

Euskal hitz ibilienak

Pello Salaburu

UPV/EHU

Laburpena

Lan honetan biltzen da gaur egungo euskaran maizen ibiltzen ditugun hitzen zerrenda. Lemaka multzokatu dira, testuetan duten agerraldien araberak. Horretarako 205 milioi hitz biltzen dituen XXI. mendeko testuak aztertu dira: literatura, egunkariak, itzulpenak, eta abar. Azterketa sistematikoa egin da, eta lemak kategoriak bilduak ematen dira, zenbateraino ibiltzen ditugun, hala sailkatuak. Anitzetan usteak ditugu gai honen gainean eta, arrazoi garbirik gabe, hitz jakin batek bertze batek baino maiztasun handiagoa duela pentsatzen dugu: orain aurkezten da azterketa sistematiko bat arazo honetaz. Jakina, ibili dugun datu basea aski handi eta fidagarria izanik ere, ez dakigu euskara mintzatuan ibiltzen ditugun hitzekin antzeko zerrenda osatuko genukeen. Ez dugu inolako arrazoirik ikusten, zernahi gisaz, arras diferentea izanzen litzatekeela pentsatzeko.

0. Sarrera*

Modan jarri ziren orain dela urte batzuk «oinarrizko testuak», «testu errazak», «testuak euskaldun berrientzat» eta holako lanak. Baziren testu batzuk, ustez, euskara errazean idatziak, edozeinek, ez horrenbertze jakinik ere, ulertzeko moduan eskribituak. Testu haiek leitzen nituenean banuen beti sentsazio bat ezin baztertua: batetik, iduri zuen halako pareta zurrun batzuen barnean bildutako testuak zirela pasarte haiek eta, bertzetik, bazirenan hainbertze hitz niri biziki arrotzak egiten zitzaizkidanak. «Jasangaitz ez banaitzaizu», «ile-apaintzaileak», «suhiltzaileak» eta holako gauzak ikusten nituen nik orrialde haietan. Sortzetik euskaraz egin badut ere, egia da zernahi

* Emaizta hauetara ekarri gaituen ikerketak Europar Batasunaren ikerketarako, teknologia garatu eta erakusteko Zazpigarren Esparru Programaren laguntza jaso du 613465 diru-laguntza hitzarmenaren bitartez eta EHuko Euskara Institutuak (www.ei.ehu.es) egiten dituen lanetan oinarritu da. Lan horiek Eusko Jaurlaritzaren Hezkuntza Saileko Unibertsitate eta Ikerketa Sailburuordetzak, eta Hizkuntza Politikako Sailburuordetzak ematen dituzten laguntzei esker egiten dira. Halaber, EHuko Euskara errektoreordetzak eta Lagun Aro Fundazioak ere ematen dituzte laguntzak.

gisaz ez nuela nik nire hizkeran haien ordezeko hitz egokirik aurkitzen, «pelukero», «bonbero» eta holakoak ez baziren. Beti galdetzen nion nire buruari, hala ere, hitz hau edo hura aukeratzean, non eta zertan oinarritzen zen idazlea. Ez hitza aukeratzeko, idazle orok du eskubide osoa nahi dituen hitzak aukeratzeko, baizik baieztatzeko hala zela, hitz horrek «euskara erraza» zelako hartan bazuela lekua, bitxi samarra baitzitzaidan ikustea euskaraz xarmanki egiten zuen jendeak ez zituela batere ibiltzen.

Nik dakidala, ez da inoiz ere egin azterketa sistematikorik honen gainean, hau da, ez da aztertu zein diren egunero gehien ibiltzen ditugun hitzak, zein diren, euskaraz mintzatu nahi badugu, ezinbertzez ibili beharko genituzkeen lehentasunezko hitzak. Ez da batere kontu erraza. Ez da kontu erraza, informazio fidagarria beharko baikenuke hitz kategoriei buruz, eta hitzak biltzen dituzten egiturak ere aztertu beharko baikenituzke. Bi kontu horiek modu objektiboan aztertuko bagenitu, hau da, modu objektiboan izaten bagenu hitz ibilienean eta hitz horiek besarkatzen dituzten egituren berri, orduan bai, orduan arrazoi eta argudio sendoak izanen genituzke errateko, oinarritzko testu bat eskribitzen badugu, hor ageri diren hitzak direla euskaldunak maizen ibiltzen dituenak. Informazio horrekin oinarritzko hiztegi bat egiten badugu, aise argudiobide handiago eta bortitzagoak izanen genituzke zinez hitz horiek direla oinarritzko hitzak pentsatzen hasteko. Eta ikusiko genuke, beharbada, erraztat jo ditugun zenbait hitz eta egitura aski arrotzak direla azkenean.

1. Hitzen kategoriak

Lehenbiziko arazoa hitzen kategorizazioarekin lotzen da. Nola sailkatzten dira hitzak euskaraz? Hemen kategorizazio lexikoez ari gara, jakina, ez kategoria funtzionalez.¹ Dakigun gisan, kategoria lexikoak sailkapen kutxak bezalakoak dira. Hala, *etxe*, *hodei*, *andre*, *karrika* edo *irin* bezalako hitzak izenak direla erraten dugu. Izenek badituzte propietate morfologiko eta sintaktiko batzuk bertze hitzek ez dituztenak. *Gaur*, *etzi* edo *aurten* bezalako hitzak adberbiotzat hartzen ditugu, eta ez dira izenen klasekoak. Hala, izenek, izenondoa (adjektibo ere deitua), har dezakete aldamenean: *karrika polita*. Baina aditzondoek (adberbio ere deituek) ez dute holakorik onartzen: **etzi polita*. Bertze hitz mota batzuk aditzak dira: *iraun*, *bota* edo *hartu*, konparazione. Eta baditugu ere izenondoak: *handi*, *eder*, edo *beltz* bezalakoak. Hitzen kategoriak bereizteko argudiobideak ez ditugu hemen garatuko, horretaz zerbait erranen badut ere.

¹ Arazo hau, berez, aise nahasiagoa da. Guk hemen hiztegi eta ohiko gramatiketan ibiltzen den sailkapena hartu dugu kontuan, baina gauzak ez dira hain gardenak. Kontu honetaz I. Bosque (2007) ikusten ahal da.

Euskaraz gehien ibiltzen diren hitzen berri eman nahi baldin badugu, komeni da hitzak kategoriaka sailkatzea, datuak pixka bat ordenaturik emateko. Ordea, batzuetan ez da batere erraza jakitea zein kutzatan sartu behar ditugun euskal hitzak. Jo dezagun, adibidez, *hurbil* hitza. Hiztegietan kontsultatuz gero, OEHN² adibidez, edo EEHan, ikusten dugu adberbioa dela, eta adjektiboa ere. Eta hori hala da: *Miren hurbil bizi da* (adb), *Miren neska hurbila da* (izond). Jakina, *Miren, hurbil zaitez!*, erraten badugu aditza da. Muga jartzea ez da izaten betere erraza batzuetan. Jo dezagun, konparazione, *atsegin* hitza: *Jainkoaren atseginak eta plazerak gehienik ariman barrena dira*, zioen Axularrek. Bistan da hor izena dela. Aldiz, *mutil atsegina da hori* diogunean izenondoa da. Eta batzuetan aditzak sortzeko erabili dugu: *atsegin du mendira joatea* edo *atsegin hartu du jakin duenean*. Holako zerbait gertatzen da bertze hitz batzuekin ere: *prest, gai, erraz, zail, trebe*, eta abar. Berrez, hitz horiek izenak (*gai*), izenondoak (*trebe, zail*), aditzondoak (adberbio ere dietua) (*prest*) izan daitezke, baina *izan* edo *egon* erantsiz gero predikatu berriak sortzen dituzte, perpaus osagarriak eskatzen dituztenak: *prest egon, zaila izan* edo *trebea izan*, konparazione, aldaera guztiekin. Horregatik diot batzuetan zehaztasunak ematea ez dela hain erraza izaten. Eta hori aski ez denez, nahasteak bertze toki batzuetatik ere sortzen dira: *kantari* bezalako hitza izena dela erranen genuke (ia beti, batzuetan izenondo bezala ageri bada ere), eta horren pareko diren *singer*, ingelesez, edo *chanteur*, frantsesez, izenen artean sailkatzen dituzte hizkuntza horietako hiztegiek. Gauza bera gertatzen da *estudiante* hitzarekin. Baina RAEk (2013) sarean duen hiztegian (<http://www.rae.es/>) *cantor* edo *cantante* adjektiboak direla erraten da. Horien pareko izan daitezkeen bertze anitzetan, ordea, ez du deusik zehazten: *pintor, escultor*, edo *escritor*, adibidez. *Aditu*, ordea, izenondotzat hartzen da OEHN, *especialista* ere izenondoa da RAErentzat, baina *specialist* ingelesez izentzat hartzen dute. Sarasola (2010) ere kontsulta daiteke.

2. Lana nola egin den

Lan hau egiteko EHUko Euskara Institutuak sarean duen Euskal Hiztegiaren Maiztasun Egitura (EHME) izeneko aplikazioaz (<http://www.ehu.es/ehg/ehme/>) baliatu naiz (Landa et al 2011). Testu masa handia dago hor, nahikoa hemen aipatzen dugun kontu honetaz zerbait funtsatua errateko. Oinarritu, *Ereduzko Prosa Gaur*-EPG (Sarasola et al 2009) izeneko corpusen oinarritzen da, baina garbiketa bat egin da lehenik, datuak fidagarriagoak izateko. Hala, aplikazio horrek 22.704.373 hitz aztertzen ditu. Badira hor 377.795 hitz desberdin eta, orotara, 53.310 lema. Hobeki konprenitzeko: *etxe, diot, etorri, apal, bihar, hura, zu, dakarkio* edo *herenegun* hi-

² Mitxelena & Sarasola (1987-2005) *Orotariko Euskal Hiztegia*.

tzak dira, hitz desberdinak. Eta horiek anitzetan errepikatzen direnez, testu masa horretan ditugun hitz guztiak, banan-bana kontatuz, 22,7 M hitz biltzen direla diogu. Errepikatu gabe kontatzen baditugu, 377.795 dira orotarar. Eta hitz horiek 53.310 lematan (*etxe, etxetik, etxeko, etxera*, hitz desberdinak izanik, lema bakarra dute: *etxe*) biltzen dira. Lemak eta hitzak bereiztea gauza sinplea da batzuetan, baina bertze batzuetan ez da, ordea, hain gardena: *ditut, zenigun, dezakegu* edo *niezaioke* hitzak, *ukan* lemaaren aldaerak direla erranen dugu (ez gara hemen **edun* eta **ezan* bereizketekin hasiko). Baina kontura gaitezen, hor ere, *ditut* hori laguntzailea izan daitekeela (*ekarri ditut*), edo aditz nagusia: *hogei urte ditut*. *Etorri* lema aditza izan daiteke (*berandu dator*), edo izena ere (*etorri handiko bertsolaria da*). Ez hori bakarrik: aditza denean ere, ez du beti erabilera bakarra. Batzuetan bakarrik heldu da, goiko adibidean bezala, eta bertze batzuetan *ongi, gaizki* eta tankera horretako aditzondoak behar ditu ondoan, eta arruntean datiboa eskatzen du: *ongi etorriko zaizue holako gauzak irakurtzen hastea*. Edo aditzondo horiek gabe ere, betiere erranahiak aldatzen direla: *ator hori ongi datorkizu*.

Ibili dugun corpusa ez da txikia, ikusten denez, baina baditu bere mugak, eta komeni da ongi zedarritzea hori ere: lema gehiegi dago, hasteko, eta lemen zerrenda laburtzea komeni da, garrantzi handiko lemak eta bigarren mailakoak bereizteko. Bestetik, erregistro jakin bateko hitzak dira guk aztertuak: batez ere literaturan (dela euskaraz zuzenean sortua, dela literatura itzulia) eta prentsa artikuluetan oinarritzen baita gure corpusa. Ez dago hor ahozko daturik, adibidez. Baina hala izanik ere, badut uste corpus egokia den, material anitz baitago hor, eta gaur egungo hiztunek ibiltzen ditugun hitzak aski ongi bildurik baitaude hor: hiztunok aise lema gutiago behar ditugu eguneroko hizkeran, eta behar ditugun guztiak prentsak edo literaturak ibiltzen ez baditu ez dakit nik nolako prentsa edo literatura egingen dugun.

Ondoko lerroetan euskaraz gehien ibiltzen ditugun hitzen zerrenda emanen dut, eta saiatuko naiz hitzak nolabaitere ordenatzen. Nolabaitere ordenatzen, diot, kategoria gurutzadura eta anbiguotasun horien argitzeak bertze lan mota bat eskatuko bailiguke: noiz jakin *etorri*, konparazione, aditza den edo izena den? Hitz horrek baditu corpusean 13.344 agerraldi, eta ez dago modu automatiko fidagarriarik bi kategoria horien artean bereizteko. Badira moduak zeharkako bideak ibiltzeko: aukera daiteke *etorri* gehi izenondo bat (*handi*, konparazione), edo gehi *izan* aditza (*etorri da*), eta orduan mugatzen dugu hobeki lema. Baina informazio hori zeharkakoa da betiere. Badira bertze izenondo batzuk izen horrekin ageri direnak (*indartsu*, adibidez), eta bakarka ikusi beharko genituzte holakoak, denekin ezin baita (**etorri gogorra*) ibili. Beraz, muga horiek guztiak kontuan hartu behar dira.

Zerrendak emateko nolabaiteko hurrenkera egingen dut. Hasiko naiz kategoriaz, itxuraz bederen, «garbiak» diren hitzekin: iduri luke *etxe, andre*,

gizon edo *aulki*, izen garbiak direla. Eta *gidatu*, *pintatu* eta *atzendu* aditzak direla. Eta *polit*, *txiki* eta *gizen* adjektiboak. Kategorizazio hauek begien bistakoak direla ematen du, baina ez dira anitzetan kontuak hain argiak. *Atzendu* aditza da, aditz bezala ibili dute idazleek, baina horren parekoa den *ahaztu*, ordea, izen gisa ere ageri da, eta ez bakarrik aditz bezala: *bidegabekeria guztien ahaztu osoa egin diete*. Norik izanen luke zalantza txikiena *urdirin* hitzaren izenondotasunaz? *Zeruak urdin paregabea du alde horietan* erraten badut, ordea, izena da. Bai frantsesak bai ingelesek izenondo eta izen bezala sailkatzen dute hitz hau, baina RAEren sareko hiztegiak dio izendondoa dela edo, bestela, «izen gisa ibilitako izenondo». Ikusten denez, bereizketa hauen guztien artean argia aurkitzea ez da batere erraza.

Bestetik, hitzak ordenatzean garbiketa gehiago egin behar dira. Corpusean 377.795 hitz (desberdin) ditugula erran dugu. Baina hitz horietatik badira behin bakarrik ageri direnak, bost aldiz, hogeitaz edo ehunka eta milaka aldiz. Jakina: holako corpus handi batean hitz bat behin bakarrik ageri bada, pisu gutxi eta aski interes mugatua duen zerbaitez ari gara.

Gezurra ematen badu ere, badira 154.837 hitz behin bakarrik ageri direnak. Kontuan izanik corpusean 377.795 hitz desberdin ditugula, hitz horien %41 egiten du horrek. Gainerako hitzak dira behin baino gehiagotan ageri direnak. Gehienez bi aldiz ageri diren hitzak (behin bakarrik aipatuak, edo bi aldiz, baina ez gehiago) 207.753 dira, %55ean jartzen gara. Gehienez 5 aldiz ageri direnak 267.550 hitz dira, %71 hurbil. Ohart gaitezen datu hauekin, aski hitz gutxi ibiltzen ditugula praktikan gure artean mintzatzeko, edo literatura eta prentsan idazteko, corpusak 22,7 M hitz dituzenez, kopuru horretan ez baita deus hitz bat 5 aldiz ibiltzea, hitza euskara arruntean adierazgarria dela errateko gehiago ibili behar baikenuke. Hitzez ari gara, ez lemez: *du* eta *dute*, bi hitz dira, baina lema bakarra.

Segi dezagun kalkuluekin. Badakigu badirela hitz batzuk aski gutxi ibiltzen ditugunak, baina interesa dugu jakiteko zein diren benetan behin eta berriz ageri direnak: demagun jakin nahi dugula 100 aldiz edo gehiago errepikatzen diren hitzek corpuseko testuen ehuneko zenbat besarkatzen duten. Badira 18.278 hitz (hitzak dira, ez lemak) gutxienez 100 aldiz ageri direnak corpusean, hortik goiti betiere. Hitz horiek 20.129.577ko kopurua osatzen dute denen artean, corpusaren % 88,66 hartzen dute. Konparazioak argiago izateko: Euskaltzaindiaren hiztegi batuaren lehenbiziko edizioak (2000) 21.235 lema hartzen zituen. Bada, orain ari gara erraten gutienez 100 aldiz ageri diren hitzak kontuan hartzen baditugu 18.278ko hitz zerrenda ageri ditzagula. Baina zerrenda horretan daude *da*, *zuen*, *zen*, *du*, *izan*, *dira*, *dute*, *zuten*, *ziren*, *zituen*, *ditu*, *nuen*, *dugu*, eta abar. Horien guztien azpian bi lema bakarrik daude, betiere: *izan* eta *ukan* (horiekin batera sartu ditugu **edin* eta **ezan* aditzen adizkiak). Eta ez ditugu sartu horien *den*, *duen*, *zela*, *diren* edo *duten* bezalakoak. Horiek ere gehien ibiltzen ditugun hitzak, betiere aipatu bi lemen azpikoak. Euskara Batuko hiztegi hartan ageri ziren 21.235 lemen parean, gure corpusean gehien ibilienean

artean 18.278 hitz ageri badira, orotaran aise lema gutxiago dago, ez da erran beharrik.

3. Hitz ibilienak

Lehenbiziko 20 hitz ibilienak honako hauek dira (parentesian agerraldiak): *eta* (987.639), *ez* (442.503), *da* (285.225), *zuen* (256.273), *ere* (224.801), *bat* (215931), *zen* (193.496), *du* (177.861), *egin* (171.844), *izan* (170.178), *baina* (133.067), *bere* (107.416), *dira* (103.067), *dute* (89.509), *behar* (88.755) *beste* (79.872), *hori* (79.623), *esan* (78.585), *zuten* (76.821), *edo* (74.369). Hitz horien agerraldien kopurua 4.036.835 da, corpus osoaren %17,78.

Hitz horiek dira gure corpusean maizen ageri direnak. Ikusi, gainera, zer desberdintasunak dauden: zerrendako lehenbiziko hitza hogeigarrena baino hamahiru aldiz gehiago ibiltzen dugu. Hogeitaz hitz horietan 14 lema daude: *eta*, *ez*, *izan*, *ukan*, *ere*, *bat*, *egin*, *baina*, *bere*, *behar*, *beste*, *hori*, *esan* eta *edo*. Jakina: *zuen* hori *ukan* aditza izan daiteke (nagusia edo laguntzailea, **edun*) edo *zuek* izenordainaren genitiboa. Eta aditz beregaina izan daiteke (*goizean ikusi zuen*), edo mendekoa (*ikusi zuen neska gaztea zen*). Eta *bere* ez da lema berez, baizik *hura* izenordain indartuaren genitiboa. Baina euskal hiztegietan lema bezala ageri da, ez gara bereizketa horietan geldituko, ez dute garrantzirik hemen.

Hitz bat adierazgarria dela errateko, zenbat aldiz agertu behar luke corpusean? Zaila da horri erantzutea, baina ikus dezagun zer gertatzen den 1.000 aldiz baino gehiago errepikatzen diren hitzekin. Denetara 2.831 dira (lemak aise gutiago, jakina), eta hitz horiekin 15.672.870 hitzeko testu masa osatzen dugu. Bertze hitz batzuetan paratzeko: hitz horiek (hitzek, ez lemek) corpusaren %69,03 hartzen dute. Ikus dezagun zein diren hitz horiek. Aha! leginduko gara sailkaturik ematen, kategoriak bereiztea batzuetan erraza ez bada ere. Zernahi gisaz, kalkuluak, ondoko lerroetan, mailaka eginen ditugu, askoz ere gehiagotan ageri diren hitzetatik hasirik.

3.1. 20.000 aldiz baino gehiago ageri diren hitzak

115 hitz errepikatzen dira corpusean 20.000 aldiz baino gehiago. Denetara 66 lema. Zenbait kasutan eztabaidagarria izan daiteke kategorien sailkapen hori: *arte*, *konparazione*, izenen artean sartu dugu, baina postposizioa izan daiteke; *bada* eta *dena* ere sail bat baino gehiagokoak izan daitezke, guk aditzetan sartu baditugu ere; *zuen* genitiboa izan daiteke, edo aditza, eta guk aditzetan sailkatu dugu; *berriz* hitza, berez, adberbio gisa sailkatu dugu, baina *berri* lemaren azpian ere sartzen ahal zen (gure ustez, adberbializatua da *gutziz*); *bere*, neurri batean, *bera* lemaren genitibotzat har daiteke, baina

bereiz jarri dugu; *batzuk*, nahi izanez gero, *bat* leAREN pluraltzat har daiteke, baina nahiago izan dugu bereiz paratu; *beste* izenordainetan sartu dugu, EGLUK egiten duen gisan, eta abar. Holako batzuk gehiago badira. Bertzetik, *euskal* hitza izenen artean sartu dugu, izen elkarketetan *euskara* hitzaren aldaera delakoan. Erran dezagun, azkenekoz, ‘lokailuak’ deitu dugula atal bat, EGLUK erraten duena segituz, baina lokailuen kontu hori gramatika horretan aipatzen dena baino zabalagoa da: ‘alabaina’ bezalako hitz bat lokailua da. Baina, testuinguruaren arabera, *hasteko* edo *laburbilduz*, konparazione, birformulatzailer gisan ere erabiltzen ahal ditugu, eta hala sailkatzen dituzte hauen parekoak beste hizkuntza batzuetako gramatikek. Lokailuak, hortaz, oso zentzu zabalean hartzen ditugu hemen, lokailu horiek EGLUn aztertu ez badira ere. Denak diskurtso markatzaileak dira. Ohar hauek ondoko kuadro guztietan egin behar genituzke, baina hemen bakarrik paratuko ditugu. Beraz, honekin erran nahi dut hemen, oro har, egin dudan lemen kategorizazio sailkapena eztabaidagarria dela.

Beraz, corpusean 20.000 aldiz baino maizago ageri diren lemak hola sailkatu ditugu:

I. TAULA

I. Hitzzen agerraldiak: 20.000-987.639		
Hitz kopurua: 7.159.835		
Corpusaren %31,54		
Kategoriak	Atal honetako lemak: 65	Lemak denetara: 65
Adberbioak	16	atzo, berriz, beti, bezala, etzi, gaur, gehiago, gero, hain, hala, lehen, ondoren, orain, ordea, orduan, oso
Aditzak	11	ari, egin, egon, eman, esan, hartu, hasi, ikusi, izan, joan, ukan
Erakusleak	5	bera, bere, hau, hori, hura
Galdetzaileak	2	zer, nola
Izenak	11	alde, arte, behar, bizi, egun, euskal, gauza, hitz, lan, nahi, urte
Izenondoak	1	berri,
Izenordainak	2	beste, gu, ni
Juntagailuak	3	baina, eta, edo
Lokailuak	1	ere
Ordinalak	1	azken
Postposizioak	1	gabe
Partikulak	5	ahal, bai, baino, ez, ezin
Zenbatzaileak	6	asko, bat, batzuk, bi, guztiak, hiru

3.2. 10.000-19.999 tarteko agerraldiak dituzten hitzak

Hemen ere ez dira hitz gehiegi: 132 bakarrik. Denetara 102 lema, baina 16 lehendik errepikatzen direnez, corpusaren % 39,42a 151 lema osatzen dute (65+102-16).

2. TAULA

II. Hitz agerraldiak: 10.000-19.999		10.000-987.639
Hitz kopurua: 1.789.223		Hitz kopurua: 8.949.058
Corpusaren %7,88		Corpusaren %39,42
Kategoriak	Atal honetako lema: 102	Lema denetara: 151
Adberbioak	20 arabera, arren, aurka, begira, behin, berriro, bertan, besterik, bezain, hala, han, hemen, hor, horrela, inoiz, kontra, non, ondo, ongi, orain	
Aditzak	31 adierazi, agertu, atera, azaldu, bete, egin, egon, entzun, erabaki, erran, esan, eskatu, etorri, galdu, geratu, gertatu, hartu, heldu, hil, ikusi, irabazi, izan, jakin, jarri, jaso, jo, jokatu, lortu, sartu, sortu, ukan	
Erakusleak	4 bera, hau, hori, hura	
Galdetzaileak	2 zein, zergatik	
Izenak	24 aita, argi, arte, atsegin, aukera, aurre, behar, bide, bitarte, buru, denbora, egia, egoera, esku, etxe, gain, gisa, gizon, herri, inguru, lagun, lan, partida, talde	
Izenondoak	4 bakar, guzti, handi, on,	
Izenordainak	4 ezer, ni, zerbait, zu,	
Juntagailuak	2 baizik, nahiz	
Lokailuak	1 <i>aldiz</i>	
Ordinalak	1 bigarren	
Postposizioak		
Partikulak	4 al, ia, omen, ote	
Zenbatzaileak	5 bat, bost, hainbat, lau, zenbait	

3.3. 5.000-9.999 tarteko agerraldiak dituzten hitzak

Multzo honetan 308 hitz daude, eta 2.102.620 hitzeko testu masa hartzen dute. Hemen 220 lema daude, baina horietarik 48 lehengoak direnez, corpusaren %48,68a (ia-ia erdia) 323 lemak osatzen dute (151 + 220-48).

3. TAULA

III. Hitzzen agerraldiak: 5.000-9.999		5.000-987.639
Hitz kopurua: 2.102.620		Hitz kop: 11.051.678
Corpusaren % 9,26		Corpusaren % 48,68
Kategoriak	Atal honetako lemak: 220	Lemak denetara: 323
Adberbioak	22 agian, are, aski, askotan, aurka, batere, benetan, berehala, bereziki, beti, bihar, erabat, guztiz, han, horretarako, kontra, laster, lehen, noski, ostean, prest, sekula,	
Aditzak	61 abiatu, agertu, agindu, aipatu, aldatu, aritu, atera, aurkeztu, aurkitu, begiratu, bildu, bota, deitu, eduki, egon, ekarri, eman, erabili, eragin, erakutsi, esan, eskatu, espero izan, ezagutu, ezarri, galdetu, gelditu, gertatu, hartu, ibili, idatzi, iragan, iritsi, irudi izan, itzuli, izan, jakin, jakinarazi, jarraitu, jarri, jo, jokatu, kendu, maite izan, nabarmendu, nahi izan, onartu, pentsatu, salatu, sartu, sentitu, sortu, ukan, utzi, zabaldu, zain egon,	
Erakusleak	1 hau	
Galdetzailak	1 noiz	
Izenak	83 alde, alderdi, aldi, ama, aste, atze, aurre, balio, barne, barru, begi, behar, behe, bide, bizitza, buru, diru, egun, emakume, era, erdi, erredakzio, esker, estatu, etxe, euro, euskara, eusko, ezker, falta, gai, gain, garai, gau, gauza, giza, gizarte, gobernu, gogo, hasiera, haur, herri, hilabete, hiri, hitz, hizkuntza, hots, igande, indar, inguru, itxura, izen, jainko, jarrera, jaun, jean, jende, kanpo, kasu, kontu, laguntza, lan, lege, leku, liburu, maila, modu, mundu, musika, nagusi, nazioarte, neska, ondo, ondorio, ordu, pertsona, puntu, seme, talde, une, ur, urte, uste	
Izenondoak	16 azken, berdin, berri, eder, garbi, gazte, handi, hurren, luze, on, oso, preso, san, soil, txiki, zuzen	
Izenordainak	9 bertze, beste, edozein, eurak, gu, inor, ni, nor, zu,	
Juntagailuak	4 ala, baina, halaber, hots	
Lokailuak	4 ba, behintzat, bestalde, zeren	
Ordinalak	1 hirugarren,	
Postposizioak	2 bila, gabe	
Partikulak	3 baldin, ea, ohi	
Zenbatzaileak	13 bat, batzuk, bi, denak, gutxi, guztiak, hamar, mila, milioi, oro, sei, zazpi, zenbat	

3.4. 1.000-4.999 tarteko agerraldiak dituzten hitzak

Multzo honetan 2.276 hitz daude, eta 4.621.192 hitzeko testu masa hartzen dute. Denetara 1.124 lema daude atal honetan, baina horietarik 195 lehendik atera zaizkigunez, corpusaren % 69,03 (bi heren baino gehiago) 1.252 lemarekin osatzen dugu (323+1124-195). Honako hauek dira atal honi dagozkionak:

4. TAULA

IV. Hitzzen agerraldiak: 1.000-4.999		1.000-987.639
Hitz kopurua: 4.621.192		Hitz kop: 15.672.870
Corpusaren % 20,35		Corpusaren % 69,03
Kategoriak	Atal honetako lemak: 1124	Lemak denetara: 1.252
Adberbioak	90 aburuz, adina, aise, alderantziz, antzina, aparte, are, arras, arrunt, aspaldi, astiro, at, atzo, auskalo, azkar, batik bat, beharbada, behin, benetan, berandu, bestela, betiere, bezala, biziki, doi, emeki, erabat, etzi, ezinbeste, finen, gainezka, gaizki, gaur, gehi, gero, gorabehera, haatik, hala, han, hemen, herenegun, hola, honela, hor, horrela, hurbil, iaz, inola, inon, jada, jadanik, jarraian, kasik, kontuz, kosta ahala kosta, lasterka, lehenbailehen, lehenik, luzaz, maiz, noizbait, nolabait, nolanahi, nonbait, oihuka, oinez, orain, ordez, orobat, osoki, ostera, ozta, pixkanaka, poliki, samar, sarri, segidan, segur, sekula, sobera, urrun, urrutii, urtero, zeharo, zehazki, zintzilik, salbu	
Aditzak	238 abian jarri, abiatu, adi egon, adierazi, ados izan, agertu, agurtu, ahaztu, aintzat hartu, aipatu, aitortu, aldarrikatu, aldatu, altxatu, amaitu, antolatatu, arduratu, areagotu, argitaratu, argitu, ari, asmatu, atera, atxiki, atxilotu, atzeman, aukeratu, aurkeztu, aurreratu, axola izan, azaldu, azpimarratu, aztertu, babestu, baieztatu, baliatu, barkatu, begiratu, behatu, behartu, bereizi, berreskuratu, berretsi, bete, bidali, bihurtu, bilakatu, bilatu, bildu, bizi izan, bota, bukatu, bultzatu, bururatu, deitu, desagertu, edan, eduki, egin, egon, egokitu, egotzi, ekarri, ekin, elkartu, eman, entzun, erabaki, erabili, eragin, eraiki, erakutsi, eraman, erantsi, erori, erosi, erran, erreparatu, errepikatu, erre, esan, eseri, eskaini, eskatu, eskuratu, esnatu, estali, estutu, etorri, etzan, eutsi, existitu, ezarri, ezkondu, ezkutatu, finkatu, frogatu, gainditu, gaineratu, gaitu, gaitzetsi, galdu, garatu, garbitu, gauzatu, gelditu, gertatu, gogorazati, gogoratu, gorde, gustatu, harrapatu, harritu, hartu, hasi, hautatu, hautsi, hazi, heldu, hil, hobetu, hurbildu, hustu, ibili, idatzi, igaro, igo, igorri, ikasi, ikusi, inguratu, ipini, irabazi, iragarri, irakurri, iraun, ireki, iritsi, irten, irudi izan, iruditatu, isildu, itxaron izan, itxi, itzuli, izan, izendatu, jabetu, jaiki, jaitu, jaitu, jakin, jan, jantzi, jardun, jarraitu, jarri, jasan, jaso, jin, jo, joan, jokatu, kaleratu, kendu, kokatu, komeni izan, konpondu, kontatu, konturatu, kritikatatu, lagundu, landu, lehertu, leporatu, lortu, lotu, luzatu, makurtu, merezi izan, mintzatu, moldatu, moztu, mugitu, nagusitu, nahasi, ohartarazi, ohartu, onartu, ordaindu, oroitu, osatu, pasa(tu), pentsatu, piztu, plazaratu, posible izan, prestatu, proposatu, saiatu, sailkatu, salatu, saldu, sartu, segitu, sentitu, sinetsi, sortu, sumatu, suntsitu, topatu, topo egin, ukan, ukatu, ukitu, ulertu, utzi, zabaldu, zaindu, zauritu, zeharkatu, zehaztu, zendu, zutitu, zuzendu	
Erakusleak	7 beraiek, berbera, haiek, hura, hauek, hau, hori	
Galdetzaileak	7 ze, zer, zertan, zelan, nola, nondik, nora	

IV. Hitzen agerraldiak: 1.000-4.999		1.000-987.639
Hitz kopurua: 4.621.192		Hitz kop: 15.672.870
Corpusaren % 20,35		Corpusaren % 69,03
Kategoriak	Atal honetako lemak: 1124	Lemak denetara: 1.252
Izenak	626 abendu, abertzale, abiadura, abokatu, abuztu, adar, adibide, adierazpen, adin, adiskide, administrazio, afera, agentzia, agindu, agintari, agiri, agur, ahal, ahalegin, ahalmen, aho, ahots, aire, aita, aitor, aitzin, akordio, alaba, albo, aldaketa, alde, alderdi, aldi, amaiera, ama, amets, amore, anaia, andana, andre, animalia, antz, apaiz (apez), apiril, arau, arazo, ardura, argazki, argi, arima, arlo, arma, armada, arnasa, aro, arrakasta, arratsalde, arrazoi, arrega, arreta, arrisku, arropa, arte, artista, askatasun, asmo, aste, astearte, asteazken, asteburu, astelehen, atal, atari, ate, atsedeen, atsegin, atzerri, aukera, aulki, aurkari, aurkitu, aurpegi, aurre, auto, autonomia, auzi, auzitegi, auzo, azal, azaro, azpi, azterketa, babes, baimen, bake, baldintza, baliabide, baloi, banaketa, banatu, barne, barren, barre, barru, batasun, batu, batzar, batzorde, bazkari, bazter, begi, begirada, behar, behe, bekatu, belarri, beldur, berba, bertso, beso, besta, bezpera, bidaia, bide, biharamun, bihotz, bikote, bilduma, bilera, bilkura, biltzar, bisita, bista, bitarte, bizkar, bolada, bonba, borondate, borroka, botila, boto, boz, buelta, bukaera, buru, buruzagi, dantza, datu, defentsa, dei, demokrazia, denboraldi, denbora, denda, deus, diputatu, diru, disko, distira, dolar, egia, egile, egin, egitasmo, egitura, egoera, egoitza, egun, egunkari, eguraldi, eguzki, ekain, ekialde, ekimen, ekintza, ekitaldi, ekonomia, eliza, elkarriketa, elkartasun, elkarte, emakume, emazte, energia, enpresa, entrenatzaile, epaiketa, epaile, epe, erabilera, eragile, eraikin, eraikuntza, erakunde, erakusketa, erantzun, eraso, erdi, eredu, eremu, erlijio, errealitate, erreforma, errege, errepede, erresuma, eronka, ertzaintza, ertze, esaldi, eskaera, eske, esker, eskola, esku, eskualde, eskubide, eskuin, esparru, espazio, esperantza, esperientzia, espe-txe, estatu, estatutu, estrategia, etapa, eten, etengabe, etorkizun, etxe, etxebizitza, euri, euro, euskara, ezaugarri, ezker, ezkontza, eztabaida, familia, festa, film, filosofia, final, forma, foru, froga, funts, futbol, gai, gain, gaitasun, gaitz, galde, galdera, garai, garaipen, garapen, garrantzi, gatazka, gau, gauza, gehiengo, gela, gerla, gerra, gertaera, gertakari, gezur, gidari, giltza, giro, gisa, gizaki, gizarte, gizon, gobernu, gogo, gogoeta, goi, goiz, gol, goresmen, gorputz, gose, greba, grin, guardia, guraso, gustu, gutun, haize, hanka, hari, harreman, harrera, harri, haur, hautagai, hauteskunde, hegazti, hego, hegoalde, hein, helburu, helegite, herri, herritar, hezkuntza, hil, hilketa, hiri, historia, hitz, hitzarmen, hizketa, hizkuntza, hizpide, horma, hots, hotz, ibai, ibilbide, idazkari, idazle, ideia, iduri, igande, igoera, ihes, ikasketak, ikasle, ikerketa, ikuspegi, ildo, ile, indarkeria, indar, industria, informazio, inguru, interes, ipuin, irail, irakasle, irakurketa, iraultza, iritzi, irri, irribarre, irudi, isiltasun, istorio, itsaso, itun, iturri, itxaropen, itxura, itzal, itzulpen, izaera, izaki, izen, jabe, jai, jainko, jarduera, jatorri, jaun, jaur-laritza, jauzi, jende, joera, jokalar, joko, justizia, kafe, kale, kalte, kanpaina, kanpo, kanta, kantari, kantu, kapitulu, kargu, kartzela, kasu, kausa, kaze-tari, keinu, kezka, kide, kilo, kilometro, kirol, kolore, kolpe, konfiantza, konpromiso, kontratu, kontseilu, kontu, kontzertu, koentzientzia, kopa, korrika, kristau, kritika, kronika, kultura, kutxa, lagun, , laguntza, laguntzaile, lan, landa, langile, larunbat, lege, legebiltzar, lehendakari, lehia, lehio, leku, lepo, lerro, letra, liburu, liga, literatura, lore, lotsa, lotura, lur, mahai, maiatz, maila, maisu, maitasun, maite, makila, makina, manifestazio, marka, mar-txa, martxo, mediku, mende, mendebalde, mendi, merkataritza, merkatu,	

IV. Hitzen agerraldiak: 1.000-4.999		1.000-987.639
Hitz kopurua: 4.621.192		Hitz kop: 15.672.870
Corpusaren % 20,35		Corpusaren % 69,03
Kategoriak	Atal honetako lemak: 1124	Lemak denetara: 1.252
Izenak (<i>jarraitu</i>)	mesede, metro, meza, mezu, militar, min, ministro, mintzo, minutu, modu, molde, mota, muga, mugimendu, multzo, mundu, musikari, musu, mutiko, mutil, nagusi, nahi, nazio, negar, neke, neska, neurri, nobela, objektu, obra, odol, ogi, ohar, ohe, ohitura, oihu, oinarri, ondo, ondorio, ontzi, ordena, ordezkari, ordu, oren, orri, osaba, osasun, ostiral, ptoitz, otsail, paper, parada, pare, partida, pauso, pena, pentsamendu, pertsonaia, pertsona, pieza, pilota, pilotari, pisu, pixka, plan, plaza, poema, poesia, politika, polizia, porrot, postu, poz, premia, prentsa, presidente, preso, proba, programa, proiektu, proposamen, protesta, prozesu, puska, sail, saio, salaketa, sare, sari, sarrera, sasoi, segundo, segurtasun, seinale, seme, senar, sexu, sindikatu, sistema, so, soinu, soldadu, soldata, sor, sorta, su, taberna, talde, tarte, teknologia, telebista, telefono, teoria, testu, tiro, toki, tortura, tratu, tren, tresna, txanda, txapel, txapelketa, txartela, txosten, uda, udal, ume, unai, une, unibertsitate, ur, urrats, urre, urri, urtarri, urte, urtebete, usain, uste, uzta, xede, xehetasun, zakur, zalantza, zaldi, zama, zati, zehar, zehazki, zelai, zerbaki, zentzu, zer, zerbitzu, zerga, zerikusi, zerrenda, zero, , zientzia, zigor, zin, zinema, zori, zorion, zuhaitz, zulo, zuzendari, zuzendaritza	
Izenondoak	95 aditu, ageri, alfer, apal, apur, arduradun, aske, aurren, bakar, bakoitz, beltz, berdin, berezi, bero, berri, bikain, bitxi, bizkor, desgerdin, don, eder, egoki, erne, erraz, eskas, estu, gaixoeuskaldun, ezagun, ezberdin, frantses, garbi, garrantzitsu, gazte, geldi, gertu, gogor, gorri, goxo, guzti, handi, harrigarri, harro, haserre, hasiera, hatz, hobe, huts, ikaragarri, ikusgarri, ilun, indartsu, isil, izugarri, justu, labur, larri, lasai, libre, luze, nabarmen, normal, ohi, oker, on, orokor, oso, polit, politiko, polizia, publiko, sakon, santa, tipi santu, seguru, sendo, serio, soil, sozial, txapeldun, txar, txiki, urdin, urduri, xuxen, zabal, zahar, zail, zibil, ziur, zorionsu, zorrotz, zut, zuzen, ipar	
Izenordainak	8 beste, elkar, gu, hi, inor, ni, norbait, norbera,	
Juntagailuak	2 edota, ezta	
Lokailuak	3 alabaina, bederen, ezen	
Numeralak	2 bana, milaka	
Ordinalak	6 azken, bigarren, bosgarren, laugarren, lehen, lehenbiziko	
Postposizioak	4 baita, barik, gabe, gaindi	
Partikulak	13 ahal, ai, alta, baietz, ei, et, ez, ezik, ezin, harik, tira, eta abar	
Zenbatzaileak	23 anitz, asko, bat, batzuk, bederatz, berrogei, bi, bizpahiru, denak, ahun, ehunka, franko, gutxi, hainbeste, hamabi, hamabost, hamaika, hoge, nahiko, oro, ugari, zernahi, zortzi	

Ez da erran beharrik sailkapen hau egiteko orduan irizpide desberdinak ibiltzen ahal direla. HOLA, konparazione, *aburuz* adberbioen sailean sartu dut, nire ustez aski adberbializatua dagoelako: lema gisan 1.280 aldiz ageri da corpusean, eta horietatik «-en + aburuz» 1.145 aldiz, esamolde egina bezala;

azkar izenondoa da, baina gehien-gehinetan, gaur egun, adberbioaren erabilera du; holako zerbait gertatzen da *finean* hitzarekin ere: egia da *fin* izenondo bezala ibiltzen dugula (*mutiko fina da hori*), baina adibide gehienak ez dira izenondoarenak, baizik erramolde mota egin batenak, *finean* edo *azken finean*, izenetik abiatuturik; sail berekoak ditugu *jarraian*, *luzaz* edo *segidan*; *segur* ere sail horretan sartu dut (*hain segur*), baina hor daude *segur izan* edo *leku segurra* bezalakoak; *atsegin* lemak aditzean ere izan dezake tokiak (*atsegin dut*), edo izenondoetan (*atsegina* da), lehen aipatu dugu; *abian jarri* bildu dut, baina hor daude *abian izan*, *abian ezarri*, *abian egon*, *abian etorri*, edo *abian ekarri* bezalakoak; gauza bera *adi egon* aditzarekin (ez ditut bildu *adi begira*, *adi entzun*, *adi geratu*, *adi ibili*...); *aditu* izenondoetan sartu dut, baina aditza ere bada; *axola* izena ere bada, *berri* izena ere, *bizi* izena eta izenondoa, *jantzi* izena, *kosta ahala* *kosta* bai, baina baztertu dudana itsasoko *kosta* ere; *maite* izentzat hartu dut, baina izenondoetan ere, ordea, badu leku, *maitea* ageri baita; *arrunt* adberbioa da (*arrunt ona*), baina izenondoa ere izaten ahal da (*aski gizon arrunta* da), eta abar. Ororen gainetik, hala ere, garrantzia ez du kategoriak, baizik lemak, lemak ari gara bereizten, eta lema horiek kategoria bat baino gehiagotan agertzen ahal dira. Guk bat hautatu dugu, eta ibiliena biltzen saiatu gara.

Ordea, *eseri* eta *eserita* lema bakarrean sartu ditugu; gauza bera *gehi*, *gehiago*, *gehien*; *guzi* eta *guzti*; *hala* eta *halaxe*; *han* eta *hantxe*; *han* eta *hara*; *hau* eta *hauxe*; *hemen* eta *hona*, eta abar. Mugak ezartzea holako konfuetan arraz zaila da, eta kontraesanak berehala agertzen ahal dira: *hego* eta *hegoalde* sartu ditut; *herri* eta *herritar*, eta abar. Adizkitzat hartu ditut, menderagailuari kasurik egin gabe *den*, *dena*, *dela*, *zela*, *diren*, *delako*, *denak*, *dituela*, *dituzten duena*, *dugun*, *eginez*, *zuena*, *zuenez*, eta abar; lema beretik bestelako lemak eratortzen dira: *diru*, *dirudun*, *dirutu*; *eder*, *ederki*; *hasi*, *hasiera*; *ikasi*, *ikasle*, *ikasketa*; *indar*, *indartsu*; *lagun*, *lagundu*, *laguntza*, *laguntzaile*, eta abar. Guztiak beregaineko lematzat hartu ditut.

Lan honek segida behar luke izan bertze hainbertze puntutan: aipatu berri ditudan «famiak»; lema eta aldaeren arteko aukerak (batzuetan lema biluzik ez da ageri, beti dira aldaerak); gehien ageri diren adizkien marka sintaktikoak; kasu marken erabilera nagusiak, eta abar. Adizkiei buruz erranen dut zerbait. Hauexek dira maizen ibiltzen direnak, ibilienetatik gutiago ibiltzen direnetara (parentesi artean multzoak): *da*, *zuen*, *zen*, *du*, *dira* (100.000tik goiti); *dute*, *zuten*, *ziren*, *den*, *dela*, *ziona* (50.000-99.999); *zituen*, *ditu*, *duen*, *dut*, *nuen*, *dago*, *dugu*, *duela*, *zela*, *diren*, *dio* (30.000-49.999); *dituzte*, *dena*, *duten*, *zegoen*, *zituzten*, *nintzen*, *naiz* (20.000-29.999); *zuela*, *direla*, *duzu*, *dituen*, *zitzaion*, *ditela*, *daude*, *zioten*, *nion*, *gara*, *ziena*, *diote*, *genuen* (10.000-19.999); *dagoen*, *zirela*, *denak*, *ditugu*, *duk*, *zeuden*, *zaio*, *dituzten*, *badu*, *daiteke*, *ginen*, *duena*, *zara*, *dagoela*, *zutela*, *zitzaidan*, *ditut*, *nituen*, *baitzen*, *delako*, *dakit*, *balitz*, *zuenez*, *zait*, *diot*, *zenean*, *dituela*, *litzateke*, *dugun*, *die*, *zegoela*, *baitu*, *baitzuen*, *nago*, *zuena*, *luke*, *zituela*, *dit*, *bazen* (5.000-9.999); *dudan*, *zuenean*, *denean*, *nuke*, *ninduen*, *datorren*, *dauka*, *baitira*, *zena*, *dizut*, *direnak*, *dituztela*,

zizkion, diat, nuela, bazuen, dutenak, duenez, badute, dezake, balu, duelako, gaude, nengoen, zieten, denez, digu, dezagun, daudela, zaie, dadin, zelako, ziola (3.000-4.999); *zitekeen, dagokionez, nintzela, daitezke, zitzaizkion, dugula, baitute, zuelako, genituen, nau, diete, zenuen, zigun, zioen, badago, doa, dituzu, zaigu, diogu, duzun, daki, dik, daukat, dirudi, badugu, dezan, zaion, zirudien, zihoan, zituztela, nauzela, zaude, ditugun, duenean, baduzu, ezazu, zelarrik, dator, dutena, dizu, duenak, diola, zarete, zidaten, dabil, dezaten, diozu, datozen, denik, delarik, dutelako, daitekeen, badakizu* (2.000-2.999); *liteke, direnean, zaizu, zekien, gaitu, badirudi, dioten, dun, zutenean, daitekeela, dizkio, baditu, badaki, badut, dituk, balira, lukete, ezakeen, zirenak, genuke, zetorren, zutenez, dituena, zitziola, dezala, digute, ninduten, zaiola, bazekien, zezan, dutenez, badakigu, zizkioten, lirateke, didate, bazuten, dutenek, ditudan, daitezzen, dakigu, zenez, dituztenak, genion, litekeena, daitezkeen, duguna, daukagu, dezakeen, duenik, nintzenean, zizkidan, baliteke, zutenak, zeuzkan, zaitu, ezak, dezakete, goazen, dadila, nekien, zuelarik, ginela, diet, genuela, diegu, zitzaizgun, nian, bedi, dezakegu, dagokion, duzula, baitzituen, zihoazen, dagoena, lukeen, balute, dakizu, dagokienez, badaude, dagoelako, daukate, diagu, zutelako, zizkien, baitut, zian, badela, dituzue, dudana, ziguten, nauzu, balego, zarela, zebiltzan, gintuen, zutena, dizuet, dezatela, dabilta, zenuke, lukeela* (1.000-1.999). Aukera handiak eskaintzen ditu euskal aditz miragarriak baina, ikusten denez, 1.000 aldiz baino gehiago errepikatzen diren hitz zerrenda horretan ez da *geniezaioke* miserable bat ageri. Ez dugu ikusten *gaitzakezue* triste bat. Egia erran behar badut lehenbizikoa 18 aldiz ageri da corpus osoan, eta bigarrena 3 aldiz bakarrik... 205 milioi hitzeko corpusean (ETC). Sarasola et al (2013) ere kontsulta daiteke, nahi izanez gero.

Aipamenak

- Bosque, Ignacio. 2007. *Las categorías gramaticales: relaciones y diferencias*. Madrid: Síntesis.
- Michelena, Luis & Ibon Sarasola. 1987-2005. *Diccionario General Vasco / Orotariko Euskal Hiztegia*. Bilbao: Desclee de Brouwer-Ed. Mensajero.
- Sarasola, Ibon. 2010. *Egungo Euskararen Hiztegia* Bilbao: Euskara Institutua, UPV/EHU. [<http://www.ehu.es/eeh/>]
- Sarasola, Ibon, Pello Salaburu & Josu Landa. 2013. *Ereduzko Prosa Dinamikoa* Bilbao: Euskara Institutua, UPV/EHU. [<http://ehu.es/ehg/epd/>]
- RAE. 2013. *Diccionario de la Lengua Española*. Madrid. [<http://www.rae.es>]
- Landa, Josu, Ibon Sarasola & Pello Salaburu. 2011. *Euskal Hiztegiaren Maiztasun Egitura* Bilbao: Euskara Institutua, UPV/EHU. [<http://www.ehu.es/ehg/ehme/>]
- Sarasola, Ibon, Pello Salaburu, Josu Landa & Josu Zabaleta. 2009. *Ereduzko Prosa Gaur* Bilbao: Euskara Institutua, UPV/EHU. [<http://www.ehu.es/euskara-orria/euskara/ereduzkoa/>]
- Sarasola, Ibon, Pello Salaburu & Josu Landa. 2013. *Egungo Testuen Corpua* Bilbao: Euskara Institutua, UPV/EHU. [<http://www.ehu.es/etc/>]