetako bat, itzultzailearen egitura informatikoa mantenduz eta ezagutza linguistikoa aldatuz, sistemak beste iturburu edo xede-hizkuntzekin lan egitea bermatzea da. Portabilitate aukerak eta mugak neurtzeko asmotan, espainieratik euskararako itzultzailearen garapena abian da.

Erregeletan oinarritutako sistemek oinarritzko hiru zati dituzte: (1) iturburuko testuaren analisia egiten duen modulua (hitz-kategorien identifikazioa, esaldien zuhaitz sintaktikoak ezagutzea eta esaldiko elementuen funtzioak eta dependentziak markatzea, esaterako); (2) baliokidetzak edo transferentzia erregelen bidez, iturburuko analisiari xede-hizkuntzako analisi baliokidea esleitzen dion modulua; eta (3) analitiko xede- hizkuntzako testua sortu eta hitzen ordena egokia ezartzen duen modulua.

Espainiera->euskara itzultzailetik ingelesa->euskara itzultzaileara pasatu nahi dugunez gero, analisi modulu guztiz berria behar dugu, Matxinek eskatzen duen analisi maila eta sakontasuna emateko gai izango dena. Ingelesaren zabalkuntza- eta teknologia-garapen maila altuak bide, tresna hauek eskuragarri daude gaur egun. Hala, guk Stanford CoreNLP analizatzailea integratu dugu Matxinen. Transferentzia moduluari dagokionez, aldaketak behar dira, hizkuntza-bikote berriarentzako propioak diren baliokidetzak zehaztu behar baitira. Hala ere, lehenengo hizkuntza-bikoterako egindako erregelak gida bezala erabiltzen dira jorratu beharreko erregelak identifikatzeko. Sarritan, erregela orokorrak birziklatzea posible da. Azkenik, sorkuntza eta berrordenaketa moduluak ia bere horretan berrerabil daitezke.

Sistemaren egitura eta tresneria finkatuta, Matxinek oinarritzko informazio linguistikoa behar du, eta hemen dator hizkuntzalararen lana. Batetik, hiztegi elebiduna lortu behar da, lexikoi ahalik eta zabalena ezagutu dezan, hitz anitzeko unitate lexikalak barne, eta kategoria itxiak ahaztu gabe. Bigarrenik, esaldi-egiturari dagokionez, ezinbestekoa da analisisian lortutako unitate, erlazio eta bestelako informazioa kodetzean erabilitako etiketen eta euskarako analisi mailan erabiltzen diren etiketen baliokidetzak-eta eguneratzea. Halaber, aditzen itzulpena bideratzeko, transferentzia erregelak idatzi behar dira paradigmak, denbora, aspektua, aktibotasuna, menpekotasuna eta abar identifikatzeko eta informazio hori sorkuntza modulura pasatzeko. Iturburuko hizkuntzako preposizioen eta euskarazko postposizioen arteko baliokidetzak ere idatzi behar dira, zeintzuek desanbiguzio ariketa sakon bat eskatzen duten maiz. Askotan, testuinguru hurbileko informazio sintaktikoaz zein semantikoaz baliatu gara desanbiguzio hau lortzeko. Sakontasun linguistiko desberdineko corpusen azterketak ere laguntzen du atal honetan. Azterketa horietan aditzek jarraitzen dituzten patrioiak identifikatzen saiatzen gara. Azkenik, sintagma barruko osagaien ordena eta esaldiko sintagmen hurrenkera berrezarri behar dira euskararen syntaxira egokitzeko.

Erregeletan oinarritutako sistemei maiz egotzi zaie garapen motel eta garestikoak izatea. Hala ere, ikusten da hizkuntza-bikote berri bat garatzerakoan berrerabili daitezkeen modulu, datu eta teknikak nabarmenak direla. Matxinen, sistema-egituran aldaketarik gabe eta tresneriaren gutxieneko ordezkapen batekin, ezagutza linguistikoa berrituz, hizkuntza-bikote berri baterako prototipo bat lortu dugu.

#### 14. BASYQUE: hizkeren arteko aldakortasun sintaktikoa aztertzeo aplikazioa.

Larraitz Uri, Gotzon Santander, Xan Aire eta Ricardo Etxepare (UPV/EHU, CNRS-IKER, IXA)

larraitz.uria@ehu.es

##### Komunikazioaren laburpena

Gaur egun, leku eta egoera guztietan komunikazio eraginkorra erdiesteko kode estandar idatzia bakarrik menderatzea ez da nahikoa. Aitzitik, tokian tokiko komunikazio moduei ere arreta berezia jartzea beharrezkoa da gure hizkuntza komunitaterako ekarpen interesgarriak egin nahi baditugu.

Proiektu honetan, BASYQUE aplikazioa<sup>3</sup> garatu da hizkeren arteko aldakortasun sintaktikoa aztertzeo helburuarekin (gure proiektua Iparraldeko hizkeretara mugatu den arren, aplikazioa erabilgarria da edozein hizkeren azterketa bideratzeko). BASYQUEren bidez hizkera jakinen ezaugarriak ezagutzeko edota aztertzeo beharrezkoak diren datuak bildu eta kontsulta daitezke. Informazio-iturrien arabera, hiru atal bereizi dira: galdeketak, literatur corpusak eta bestelako iturriak.

Galdeketen atalari dagokionez, 6 galdeketa egin zaizkie Iparraldeko 30 herritan bizi diren 46 datu-maileri. Galdeketa bakoitza fenomeno linguistikoko jakin bati dagokio (datibo-komunztadura, postposizio-egiturak, determinatzailearen erabilera, omen eta ote partikulen posizioa eta zenbatzailearen eta aditzaren arteko komunztadura). Guztira, 13.200 erantzun inguru bildu dira orain artean Iparraldean egindako galdeketen bidez.

Literatur corpusari dagokion atalean, XX. mendeko bost autoreren lanetan aurkitutako adibideak (aztergai ditugun fenomeno linguistikoei lotutakoak) bildu dira orain artean. 2.940 adibide, guztira.

Azkenik, bestelako iturriak atalean, beste proiektu batzuetan bildutako informazioa edota testigantzak jasotzen dira, euskal dialektologiaren alorrean lanean dihardugun ikerlari batzuen artean sortutako elkarlanari esker. Orain artean, Norantz<sup>4</sup> proiektuan bildutako hainbat adibide ekarri dira datu-base honetara eta Ahotsak.com<sup>5</sup> proiektuan egindako elkarrizketetan entzuten diren adibideak jasotzeko asmoa ere badago.

Ongi hornitutako datu-baseak askotariko azterketak egin eta ondorio linguistikoak ateratzeko aukera eskaintzen du, hizkeren artean dagoen bariazio sintaktikoaren inguruan dauden hipotesiak betetzen diren edo ez ikusteko aukera. Eta azterketa eta ondorio linguistikoak ateratzeaz gain, proiektu honek badu beste helburu bat ere. Izan ere, bildutako informazioa oinarri hartuta, EHUko IXA taldean aztertutako hizkeren tratamendu automatikoa nola bidera daitekeen aztertzeo lehen urratsak egiten gabiltza, izaera dialektala duten corpusak aztertu eta automatikoki tratatzeko tresnak garatzeko helburuarekin.

3 <http://ixa2.si.ehu.es/atlas2/index.php?lang=eu>

4 <http://www.iker.cnrs.fr/~norantz-contact-de-langues-et-.html?lang=f>

5 <http://www.ahotsak.com>

## 15. Gazteak lagun artean idazten.

Elixabete Perez eta Esther Zulaika (Deusto)

elixabete.perez@deusto.es esther.zulaika@deusto.es

Azken urteotan lantzen ari garen komunikazio-gaitasunen ikerlerroan kokatzekoa da gure komunikazioa. Zehatzago esan, euskaldun gazte-nerabeek komunikatzeko gaitasuna idatziz nola gauzatzen duten dugu oraintsuko jomuga. Honenbestez, esan daiteke hizkuntzalaritza aplikatuko ikerketa dela hauxe. Komunikazio-esparru formal, maila akademikokoa ikertu ondoren (Perez Gaztelu, E, Zulaika, E. (2012). Idatziz komunikatu. Gazteak eskolan idazten), lagunarteko esparru “askeagoan” euskal gazte-nerabeek bere artean euskaraz nola idazten duten ikertzea dugu xede.

Horretarako, sare sozial bateko gazte-nerabeen arteko euskal elkarrizketa naturalen corpusa eskuratuta, elkarrizketa horietan darabil(tza)ten kodearen, kodeen ezaugarriak xehe deskribatu eta lerro nagusiak sistematizatzerira iritsi nahi dugu.

Landu dugun corpusa solaskideek bere borondatez eskuetaratu digute, hots, benetako elkarrizketa naturalak gure esku utzi dituzte iker ditzagun.

Gure arreta ez dago bideratua gazteen hizkuntzaren aukeran (demagun noiz gaztelaniaz noiz euskaraz egitea hautatzen duten) eragiten duten baldintzak aztertzerara; solasean darabilten bitarteko idatzia, “hizkuntza lehengaia” ikertzera baizik.

Aipatu ditugun bi alderdiok bereizten gaituzte Soziolinguistikako klusterrak azken urteotan —bereziki Gazteak, sare sozialak eta hizkuntza kudeaketa [2012ko txostena](#) da aipatzekoa— egin duen lan eskergatik, eta, aldi berean, lan horien osagarria ere bada gurea<sup>6</sup> Nolabait esan, txosten horretan “fresko” eta “informaltzat” jotzen den hori zer eta nolakoa den arakatu nahi dugu.

Hainbat hizkuntzalariren (Crystal 2003, 2004, 2011..., Amonarriz [“Algara edo ez gara”](#) bloga ... ) ildo beretik, gure hipotesia da euskal gazteok ere “askeagotzat” dituzten komunikazio esparru digital eta lagunarteko komunikaziorako erabiltzen dituztenetan bere euskal kode idatzia eratzen ari direla.

Erregistro jaso idatzian (akademikoan batik bat) darabiltena ez bezalakoa da eta erabilesparua mugatua du kode horrek. Alegia, gazte-nerabe euskaldunek garbi bereizten dute komunikazio-egoera, helburu, solaskideen arteko harremanen eta komunikazio saioetan eragina duten aldagaien arabera zein kodez eta nola baliatu.

Bistan da, beraz, gure ikuspegiaren arabera, ez dugula hizkuntza idatziaren erabilera hau hizkuntzaren hondamendiarekin parekatzen, adierazteko modu berriak irabaztearekin baizik.

<sup>6</sup> Esate baterako, 2012ko txosteneko kasuen deskribapenean gisa honetako informazioa ematen dute. Zerk eusten dio, ordea, hiztunak “(in)formaltzat” daukan hizkerari, zein hizkuntza lehengai? Horixe nahi genuke zehaztu:

Zein motatako hizkera darabilzue? “Formalismotik” urruntzen saiatzen dira, jendearengana iristeko hobea dela iruditzen zaielako.

Euskalkia ala batua? Euskara batua baina “informalagoa” erabiltzen dute.

Formala ala informala? Modu informal batean idazten saiatzen dira. Hala ere, ez dute beti lortzen. (Soziolinguistika klusterra 2012: 15).

## 16. Aldakortasun sintaktikoa euskararazko perpaus nominalizatueta.

Maia Duguine (UPV/EHU)

Ingelesa edo gaztelania bezalako hizkuntzetan, morfo-fonologikoki gauzatzen ez denean, perpaus jokatueta subjektuak gehienetan interpretazioa perpaus nagusiko DS baten arabera du:

- (1) [Al llegar  $\Delta i/*j$  a casa], Juani preparará la cena. *Gaztelania*  
 (2) Johni wants [ $\Delta i/*j$  to go home]. *Ingelesa*

Goenagak (1984), Ortiz de Urbinak (1989) eta Artiagoitiak (2003) (besteren artean) azpimarratu dutenez, euskarazko -t(z)e nominalizazio egituretan, gauzatzen ez den subjektu batek irakurketa askea eduki dezake:

- (3) Aitorreki nahi luke [ $\Delta j$  etxera joatea]. *Euskara*

Hitzaldi honetan, datu hauek modu xehe batean aztertuko ditut, bi ezaugarritan oinarrituz bereziki: nominalizazioen egitura morfosintaktikoa eta euskalkiz euskalki behar daitezkeen aldakortasuna.

Ezaugarri morfosintaktiko desberdineko perpaus nominalizatuak daude; bost multzo bereiz daitezke (ik. Artiagoitia 2003): (A) perpaus nagusiko argumentu posizioan agertzen direnak (kasu ergatibo, absolutibo edo datiboarekin; ikus (3) adibidea), (B) perpaus nagusiko adjuntu direnak, (C) aginte aditzek osagarri gisa hautatzen dituztenak (-t(z)eko atzizkiarekin gauzatzen direnak), (D) mugimendu aditzek osagarri gisa hautatzen dituztenak (-t(z)era atzizkiarekin gauzatzen direnak), eta (E) -t(z)en atzizkiarekin gauzatzen direnak (hainbat motako aditzek hautaturik: pertzepzio aditzak, aditz aspektualak, etab.).

Gai honi buruzko aurreko lanetan aurki daitezkeen datuetan oinarriturik, bai eta nire aldetik bildutako datuak erabiliz, erakusten dut orain arte oharkabean egona zen aldakortasuna dagoela subjektu isilaren irakurketa askearen posibilitatean.

Zehazkiago, gutienez ondoko bi barietateak bereizi behar direla erakusten dut:

- (4) 1. barietatea (erdialdeko euskara).

- Subjektu isilak irakurketa askea eduki dezake A, B eta C motako perpausetan.
- Subjektu isilak ezin dezake irakurketa askea eduki D eta E motako perpausetan.

- (5) 2. barietatea (nafar-lapurtera).

- Subjektu isilak irakurketa askea eduki dezake A motako perpausetan.
- Subjektu isilak ezin dezake irakurketa askea eduki B, C, D eta E motako perpausetan.

Datuak honakoak dira. A motako perpausa (3)n itxuratua da: irakurketa askea bi barietateetan lortzen da. B-tik E-rako nominalizazioei buruzko datuak (6)tik (9)rako adibideetan erakusten dira:

- (6) B: [ $\Delta j$  etxetik ateratzean] konturatu da Joni giltzak barruan zeudela. (1. barietatean ✓; 2. bar. ✗)  
 (7) C: [ $\Delta j$  ixilik egoteko] eskatu zidan (nirii). (1. barietatean ✓; 2. bar. ✗)  
 (8) D: Joni [ $\Delta j$  artikulua idaztera] behartu dugu. (1. barietatean ✗; 2. bar. ✗)  
 (9) E: Joni [ $\Delta j$  artikulua idazten] ikusi dugu. (1. barietatean ✗; 2. bar. ✗)



Irakurketa askea duten subjektu isilak syntaxian baimenduak izan behar direlako hipotesia aintzat harturik (Rizzi 1986, Holmberg 2010, Barbosa 2011), 1. eta 2. barietateetan B eta C motako perpausek egitura sintaktiko desberdina dutela argudiatzen dut. Zehazkiago, proposatzen dut subjektu isilen baimentzea Kasu sintaktikoaren ezarketaren arabera dela, eta 2. barietatean, B eta C egituretan subjektuak ez duela Kasurik jasotzen.

Amaitzeko, mendebaldeko euskarako datuak eztabaidatzen ditut laburki, hirugarren barietate bat ere existitzen dela erakutsiz. Barietate honetan, subjektu isilaren irakurketa askea posible izateko funtsezkoa ematen duena, nominalizazioak -a determinatzailea buru gisa edukitzea da. Artiagoitia (1994), San Martin (2004), eta Etxepareri (2006, 2012) jarraiki, determinatzaileak perpaus nominalizatuen barruko syntaxian eduki dezakeen eragina eztabaidatuko dut.