

Hitzak sarean

Pello Salabururi esker onez

Itziar Laka (arg.)

eman ta zabal azu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

CIP. Unibertsitateko Biblioteka

Hitzak sarean [Recurso electrónico]: Pello Salabururi esker onez / Itziar Laka (koord.). – Datos. – Bilbao : Universidad del País Vasco / Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua = Servicio Editorial, [2020]. – 1 recurso en línea : PDF (170 p.)

Textos en euskara, inglés y español.

Incluye referencias bibliográficas.

Modo de acceso: World Wide Web

ISBN: 978-84-1319-111-9.

1. Euskara (Lengua) – Lexicografía. 2. Euskara (Lengua) – Gramática. I. Laka, Itziar, coord. II. Salaburu Etxeberria, Pello, homenajeado.

(0.034)809.169-3

(0.034)809.169-5

Atarikoa¹

Hona hemen, irakurle, Pello Salaburu maisu handiari eskaintzen diogun esker on liburua. Hamaika lan bildu ditugu, hainbat alorretakoak: euskararen lexikografia, terminologia, historia, gramatika, corpus ikerketa eta baita euskara eta gaztelaniaren prozesamenduari buruzkoak ere. Askotarikoa izan baita, zalantza gabe, Salaburuk utzi digun uzta oparoa; hizkuntzalaritzan, euskararen fonologian, gramatikan, lexikografian aitzindari izan dugu, gramatika sortzailearen esparruan lehenik, gero corpus eta hiztegi digitalak sortuaz, gure hizkuntza hobeto ezagutzeko, zaintzeko eta erabiltzeko behar ditugun XXI. mendeko lanabesak eraikiaz eta gizartean zabalduaz. Salabururen uzta ikerketaz eta akademiak harago zabaltzen da, gure gizarteak ezagutu dituen hainbat erronkatarara. Pello ez da inoiz uzkuritu gogoeta egin behar izan denean, bakegintzaren bideari dagokiola edo unibertsitatearen etorkizunari, beldur gabe agertu izan zaigu beti plazan, argi mintzatuaz, zuzen, hitzen kiribiletan sekula bere burua ezkutatu gabe.

Liburu honetan parte hartzen dugun guztiok baditugu motiboak soberan Salabururi eskerrak emateko, baina neuk hartu dut liburua argitaratzeko ardura, eta abagunea hartuko dut neure eskerrak emateko; zilegi bekit bada ñabardura pertsonalak ekartzea hona. Askotan azaldu dudanez, Pello Salaburu agertu zen sintaxiko irakasle karrerako (orain gradua deritzona) azken urtean eta han berpiztu zen nire sintaxi sortzailerako zaletasuna (Patxi Goenagak piztu zuena, bere *Gramatika Bideetan* liburuarekin). Salaburu MITeko hizkuntzalaritza sailean egonaldia eginda zetorren zuzenean gurera eta, nire zaletasuna ikusita, berak galdetu zidan ea gustatuko litzaidakeen Bostongo MITera tesia egitera joatea, amestera ere ausartzen ez nintzena. Ez dut gogoan hau inoiz idatzi izana, Salaburu maisu handiari zor diodan laguntza eskertzeko: nire etengabeko eskerrak, Pello, bidean erakutsi didazun laguntza guztiagatik, ez baita gutxi izan.

Hamaika lan bildu ditugu Pello Salabururen omenez: terminologiaren alorretik, lege hizkerari buruzko lan bi, bata EHUKo *Hiztegi Juridikoaren* sorrera eta eraketa azaltzen digu, Alberdi irakaslearen eskutik eta besteak terminologia juridikoak euskaraz izan duen ibilbide historikoa ezagutarazten digu, Azkarate eta Lopez Basaguren adituen eskuetatik. Hiztegi gintzaren alorrean, Lindemann eta San Vicentek Larramendiren neologismoei begiratzen diete egungo baliabide lexikoen ikuspegitik, eta Landak eta Sarasolak beharrezkoa dugun euskal hiztegi konbinatorial baterako lehen urratsak aurkezten dizkigute. Goenagak euskaldunok euskararen inguruan ditugun hainbat *uste* oker eztabaidatzen ditu eta Makazagak bere tesi lanak gogoan hartuta *benetako* euskara-

¹ Eskerrak eman nahi dizkiogu Eusko Jaurlaritzari ikerketarako emandako laguntzagatik (IT1169-19).

ren ikerketari buruzko gogoeta pertsonala egiten digu. Gramatikagintza garaikidearen isla, Berrok eta Fernandezek euskararen argumentu egituren gauzatze ezberdinak dakarzkigute, gurean oraindik oso ikertua ez den kontua. Egungo corpus hizkuntzalaritzaren erakusgarri, Pastorrek euskararen hitzorden kanonikoari buruz egindako ikerketa berria biltzen dugu lan sorta honetan. Azkenik, gramatikaren prozesamenduari buruzko hiru lan: lehena, Egusquiza eta Zawiszewski adituen eskutik, gaztelaniazko korreferentzia harremanen izaerari buruzkoa, Salaburuk hainbeste jorratu zuen *Binding Theory* haren ondorengo garaikidea; bigarrena, Andeetako gaztelaniak eta penintsulakoak perpausaren ordena kanoniko bera ala ezberdina ote duten galdetzen duena, Yozak, Erdociak eta Lakak elkarlanean burutua, eta Salaburuk Cuzcoko hiriarekin izan duen harreman bereziaren oroigarri; hirugarrena, Yetano, Duñabeitia eta Laka egile dituen, euskararen erlatibozko perpaus postnominalen prozesamenduaz, Salaburuk euskararen ikerketa modernoa bultzatzeko egin dituen ahalegin guztien oroigarri. Hiru hizkuntza biltzen dira liburu honetan bildu ditugun artikuluetan: euskara, gaztelania eta ingelesa. Izan ere, ez dugu ahaztu nahi Salaburuk noranahikoa goa egin duela gure hizkuntza txikia, inguruan diren hizkuntza handietan igeri eginaz, beste bideirik ez baitago izaten jarraituko badugu.

Bihoaz Lizardi poetaren hitzak zuregana, Pello, egin duzun lanaren laburbiltzerik egokiena baitira: *zuk, hizkuntza larreakoa nahi duzu ere noranahikoa, jakite-hegoek igoa, soina zahar, berri gogoa, azal orizta, muin betirakoa.*

ITZIAR LAKA

Lankide eta adiskide leialari

1

Urteak ongi beteta
iritsi zaik erretreta,
egindakoaren ordaina duk
ongi merezita.
Elkarri esku emanda
egonarren nekatuta,
eginen diagu atzera berriz
egin behar bada.

**Larai lailai, larai lailai
elkarri esku emanda.
Larai lailai, larai lailai
egin behar bada.**

2

Adiskide ta lankide
ez dituk baliokide.
Oraingo honetan, orde, biak
ditiagu hizpide.
Aspaldion gu lankide
eta betiko adiskide:
harro sentitzeko eran gaituk
horixe, hala fede.

**Larai lailai, larai lailai
lankide ohi ta adiskide
Larai lailai, larai lailai
Horixe, hala fede.**

3

Aurrekoaren oinatzak
Estaltzen dizkik lokatzak.
Askotan izaten ohi dituk
Zabu ta balantzak.
Haizearen noraezak
sortzen zizkiguk zalantzak.
Norabait joan nahi eta ezin
holaxe gabiltzak.

**Larai lailai, larai lailai
Zabu ta balantzak.
Larai lailai, larai lailai
Holaxe gabiltzak.**

4

Eskerrak hiri eman nahirik
norbaitek gu bildu gaitik.
E-Hatxe -U ta euskara zaudek
higan itsatsirik.
Hiru hizki ta zenbat hitz!
Guk ez behar gehiagorik.
Eskerrik asko, lagun jatorra,
hire lanagatik!

**Larai lailai, larai lailai
elkarri esku emanda.
Larai lailai, larai lailai
hire lanagatik**

Pello Salabururi eskainitako bertsoak (Bilbon, 2019-12-5)

2019ko abenduaren 5ean Bizkaia Aretoko de Rijk liburutegian Pello Salabururi egindako omenaldian parte hartu genuen adiskideak.



Argazkia: Mitxi. UPV/EHU

Atzeko lerroan: Maitena Etxebarria, Alberto Lopez Basaguren, Jesus Mari Makazaga, David Lindemann, Josu Landa, Patxi Goenaga.

Aurreko lerroan: Miren Azkarate, Xabier Alberdi, Beatriz Fernandez, Pello Salaburu, Itziar Laka.



Omenaldi egunean Patxi Goenagak Pellori opari eman zion akuarelaren erreproduzkoa

Aurkibidea / Contents / Índice

1. UPV/EHUko Zuzenbide Fakultatearen Hiztegi Juridikoa: hiztegi-sarreraren egitura eta definizioen idazkera <i>Xabier Alberdi Larizgoitia</i>	9
2. Lege-hizkera euskaraz: begiratu bat historian zeharrekoari eta zenbait gogoeta XXI. mendekoaz <i>Miren Azkarate, Alberto Lopez Basaguren</i>	24
3. (Euskal) Bozak: erdikoak, inkoatiboak eta arazleak <i>Ane Berro, Beatriz Fernández</i>	39
4. Coreference and antecedent frequency effects in Spanish: An eye-tracking study <i>Nerea Egusquiza, Kepa Erdozia, Adam Zawiszewski</i>	51
5. Eder baliz, on ez eiliz: Euskaldunok hizkuntzarekin ditugun maitasun/gorroto harremanez <i>Patxi Goenaga</i>	63
6. Baliabide lexikoen sarea: Baldintza filologiko eta tekniko zenbait <i>David Lindemann, Iñaki San Vicente</i>	79
7. Tesiko bosgarren kapituluaren historia txikia («Zer arraio idatzi dok zorioneko bosgarren kapitulu horretan?!») <i>Jesus Mari Makazaga</i>	97
8. Sobre el orden básico de palabras en euskera: un nuevo estudio de corpus <i>Luis Pastor Santamaría</i>	112
9. Euskararen konbinazio hiztegi baterako <i>Ibon Sarasola, Josu Landa</i>	130
10. Processing Preferences in an Ergative Language: Evidence from Basque Postnominal Relative Clauses <i>Iraia Yetano, Jon Andoni Duñabeitia, Itziar Laka</i>	137
11. Sentence processing in Andean and Peninsular Spanish: word order and animacy <i>Natalia Yoza, Kepa Erdocia, Itziar Laka</i>	151

UPV/EHUko Zuzenbide Fakultatearen Hiztegi Juridikoa¹: hiztegi-sarreraren egitura eta definizioen idazkera

XABIER ALBERDI LARIZGOITIA

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

Abstract

Artikulu honetan UPV/EHUko Zuzenbide Fakultatearen Hiztegi Juridikoaren nondik norakoak eta ezaugarriak aurkezten dira: hiztegiaren sorrera, terminologiari buruzko ikuspegia, metodologia, egitura, sarreraren egitura... Orobat aurkezten dira hiztegi honetan definizio argi eta ulergarriagoak idazteko baliatu diren estilo-irizpide batzuk (oro har, atzerakarga arintzeko eta definizio komunikatiboagoak egitekoak: nola eta noiz aurreratu deskriptorea, nola saihestu definizio zirkular, tautologiko eta parafrastiko hutsak, eta abar).

Hiztegiaren testuingurua: sorrera eta helburua

UPV/EHUko Zuzenbide Fakultateko Euskara Batzordearen ekimenez sortu zen euskarazko hiztegi juridikoaren proiektua, eta fakultateko irakasle elebidunen langintza eta lankidetzari esker gauzatu da orain dela urte batzuk egitasmo hutsa zena. Une egokian argitaratuko da lan hau; hain zuzen ere, Zuzenbide Fakultatearen 50. urteurrena ospatzen ari garen honetan. Hiztegi juridiko honek, mende Erdiko ibilbidean egindako beste lanek bezalaxe, agerian uzten du UPV/EHUko Zuzenbide Fakultateak euskara juridikoaren garapen- eta normalizazio-prozesua sustatzeko duen konpromiso irmoa eta helburu horri begira egiten duen lan etengabea. Ez gabilta oso oker esaten dugunean zuzenbideko irakasle euskaldunen hiru belaunaldi, behintzat, daudela ordezkaturik lan honen egile-taldean. Haietako batzuk aitzindariak izan dira dagokien alorrean euskaraz irakasten; beste batzuk, berriz, zuzenbideko ikasketak osorik euskaraz egin dituzten irakasle gazteak.

Urte asko joan dira UZEIk 1985ean lehen Zuzenbide Hiztegia argitaratu zuenetik, eta harrezkero ibilbide luzea egin da euskara juridikoa garatzeko. Aipagarriak dira eragile askoren lanak: besteak beste, Euskal Herriko administrazioen eta bereziki IVAPen lana, UPV/EHUK eta Deustuko Unibertsitateak egindakoa, Eusko Jaurlaritzako Justizia Sailak eta Terminologia Batzordeak egindakoa... Erabateko normalizaziotik urrun bagaude ere, ikaragarria da egin den aurrerapena;

¹ 2019an argitaratu du UPV/EHUK hiztegia (Alberdi, Azkarate eta beste, 2019). Eusko Jaurlaritzak finantzaturako IT1169-19 (2019-2021) ikerketa-proiektuaren emaitzetariko bat da lan hau.

horren lekuko, Zuzenbide Corpusak² jasotzen dituen emaitzak. Azken hamarkada hauetan gaikako edo arlokako hiztegi juridiko asko egin eta argitaratu dira, baina, hala eta guztiz ere, nabaria zen hiztegi juridiko orokor edo orotariko baten falta. Eta, hain zuzen ere, hutsune hori betetzera etorri da orain *UPV/EHUko Zuzenbide Fakultatearen Hiztegi Juridikoa*.

Euskal Herri osoan UPV/EHUko Zuzenbide Fakultatea da ikasketa guztiak euskaraz eskaintzen dituen eragile bakarra eta, dimentsio hori kontuan hartuta, ezinbestekotzat jo genuen bertako ikasle, irakasle eta langileek erabili ahal izango zuten hiztegi juridikoa sortzea. Funtsezko tresna da hiztegia, gure egunerokoan erabiltzen dugun terminologia juridikoari batasuna eta oinarri sendoagoa emateko. Gure komunitate juridikoan terminologia bateratua adosteko eta erabiltzeko aukera ahalbidetzen duen euskarri elebiduna da hiztegi juridiko hau. Irakaskuntzan, ikerketan zein kudeaketan erabiltzeko tresna aproposa. Fakultatea osatzen dugunon lanaren kanpo- eta barneremu guztietan erraz integratzen den baliabidea. Fakultatean egiten dugun lana ikusgai jartzen duena.

Hiztegiaren lehen erabiltzaileak, lana egiterakoan bereziki gogoan izan ditugunak, UPV/EHUko ikasle, irakasle eta langileak dira, baina ez dago esan beharrik arlo juridikoko eragile guztientzat pentsatuta dagoela hiztegia. Are gehiago, zuzenbideak gizarte-bizitzan eta hedabideetan duen eragin ikaragarria kontuan harturik, zilegi da esatea gizarte osoa dela hiztegi honen hartzailea.

Esan bezala, UPV/EHUko Zuzenbide Fakultateko irakasle askok lan egin dute hiztegian: 5 koordinatzailek, 15 taldeburuk eta 51 autorek. Beraz, erabatekoa izan da fakultateko irakasle elebidunen inplikazioa. Hiztegia arloka antolatu zen eta arlo bakoitzean taldeburu bat arduratu zen lana koordinatzeaz. Hamalau arlo hauek bereizi dira hiztegian:

Administrazio Zuzenbidea (*Admin. Zuz.*).
Europar Batasuneko Zuzenbidea (*EB. Zuz.*).
Finantza Zuzenbidea (*Fin. Zuz.*).
Konstituzio Zuzenbidea (*Konst. Zuz.*).
Kontzientzia Askatasunari buruzko Zuzenbidea (*Kontz. Ask. Zuz.*).
Lan Zuzenbidea eta Gizarte Segurantza (*Lan Zuz.*).
Merkataritza Zuzenbidea (*Merkat. Zuz.*).
Nazioarteko Zuzenbide Pribatua (*Nart. Zuz. Pr.*).
Nazioarteko Zuzenbide Publikoa (*Nart. Zuz. Pu.*).
Zigor Zuzenbidea (*Zig. Zuz.*).
Zuzenbide Erromatarra (*Zuz. Err.*).
Zuzenbide Prozesala (*Zuz. Proz.*).
Zuzenbide Zibila (*Zuz. Zib.*).
Zuzenbidearen Teoria (*Zuz. T.*).

Egileen artean, Zuzenbide Fakultatean hizkera eta komunikazio juridikoa irakasten eta jurilinguistikaren alorrean lan egiten duten hiru hizkuntzalari aritu dira proiektuan. Haiek arduratu dira hiztegiaren diseinuaz, terminologia erauzi eta lantzeko metodologiaz, definizioen idazketa-gidalerroez eta hizkuntza-irizpideez zein edizioaz.

² <http://www.ehu.es/ehg/zuzenbidea/>

Terminologiarekiko ikuspegia, definizio-moldeak eta izaera deskriptiboa

Hiztegitantzari buruzko ikuspegiari dagokionez, *Diccionario del español jurídico (DEJ)*³ izan dugu eredu eta gidari. Eta horrek esan nahi du hizkuntza-hiztegiak egiteko erabili ohi diren metodologia eta irizpide lexikografikoak erabiltzea: definizio entziklopediko luzeak eman beharrean, ahalik eta definizio laburren eta argienak ematen ahalegindu gara, argibide osagarriekin osatuz terminoari buruzko informazioa.

Gure ikuspegia, beraz, Terminologiaren Teoria Komunikatiboaren ildoan kokatzen da (Cabr , 1993), eta guztiz aldentzen da W ster-en (1979) Terminologiaren Teoria Orokorrean. Izan ere, W ster-en teoriaren arabera, kontzeptu espezializatuak (haien ustez, hizkuntzaren aurretikoak eta unibertsalak) ditu aztergai terminologiak, eta ez hitzak eta haien esanahiak; ez da onartzen polisemiarik eta sinonimiarik; hiztegi espezializatuetatik kanpo geratzen dira kategoria predikatiboak (aditzak, esaterako); definizio-bide bakarra onartzen da (deskriptore generikoa + ezaugarri bereizgarriak); kontzeptuen antolaketa hierarkiko eta zurrunda du oinarri terminologiak...

Oso bestelakoa da Terminologiaren Teoria Komunikatiboaren ikuspegia (Lorente, 2001). Teoria horren arabera, terminologiak balio espezializatuko unitate linguistikoak (BEUak) ditu aztergai eta hizkuntza naturala du oinarri. Eta horrek esan nahi du ez dagoela hitzen eta terminoen arteko bereizketa zurrun eta erabatekorik: izan ere, terminoak, azken batean, komunikazio espezializatuaren jardunean balio terminologikoa bereganatzen duten unitate linguistiko arruntak baino ez dira (hau da, komunikazio espezializatuan erabiltzen direnean aktibatzen da hitzen balio espezializatua). Hortaz, kontzeptuak ez dira existitzen izaki bakartu gisa, testu espezializatuari esker baizik. Ondorioz, izaera linguistikoa du terminoen esanahiak, eta hizkuntzaren beste edozein unitatetan bezain ohikoak dira sinonimia eta aldakortasuna terminoen (BEUen) alorrean.

Teoria komunikatiboen arabera jokatu dugu euskarazko hiztegi juridiko honetan definizioak (*definiens*) moldatzean: a) erabiltzailearen soslaiaren arabera egokitu ditugu definizioak (erabiltzaile ez-adituak ulertzeko moduko definizio labur eta albait argienak); b) sarrera polisemikoak ere jaso ditugu; c) askotariko estrategiak erabili ditugu definizioak idaztean (definizio inklusibo klasiko «aristotelikoa» —deskriptore generiko hiperonimikoa + ezaugarri bereizgarriak—, parafraasi bidezkoa, eta sinonimoen bidezkoa); d) aldakortasun kontzeptuala (sinonimia) jasotzen ahalegindu gara. Labur esanda, lexikografiaren eredutik gertu ibili gara terminoak definitzean: definizioaren helburua ez baita ahalik eta kontzeptuari buruzko xehetasun gehien ematea, baizik eta irakurleari balio espezializatuko unitate (BEU) horren esanahiaz jabetzeko adinako informazioa ematea; eta horregatik jo izan dugu definizio laburretara.

Terminologia juridikoa bateratzea eta finkatzea da hiztegi honen helburu nagusietako bat, baina, hala ere, horrek ez du esan nahi hiztegi arauemaitzat jo behar denik. Hain zuzen ere, geure egin dugun ikuspegi komunikatibo horretatik eratortzen da hiztegiaren izaera deskriptiboa: UPV/EHUko testuetatik elikatzen da hiztegi hau, eta, ezinbestean, Euskaltzaindiaren hiztegi batua du oinarri. Helburua, beraz, aldakortasuna onartzen duen eredu bat eskaintzea da eta ez hainbeste eredu zurrun bat inori ezartzea.

³ <https://dej.rae.es/>

Aldi berean, euskaraz komunikatzean segurtasun juridikoa eta herritarrek ulertzeko duten eskubidea⁴ bermatzen laguntzea da hiztegi honen beste helburu garrantzitsu bat. Argitalpen hau hizkera argiaren bidetik aurrera egiteko beste urrats erabakigarri bat izatea nahi genuke.

Lanerako metodologia

Terminologiari buruzko ikuspegi komunikatiboari jarraituz, UPV/EHUko Zuzenbide Fakultateko irakasleen testuetan oinarritu ginen euskarazko hiztegi juridiko hau egiteko: alegia, hasteko, irakasleen testugintza izan genuen oinarri eta hartatik erauzi genituen aukeraturiko arlo juridikoe-tako glosarioak osatzeko terminoak (UPV/EHUko «Terminologia Sareak Ehunduz» programaren bidez)⁵. Beraz, UPV/EHUko testu errealak eta testu haietatik erauzitako terminoak izan ziren hiztegiaren lehen iturria, barne-iturria, alegia.

Bestalde, UPV/EHUko ekoizpenarekin jarraituz, gure hiztegi juridikoan definitu beharreko terminoen zerrenda finkatzeko orduan oso kontuan hartu genuen, halaber, UPV/EHUk argitaratutako Legeak/Leyes⁶ bildumako terminologia.

Azkenik, kanpo-iturriak ere erabili genituen hiztegiko sarrerak finkatzeko: alde batetik, erdal hiztegi juridikoak⁷, Terminologia Batzordeak arlo juridikoan landutako glosarioak⁸, Euskalterm⁹, Deustuko Unibertsitateak argitaratutako berbategiak eta abar.

Behin terminoen erauzketa eginda eta barne- eta kanpo-iturriak kontsultatuta, definitu beharreko terminoen zerrenda —hiztegiaren oinarria— finkatu zen koordinatzaileen artean eta taldeburuen laguntzaz: ezinbestean, termino ugari utzi behar izan ziren kanpoan, argitalpena arrazoizko luzera-neurri batera ekartzearren. Handik aurrera, arloz arlo, definizioak idazteari eta sarrerak osatzeari ekin zioten hamalau arlo juridikoetako taldeburu eta autoreek.

Bestalde, hasieratik eta bukatu arte, hizkuntza-arduradunek arreta handiz gainbegiratu eta zuzendu dute erredakzio-lana: hasieran, definizioak idazteko estilo-liburu baten bidez, eta prozesuaren erdian eta amaieran testuak eta moldiztegiko probak zuzenduta.

⁴ Ikus lan hauek, besteak beste:

Herritarrek Justiziaren aurrean dituzten eskubideen Gutuna (2002): https://www.justizia.eus/servlet/Satellite?blobcol=urldat&blobheader=application%2Fpdf&blobheadname1=Content-Disposition&blobheadvalue1=filename%3Ddocumentos-303-carta-derechos-EUSK_0.pdf&blobkey=id&blobtable=MungoBlobs&blobwhere=1290510252341&ssbinary=true&miVar=1524343183302

Hizkera Juridikoa Modernizatze Batzordearen txostena: <http://www.upv.es/entidades/VRSC/info/U0711342.pdf>

Ainhua Altzuguren-en «Hizkera juridiko argia» web-orria: <https://ainhoaaltzuguren.wixsite.com/hizkerajuridikoargia>

⁵ Zuzenbide Fakultateko irakasle elebidun ia guztiek parte hartu zuten UPV/EHUko «Terminologia Sareak Ehunduz (TSE)» programan: <https://www.ehu.es/eu/web/euskara/terminologia-sareak-ehunduz>

⁶ <https://www.ehu.es/es/web/euskara/legeak-leyes-bilduma>

⁷ Funtsezkoa den *Diccionario del español jurídico (DEJ)* hiztegiak aparte (<https://dej.rae.es/>), honako hauek, besteak beste: Arco Torres, M.A.; Villa-Real, R. (2006): *Diccionario de términos jurídicos*. Granada: Comares.

Fernández Martínez, J. M. (2012): *Diccionario jurídico*. Pamplona: Aranzadi.

Fonseca Herrero, J.I.; Iglesias Sánchez, M.J. (2010): *Diccionario jurídico*. Madrid: Colex.

Fundación Tomás Moro (2004): *Diccionario jurídico Espasa*. Madrid: Espasa.

Gómez de Liaño, F. (1979): *Diccionario jurídico*. Oviedo: Forum.

Ortiz Sánchez, M.; Pérez Pino, V. (2012): *Diccionario jurídico básico*. Tecnos.

Villa-Real, R.; del Arco Torres, M.A. (2014): *Diccionario jurídico elemental*. Granada: Comares, 7. argitaraldia.

⁸ <http://www.euskadi.eus/hiztegiak-banku-terminologikoak-eta-entziklopediak/web01-a2eutres/eu/>

⁹ <http://www.euskadi.net/euskalterm>

Hiztegiaren egitura

Atal batzuez osatuta dago euskarazko hiztegi juridiko hau, baina, euskarazko sarrerak alfabetikoki ordenatuak dituen atala —definizioz eta oharrez hornitua— da, zalantzarik gabe, lanaren muina. Izan ere, lehenengo eta behin, erabiltzaile euskalduna izan dugu gogoan hiztegia osatzera-koan.

Hala ere, mesede gutxi egingo genion geure buruari eta gizarteari, euskara hutsezko hiztegi juridiko bat egin izan bagenu. Izan ere, askotan azpimarratu izan denez, onenean mundu juridikoa-ren errealitatea elebiduna da gure artean: alegia, ezinbestean, gaztelania alboan edo gogoan duela ibili beharra dauka jurista euskaldunak.

Beraz, errealitate elebidunak baldintzatu du hiztegiaren egitura. Hasieran, eranskin elebidun batzuk jarri ditugu, erabiltzaileari lana erraztuko diotelakoan:

- Zuzenbidearen arloak edo adarrak, eta haien laburdurak.
- Siglak: auzitegien eta arauen izenak eta siglak.
- Erakundeen izenak.
- Epaitegi eta auzitegien izenak.
- Arau-xedapenak (maila guztietakoak: Nazioarteko Zuzenbide Publikokoak, Zuzenbide Pri-batukoak, Estatukoak, Euskal Autonomia Erkidegokoak, Nafarroako Foru Erkidegokoak eta lurralde historikoenak).
- Zergen izenak.

Gero, atal nagusia dator: euskarazko sarrera definiziodunen atala, hurrenkera alfabetikoa-ordenatua. Oro har, laburrak dira sarrera gehienak, nahita baztertu baitugu hiztegi entziklopedikoe-tako artikuluen estiloa: funtsezko argibideak ematen dira definizioan, eta gero argibide osagarriak (ikus beherago sarreraren egiturari buruzko atala, gainerako xehetasunen berri jakiteko). Guztira, si-nonimoak kontatuta, 3.500 termino inguru ditu hiztegiak.

Eta, azkenik, termino guztien zerrenda elebidun hutsa dator: gaztelania-euskara zerrenda de-finiziorik gabea. Esan bezala, atal hau ere funtsezkoa izango da hiztegiaren erabiltzailearentzat.

Hiztegiaren mikroegitura: hiztegi-sarreraren egitura

Sarrera lexikografiko batek honelako egitura izan ohi du:

definigaia	+	kategoria lexikoa	+	definizioa
(definitu beharreko hitza, ‘definiendum’)		(izena, izenondoa, aditza, adberbioa...)		(definiens)

Hiztegi juridiko honetan —*Diccionario del español jurídico* hiztegian eta Euskaltermen be-zala¹⁰— ez da zehazten kategoria, baina horrek ez du esan nahi kategoria desberdinak jaso ez di-renik. Izan ere, Terminologiaren Teoria Orokorren arabera lan terminografikoetan ez bezala,

¹⁰ Cercaterm terminologia-bankuan (<https://www.termcat.cat/ca/cercaterm>), berriz, kategoria lexikoa zehazten dute: adibi-dez,

— *ca* **acusar**, v tr

— *es* **acusar**, v tr

<Dret penal i penitenciari > Dret processal penal>

kategoria predikatibo ugari jaso ditugu hiztegian (batez ere, aditzak, hala nola *aldarrikatu*, *akusatu*, *alegatu*, *artxibatu*, *aurkaratu*, *aurrejuzgatu*, *baliozkotu*, *derogatu*, *deuseztatu*, *errebokatu*, *eskualdatu*, *epatu*, *errefusatu*, *errekusatu*, *espetxeratu*, *heredatu*, *inkulpatu*, *jarauntsi*, *kalifikatu*, *mugarritzatu*, *ordezkatu*, *tatxatu*, *xedatu*, *zedarritzatu*...).

Sarrera bi hauek ditugu jokaera horren erakusgarri (lehen terminoa izena da; bigarrena, berriaz, aditza):

akusatu 1 [*acusada*, *acusado*] ♦ *Zuz. Proz.* Zigor-prozesuko beharrezko alderdi pasiboa, akusazio-idazkiaren bitartez gauzatutako akusazioa pairatzen duena. Oh. PKL, 760 III, 368., 371., 373., 448., 512., 687. art. **Ik.** *alderdi akusatu*.

akusatu 2 [*acusar*] ♦ *Zuz. Proz.* Zigor-akzioa egikaritu pertsona zehatz baten kontra.

Beraz, egia bada ere sarrera terminologiko gehienak izenak edo izen-kategoriakoak direla (*ahalmen*, *alderdi*, *aldarrikapen*, *aplikazio-eremu*...; *jabetza pribatu*, *itxurazko kontratu*...), me-rezi du azpimarratzea arlo juridikoko aditz espezializatuen garrantzia.

Bestalde, hizkuntza-hiztegi orokorretan ez bezala —eta egungo terminografia-lanetan egin ohi denez—, askotariko «balio espezializatuko unitateak» (BEUak) jaso dira hiztegi juridiko honetan sarrera gisa, eta ez bakarrik izen-kategoriakoak edota kategoria lexiko soilak (Estopá, 2001):

BEU lexikoak: bakunak / konplexuak (eratorriak, elkartuak...):

- izenak (*abokazio*, *abokatu*...) / (*ahozkotasun*, *agerraldi*...)
- aditzak (*alegatu*, *kereilatu*, *zedarritzatu*...) / (*aurrejuzgatu*, *eskualdatu*...)
- ...

BEU ez-lexikoak:

- izen-sintagmak:

Europako Banku Zentrala, *ahozko bista*, *ahozkotasun-printzipio*, *akats judizial*, *akusazio-sistema formal*, *alderdien galdeketa*, *demandaren prebentziozko oharpen*...

- aditz-sintagmak:

sorospen-betebeharra ez betetze, *kanpora-bidaltze*, *atzera-itzultze*, *bidegabeko kaleratze*, *zor ez diren sarrerak itzultze*...

Behin kategoriaren kontua argituta, ikus ditzagun sarrera batzuk, hiztegi honetarako hautatu dugun mikroegituraren nondik norakoak azaltzeko:

afektazio [*afectación*] ♦ *Admin. Zuz.1.* Eskubideak zein ondasunak interes orokorreko erabilera publiko batera zuzentzea. Normalean nahitaezko desjabetze-prozeduretan gauzatzen da. **Oh.** 1954ko Nahitaezko Desjabetze Legea, 9. art. **2.** Ondasun edo eskubideek jabari publikoko izaera eskuratzeko beharrezko den egintza. Haren bidez ondasun edo eskubidea erabilera orokorrari edo zerbitzu publikoari atxikitzen zaio. **Oh.** 33/2003 Legea, 65. art. **Sin.** *xedemate Ik. jabari publiko, zerbitzu publiko*.

arrazoitze [*motivación*] ♦ *Zuz. T.* Ebazpen jakin bat oinarritzen duten arrazoien adierazpena. Ordenamendu juridikoak hainbat ebazpen edo egintza juridiko motibatzea eskatzen du, hala nola

ebazpen judizialak eta Europar Batasuneko egintza juridikoak. **Oh.** 6/1985 Lege Organikoa, 247. art. eta 248. art.; Europar Batasuneko Funtzionamendu Hitzarmena, 296. art. **Sin. motibazio** ♦ *Admin. Zuz.* Administrazio-egintza funtsatzen duten egitateen eta zuzenbideko oinarrien adierazpena. Administrazioak bere jardunean izan behar duen gardentasunak eta administrazio onaren eskubideak administrazio-egintza motibatzea eskatzen dute. Horrez gain, administrazio-erabakien kontrola ere errazten du. 39/2015 Legearen 35. artikulua zehazten du zein administrazio-egintzatan den nahitaezko arrazoitzea. **Oh.** 39/2015 Legea, 35 art. **Sin. motibazio** **Ik.** *administrazio-egintza*.

Aurreko adibideetan ikus daitekeenez, lehen-lehenik euskarazko terminoa (lema) ageri da, letra lodiz; ondoren, gaztelaniazko ordaina (letra etzanean eta parentesi karratuen artean): *afektazio* [*afectación*].

Sarreretan, ♦ ikurra erabili da adierak zuzenbidearen adarren arabera bereizteko: goragoko lehen adibidean (*afektazio*), Administrazio Zuzenbidekoak (*Admin. Zuz.*) dira terminoaren bi adierak (1. «Eskubideak zein ondasunak...». 2. «Ondasun edo eskubideek...»). Bigarren adibidean (*arrazoitze*), berriz, Zuzenbidearen Teoriakoa (*Zuz. T.*) da lehen adiera («Ebazpen jakin bat...»), eta Administrazio Zuzenbidekoa (*Adm. Zuz.*), bigarren adiera («Administrazio-egintza...»).

Sarreretan, bestalde, zenbakiak (1, 2, 3...) erabili dira zuzenbidearen adar baten barruan termino batek adiera bat baino gehiago ditutenean: esate baterako, *afektazio* sarreran, Administrazio Zuzenbidekoak dira terminoaren bi adierak (1. «Eskubideak zein ondasunak...». 2. «Ondasun edo eskubideek...»). *Arrazoitze* sarreran, berriz, ez da zenbakirik erabili; izan ere, bi adierak zuzenbidearen adarraren arabera (♦ *Zuz. T.* / ♦ *Admin. Zuz.*) bereizten dira.

Formatuari dagokionez, letra lodi etzanez eman ditugu latinezko eta erdal terminoak: *ius cogens*, *ius praetorium*, *iuris tantum*, *iuris tantum* presuntzio...

Era berean, letra etzanezko laburduren bidez adierazi dugu zuzenbidearen arloa edo adarra (hiztegiaren hasieran ageri da laburduren zerrenda): *Admin. Zuz.* (Administrazio Zuzenbidea), *Err. Zuz.* (Erromatar Zuzenbidea), *Nart. Zuz. Pu.* (Nazioarteko Zuzenbide Publiko) eta abar.

Sarrera askotan oharrak gehitu dira, gehienbat lege-esparruaren erreferentzia zehatza emateko; alegia, delako terminoa jasotzen, aipatzen edo definitzen duten legeen artikuluetara bidaltzen da irakurlea, nahi izanez gero, argibide zehatzagoak eskuratzeko. Horretarako, «Oh.» laburdura erabili da. *Arrazoitze* terminoaren lehen definizioaren ostean, esaterako, honako xehetasun hauek ematen dira: «Oh. 6/1985 Lege Organikoa, 247. art. eta 248. art.; Europar Batasuneko Funtzionamendu Hitzarmena, 296. art.». Erabaki horretan ere, *Diccionario del español jurídico (DEJ)*¹¹ hartu dugu eredutzat; izan ere, sinetsita gaude erreferentzia osagarri horiek guztiz aberasgarri eta lagungarriak izango direla irakurlearentzat.

Lexikografian, ezinbestean, ohikoa izaten da hitz batzuk sinonimiaren bidez «definitzea». Eta hiztegi juridiko honetan ere, sinonimo ugari ageri dira. Arlo espezializatueta oro har, eta arlo juridikoan zer esanik ez, eguneroko gertaera ezin saihestuzkoa da aldakortasuna, eta aldakortasun horren isla da gure hiztegian sinonimo ugari jaso izana. Horretarako, «Sin.» laburdura erabili dugu sarreraren amaieran; *arrazoitze* terminoaren kasuan: «Sin. *motibazio*».

Bestalde, termino batzuetan «Ik.» laburdura gehitu da sarreraren amaieran, modu laburrean kontzeptuaren inguruko argibide gehiago emateko edota definigaiarekin (*definiendum* delakoare-

¹¹ <https://dej.rae.es/>

kin) lotura duten beste termino batzuk ikusteko. *Arrazoitze* terminoaren kasuan, adibidez: «Ik. *administrazio-egintza*».

Batzuetan, hala ere, sinonimoen berri emateko erabili da «Ik.» laburdura: lehenesten den terminoaren ostean jaso dira definizioa eta sinonimoa (ikus *arrazoitze* sarrera); beste hitz sinonimoen sarreran, berriz, termino lehenetsira bidaltzen da irakurlea, definiziorik eman gabe (Ikus *motibazio* sarrera). *Arrazoitze* eta *motibazio* terminoen kasuan, honela:

arrazoitze [*motivación*] ♦ Zuz. T. [DEFINIZIOA: «♦ Ebazpen jakin bat oinarritzen duten arrazoi adierazpena(...)»] Sin. *motibazio*; ♦ Admin. Zuz. [DEFINIZIOA: «»Administrazio-egintza funtsatzen duten egitateen eta zuzenbideko oinarrien adierazpena. (...)] Sin. *motibazio*
motibazio [*motivación*] ♦ Zuz. T., Admin. Zuz. Ik. *arrazoitze*.

Sarreraren mikroegiturari buruzko atal hau amaitzeko, argitu dezagun zer definizio-eredu hautatu dugun hiztegi honetarako. Jakina denez, definizioa funtsezko osagaia da espezialitate-hiztegi bateko sarreretan. Bada, gorago azaldu dugunez eta aurreko adibideetatik ondoriozta daitekeenez, oro har, definizio laburraren bidetik jo izan dugu hiztegiaren zehar, lexikografiako eredua geure eginez eta definizio entziklopedikoak bazterturik. Gehienbat, hiztegiaren erabiltzaile ez-aditua izan dugu gogoan: izan ere, helburua ez da kontzeptuaren inguruko xehetasun guztiak jasotzea, baizik eta azken batean irakurlea kontzeptu juridiko bakoitzaren muinaz jabetzea.

Ohiko definizioen egitura: definizio-eredu soilen aldeko apustua

Hiztegi-sarrerek aurreko atalean zehaztutako irizpideak zainduko badituzte, terminoen definizioek laburrak eta zehatzak izan behar dute ezinbestean. Edonola ere, gogora ekarri behar da hiztegi juridiko honen helburua ez dela kontzeptu juridikoen adieran eta ñabarduretan sakontzea. Izan ere, definizioen helburua beste bat da; alegia, euskara/gaztelania interfazeaz harantzago joan eta erabiltzaileei euskarazko diskurtso juridikoaz jabetzen lagunduko dien balio erantsiko baliabide bat eskaintzea.

Jakina denez, legelari euskaldun askok gaztelaniazko diskurtsoa dute nagusiki gogoan euskarak bizi duen egoera diglosikoaren eraginez. Haientzat oso baliagarria izango da terminoen definizio erraz eta ulergarriak izatea eskura. Esan gabe doa biziki eskertuko dituztela definizio hauek (xumeak izan arren) ikasketak euskaraz egitea aukeratu duten ikasleek.

Horrenbestez, gure hiztegi juridikoan jaso diren definizioek ez dituzte gehienetan jaso termino juridikoak bere barruan gordetzen dituen adiera-elementu guztiak. Adituentzat jakingarri diren ñabardurak ere albo batera utzi ditugu ezinbestean. Eta, noski, termino batek teoria juridikoaren arabera definizio desberdinak izan baditu, aukera bat egin da ahalik eta definizio orokor eta neutroena eskaintzeko. Betiere, gogoan izan da zer den ikasleek termino horri buruz jakin behar dutena, lehen aldiz edo, kontzeptu juridiko horretara hurbiltzen direnean.

Mikroegituraren ikuspegitik ere ahalik eta definizio soil eta laburrenak hautatu dira. Oro har, bi edo hiru osagai izaten dituzte (ik. Arrieta eta beste, 2003): a) deskriptorea (zein kontzeptu-kategoria definitzen ari garen adierazten du; gehienetan hiperonimo baten bidez adierazia); b) ezaugarri espezifikoak (definitzen ari garen kontzeptua kategoria bereko beste kontzeptuetatik bereizteko balio dutenak —ik. Seco, 1987—); eta c) iruzkinak (oharrak edo ñabardurak eransteko balio dutenak).

Kontuak kontu, oinarritzko definizioari informazio osagarria gehitu behar bazaio, aparteko esaldi batean gehitu dira halako argibide, azalpen edo deskripzio osagarriak. Hona hemen hiru osagaiak dituzten definizio batzuk:

EGILEKIDETZA [COAUTORÍA] ♦ *Zig. Zuz.* Delituan parte hartzeko modu bat [deskriptorea], delituaren egile nagusi bat baino gehiago eta erantzukizun-maila berean aritzean datzana [ezaugarriak]. Egilekide guztiek zigor bera jasotzen dute [iruzkina]. (...)

AKUSATZAILE PRIVATU [ACUSADORA PRIVADA, ACUSADORPRIVADO] ♦ *Zuz. Proz.* Zigor-prozesuko alderdi akusatzailea [deskriptorea], delitu pribatu batek ofenditua izateagatik kereila aurkeztu duena [ezaugarria]. Delitu pribatuak jazartzeko prozesuetan bera da zigor-akzioaren titular bakarra eta, horregatik, akusatzaile bakarra [iruzkina].

Definizio batzuetan lehen bi elementuak baino ez dira erabili, eta definizio konplexuetan ezaugarri bat baino gehiago eta iruzkin bat baino gehiago. Alabaina, gure hiztegi juridikoak dituen ezaugarriak eta helburuak gogoan izanda, ahal den neurrian, definizio-eredu laburrak eta soilak lehenetsi ditugu.

Hizkuntzaren aldetik ere egitura soil, argi eta ulergarriak erabiltzen ahalegindu gara (ik. Arrieta eta beste, 2003: 172-187)¹². Hurrengo atalean azalduko ditugu definizioak osatzeko erabili diren estilo-irizpide zehatzak.

Definizioargi eta ulergarri(ago)ak idazteko estilo-irizpide batzuk

Oro har, hiztegiaren helburuaren eta hartzaile motaren arabera egokitu ditugu definizioen konplexutasuna, teknizismo-maila eta idazkera.

Hitz edo termino ezagunak erabili ditugu definizioetan. Edozein kasutan, definizio batean termino teknikoak erabiltzen badira, terminook beste nonbait egon behar dute definituta (alegia, hiztegian bertan behar dute definizioa).

Definizio gehienetan, hiperonimo bat —ahalik eta zehatzena— erabili da deskriptore gisa (definizio formal edo hiperonimikoa —ik. Bosque, 1982—):

AKUSATZAILE [ACUSADOR, ACUSADORA] ♦ *Zuz. Proz.* Zigor-prozesuko posizio aktiboko *alderdia* [hiperonimoa], zigor-akzioa egikaritzuz akusazioa gauzatzen duena. (...)

KARRERA JUDIZIAL [CARRERA JUDICIAL] ♦ *Zuz. Proz.* Magistratura osatzen duen *personal juridikozionala* [hiperonimoa].(...)

KATASTRO [CATASTRO] ♦ *Zuz. Zib.* Hazienda Ministerioan kokaturiko ondasun higiezinaren *erregistro administratiboa* [hiperonimoa].

Definizioetan isilpean utzi dugu predikatua (*da, deritzo, deitzen zaio...*), hiztegitantz jakintzat ematen baita «X (definigaita) Y (definizioa) da» predikazioa (Rey-Debove, 1971):

↓ LURRALDE HISTORIKO [TERRITORIO HISTÓRICO] ♦ *Konst. Zuz.* Foru-lurraldeek Euskal Autonomia Erkidegoan *duten izena*.

LURRALDE HISTORIKO [TERRITORIO HISTÓRICO] ♦ *Konst. Zuz.* Euskal Autonomia Erkidegoko foru-lurraldea.

¹² URLa: http://www.ivap.euskadi.eus/contenidos/informacion/euskalduntze_argit/eu_def/adjuntos/Galdezka1.pdf

↓EGILESLE [OTORGANTE] ♦Zuz. Zib. Negozio juridiko bati bere adostasuna ematen dion pertsona *da*. [...]

EGILESLE [OTORGANTE] ♦Zuz. Zib. Negozio juridiko bati bere adostasuna ematen dion pertsona *da*. [...]

Oro har, saihestu egin ditugu honelako hitz eta egiturak deskriptore gisa (Arrieta, 2003): (... *izendatzeko erabiltzen den izena*, *-t(z)earen egintza*, *-t(z)eko egintza*, (...) *zera*, (...) *duena*...

↓ABOKAZIO [AVOCACIÓN] ♦Admin. Zuz. Administrazio-organo batek hierarkikoki beheago da-
goen beste organo batek daukan gai bat ezagutu edo ebazteko eskumena *bereganatzearen egintza /*
bereganatzeko egintza. (...)

ABOKAZIO [AVOCACIÓN] ♦Admin. Zuz. Administrazio-organo batek hierarkikoki beheago da-
goen beste organo batek daukan gai bat ezagutu edo ebazteko eskumena *bereganatzea*. (...)

Definizio laburretan arazorik gabe eman da deskriptorea (ezaugarria adierazten duen) erlati-
bozko egituraren ostean (egitura sintetikoa: ezaugarriak + deskriptorea):

KARRERA JUDIZIAL [CARRERA JUDICIAL] ♦Zuz. Proz. Magistratura osatzen duen *pertsonal juris-
dikzionala* [hiperonimoa]. (...)

ALDABERRITZE-GREBA [HUELGA NOVATORIA] ♦Lan. Zuz. Hitzarmen kolektiboaren edukia alda-
tzeko egiten den *greba* [deskriptorea].

Hala ere, definizio luzeetan deskriptorea aurreratu dugu modu sistematikoan:

↓AKUSAZIO-IDAZKI [ESCRITO DE ACUSACIÓN] ♦Zuz. Proz. Ahozko epaiketaren hasiera eskatzeko
eta pertsona akusatua, delituzko egitateak, haien kalifikazio juridikoa, egiletasun-maila eta ingu-
ruabar aringarri edo astungarriak eta ezarri beharreko zigorra zehazteko, prozedura laburtuan Mi-
nisterio Fiskalak eta gainerako alderdi akusatzaileek aurkeztutako *dokumentua*. (...)

AKUSAZIO-IDAZKI [ESCRITO DE ACUSACIÓN] ♦Zuz. Proz. Prozedura laburtuan Ministerio Fiskalak
eta gainerako alderdi akusatzaileek aurkeztutako *dokumentua*; *haren helburua da* ahozko epaike-
taren hasiera eskatzea eta pertsona akusatua, delituzko egitateak, haien kalifikazio juridikoa, egile-
tasun-maila eta inguruabar aringarri edo astungarriak eta ezarri beharreko zigorra zehaztea. (...)

Idazle onenen tradizioari jarraituz (Alberdi, 2007 eta 2008), «deskriptorea + egitura aposatua
(erlatibozko -na/partizipioa)» erabili da askotan atzerakarga saihestearren (egitura analitikoa: des-
kriptorea + ezaugarriak):

AKUSATZAILE [ACUSADOR, ACUSADORA] ♦Zuz. Proz. Zigor-prozesuko posizio aktiboko alderdia
[deskriptorea], zigor-akzioa egikaritzat akusazioa gauzatzen duena. (...)

FUNDAZIO [FUNDACIÓN] ♦Zuz. Zib. Irabazi asmorik gabeko nortasun juridikoa duen interes pu-
blikoko erakundea [deskriptorea], interes orokorreko helburua lortzeko sortua izan dena. (...)

HITZARMEN [ACUERDO, CONVENIO] ♦Konst. Zuz. 1. Akordio, itun edo kontratua [deskriptorea],
kontsentsu bidez sortua eta ondorio juridikoak dituen.

INSTRUKZIO-EPAILE [JUEZ DE INSTRUCCIÓN] ♦Zuz. Proz. Instrukzio-epaitegiko titularra [deskrip-
torea], atariko prozedura edo instrukzio-fasearen ardura duena.

KARTZELA-ZIGOR IRAUNKOR BERRIKUSGARRI [PEN DE PRISIÓN PERMANENTE REVISABLE] ♦Zig.
Zuz. Delitu larri batzuentzat aurreikusten den zigor askatasun-gabetzailea [deskriptorea], eskepen-
baldintza eta -epe bereziak dituen. (...)

AUZITEGI [TRIBUNAL] ♦Zuz. Proz., Admin. Zuz., Konst. Zuz., Zuz. T. Kide anitzeko juris-
dikzio-organoa [deskriptorea], hiru magistratuz edo gehiagoz osatua.

ELKARGO OFIZIAL [COLEGIO OFICIAL] ♦ Admin.Zuz. 1. Titulazio bereko pertsonen osatzen duten zuzenbide publikoko erakundea [deskriptorea], titulazio horri lotutako lanbidea arautu, kideen interesak babestu eta ordezkatzeko xedez eratua.

Definizio terminografikoak idazteko irizpideen arabera (ISO 704, 18. or.), definizioaren muina —kontzeptu bat gainerakoetatik bereizten duena— esaldi bakar batean ematea komeni da (puntutik puntura doana da esaldia eta sarritan perpaus batzuez osatuta egoten da). Baina horrek ez du esana nahi ezaugarri guzti-guztiak esaldi bakar batean eman behar direnik. Aitzitik, kasu horietan gomendatzen da informazio osagarria (argibide, azalpen edo deskripzio osagarria) aparteko esaldien bidez ematea eta, beraz, -na aposatuen erabilera gehiegizkoa saihestea.

↓HORNİKUNTZA-ERREGIMENĒKO EPAILE [JUEZ DE RÉGIMEN DE PROVISIÓN] ♦ Zuz. Proz. Lehiaketa-tan hutsik geratu den epaile-hutsunea behin-behinean betetzen duen epailea [definizioaren muina], *prozedura arrunten bidez hutsunea bete arte jardunean ari dena* [argibide osagarria].

HORNİKUNTZA-ERREGIMENĒKO EPAILE [JUEZ DE RÉGIMEN DE PROVISIÓN] ♦ Zuz. Proz. Lehiaketa-tan hutsik geratu den epaile-hutsunea behin-behinean betetzen duen epailea [definizioaren muina]. *Prozedura arrunten bidez hutsunea bete arte aritzen da jardunean* [argibide osagarria].

Bestalde, baditugu baliabide estrategiko batzuk definizioetako erlatibo aposatuen erabilera gehiegizkoa murrizteko: besteak beste, alborakuntza (puntuazioa) eta juntadura. Hona hemen adibide batzuk:

IKUSKARITZA [AUDITORÍA] ♦ Merkat. Zuz. Dokumentu kontableen (eta horien euskarri diren beste dokumentu batzuen) berrikuste- eta egiaztatze-lana. *Ikuskaritzak hirugarren pertsonen aurrean eragina duen txostena igortzea du helburu* [alborakuntza]. Txosten hori oinarri hartzen da enpresaren informazio ekonomiko finantzarioaren fidagarritasunari buruzko irizpenak egiteko. (...)

IMMUNITATE [INMUNIDAD] ♦ Konst. Zuz. (...) 2. [immunitate diplomatiko] Atzerriko diploma-zialariei eta haien senitartekoei, bai eta batzuetan enbaxadetako langileei aplikatzen zaien pribilegioa; *harrera-lurraldearen jurisdikziotik at daudela esan nahi du* [alborakuntza]. (...)

KASAZIO-ERREKURTSO [RECURSO DE CASACIÓN] ♦ Zuz. Proz., Admin. Zuz. Aparteko errekurtsu debolutiboa. *Legeak aurreikusitako zio jakin batzuetan oinarrituta soilik jar daiteke* [alborakuntza]. Ezagutzeko eskumena Auzitegi Gorenari edo Justizia Auzitegi Nagusiari dagokio, esleitutako arauaren izaeraren arabera. (...)

JAİOTZA-İNSKRIPTION [INSCRIPCIÓN DE NACIMIENTO] ♦ Zuz. Zib. Erregistro Zibilean egiten den idazpena, umea noiz (zein egunetan eta ordutan) eta non jaio den adierazten duena, *eta bere sexua eta seme-alabatasuna zein den ezagutzera ematen duena* [juntadura].

↓AKUSAZIO-IDAZKI [ESCRITO DE ACUSACIÓN] ♦ Zuz. Proz. Ahozko epaiketaren hasiera eskatzeko eta pertsona akusatua, delituzko egitateak, haien kalifikazio juridikoa, egiletasun-maila eta inguruabar aringarri edo astungarriak eta ezarri beharreko zigorra zehazteko, prozedura laburtuan Ministerio Fiskalak eta gainerako alderdi akusatzaileek aurkeztutako *dokumentua*. (...)

AKUSAZIO-IDAZKI [ESCRITO DE ACUSACIÓN] ♦ Zuz. Proz. Prozedura laburtuan Ministerio Fiskalak eta gainerako alderdi akusatzaileek aurkeztutako *dokumentua*; *haren helburua da ahozko epaiketaren hasiera eskatea eta pertsona akusatua, delituzko egitateak, haien kalifikazio juridikoa, egiletasun-maila eta inguruabar aringarri edo astungarriak eta ezarri beharreko zigorra zehaztea*. (...)

Definizio funtzionalen kasuan ere joskera bat baino gehiago erabili ditugu:

- a) (...) -*t(z)eko* + Y (deskriptorea)
- b) (...) -*t(z)ea helburu duen Y* (deskriptorea)

- c) *Y* (deskriptorea), (...) -*t(z)ea helburu duena*
- d) *Y* (deskriptorea), (...) -*t(z)eko helburua duena*
- e) (...) -*t(z)eko helburua duen Y* (deskriptorea)

ABSTENTZIOA [ABSTENCIÓN] ♦ Zuz. Proz. Epaile eta magistratuen inpartzialtasuna *bermatzeko bitartekoa*. Horren bidez, legeak aurreikusitako arrazoiak daudenean, epaileak zein magistratuak auzi bat ezagutzeari uzteko eskaera egiten du.

HERRI-SEGURTASUN [SEGURIDAD CIUDADANA, SEGURIDAD PÚBLICA] ♦ Admin. Zuz. Pertsonak eta euren ondasunak babestea, eta herritarren lasaitasuna *bermatzea helburu duen jarduera multzoa*. (...)

JARAUNSPENAREN ZAPUZTE [REPUDIACIÓN] ♦ Zuz. Zib. Jaraunspena errefusatzeko deituaren *alde-bakarreko egintza*, deituak jaraunspena ez duela onartzen espresuki *adieratzea helburu duena*. (...)

ERREKERIMENDU [REQUERIMIENTO] ♦ Zuz. Proz., Admin. Zuz. Jakinarazpen prozesalaren *jarduketa*, prozesuaren alde bati edo hirugarren bati zuzendua eta ez-egiteren bat edo jokabideren bat egin dezala *agintzeko helburua duena*.

FIDANTZA [FIANZA] ♦ Zuz. Zib. Zordunak hartzekodunari zorra ordainduko diola *ziurtatzeko helburua duen berme pertsonala*. Fidantzaren bidez, fidatzaileak, hartzekodunaren baimenarekin, bere gain hartzen du zorraren ordainketa zordunak ordaintzen ez duen kasuetan.

Oro har, deskriptorea asko atzeratzen duten egiturak saihesten ahalegindu gara; honelakoak, esate baterako:

↓ «(sintagma luzea) -*t(z)ea helburu duen* + deskriptorea»

↓ «(...) -*t(z)ea*, (...) -*t(z)ea eta* (...) -*t(z)ea... helburu duen* + deskriptorea»

↓ IRUZUR FISKAL [FRAUDE FISCAL] ♦ Fin. Zuz. Zergapekoak burutako operazio erreala, nolabait teko amarrubidea adierazten duena, *bilatzen diren azken efektuak lortzeko, ohikotasunean erabiliko ez liratekeen negozio eta arauak aplikatzea helburu duena*. Eta hori guztia, zerga-zorra txikitze asmoz. (...)

IRUZUR FISKAL [FRAUDE FISCAL] ♦ Fin. Zuz. Zergapekoak burutako operazio erreala, nolabait teko amarrubidea adierazten duena. *Haren helburua da, bilatzen diren azken efektuak lortzeko, ohikotasunean erabiliko ez liratekeen negozioeta arauak aplikatzea*. Eta hori guztia, zerga-zorra txikitze asmoz. (...)

Erlatibozko perpausaren barruko osagaiak handik kanpora ateratzea saihestu dugu:

↓ EGINBEHAR PUBLIKO [DEBER PÚBLICO] ♦ Admin. Zuz. 1. Arduradun publikoek zein bestelako pertsonak bete beharrekoa, dela akzioz dela omisioz, legediak ezarrita zein agintariak hala aginduta, *interes publikoa babestearren*. (...)

EGINBEHAR PUBLIKO [DEBER PÚBLICO] ♦ Admin. Zuz. 1. *Interes publikoa babestearren*, legeak ezarrita zein agintariak hala aginduta, arduradun publikoek zein bestelako pertsonak bete beharrekoa, dela akzioz dela omisioz. (...)

↓ EGINBEHAR PUBLIKO [DEBER PÚBLICO] ♦ Admin. Zuz. (...) 2. Zenbait pertsonari ezartzen zaizkien betebeharrak *interes publikorako onuragarriak izateagatik*. (...)

EGINBEHAR PUBLIKO [DEBER PÚBLICO] ♦ Admin. Zuz. (...) 2. *Interes publikorako onuragarriak izanik*, zenbait pertsonari ezartzen zaizkien betebeharrak. (...)

↓ EGINBEHAR PUBLIKO [DEBER PÚBLICO] ♦ Admin. Zuz. (...) 3. Administrazioak egin beharreko ekintzak, *ordenamendu juridikoa betetzen dela bermatzearen*. (...)

EGINBEHAR PUBLIKO [DEBER PÚBLICO] ♦ Admin. Zuz. (...) 3. *Ordenamendu juridikoa betetzen dela bermatzearen* administrazioak egin beharreko ekintzak. (...)

↓IMMATRIKULAZIO [INMATRICULACIÓN] ♦Zuz. Zib. Jabetza Erregistroan finka bat lehen aldiz sartzeko egin behar den eragiketa, *beharrezkoa den kalifikazio erregistrala gainditu ondoren* (...)

IMMATRIKULAZIO [INMATRICULACIÓN] ♦Zuz. Zib. Jabetza Erregistroan finka bat lehen aldiz sartzeko egin behar den eragiketa, *beharrezkoa den kalifikazio erregistrala gainditu ondoren egiten dena*. (...)

Ez da komeni «erlatibo kateatuak» erabiltzea definizio batean, definizioa ulergaitz bihurtzen da eta:

↓MATXINSALTO. Jauzi egiteko moldatuta *dauden* hanka luze eta sendoak *dituzten* ortopteroen ordenako *intsektua* [deskriptorea].

MATXINSALTO. Ortopteroen ordenako *intsektua* [deskriptorea]; hanka luze eta sendoak ditu, jauzi egiteko moldatuak.

Ez da komeni definizio tautologikoak ematea, jakina. Alegia, definizioak ez dira osatu behar terminoa osatzen duten hitzak errepikatuz edo parafrasi lausoak —terminotik beretik ondoriozta daitekeen informazioa ezertan osatzen ez dutenak— erantsiz.

↓ AKUSAZIO-IDAZKI. Akusazioak jasotzen duen idazkia

AKUSAZIO-IDAZKI [ESCRITO DE ACUSACIÓN] ♦Zuz. Proz. Prozedura laburtuan Ministerio Fiskalak eta gainerako alderdi akusatzaileek aurkeztutako *dokumentua*; haren helburua da ahozko epaiketa-ren hasiera eskatzea eta pertsona akusatua, delituzko egitateak, haien kalifikazio juridikoa, egileta-sun-maila eta inguruabar aringarri edo astungarriak eta ezarri beharreko zigorra zehaztea. (...)

Horrek ez du esan nahi terminoaren osagaiak definizioan ezin ager daitezkeenik: terminoa-ren muina batzuetan terminoaren generikoa izaten da, eta, deskriptorea denez, definizioaren oina-rra izaten da.

AUZITEGI GORENA. Erakunde judizialaren gailurrean dagoen eta Estatu osoan jurisdikzioa duen *auzitegia*. [Euskalterm: Zuzenbidea / Zuzenbide Prozesala]

KASAZIO-ERREKURTSO [RECURSO DE CASACIÓN] ♦Zuz. Proz., Admin. Zuz. Aparteko *errekurtsa* debolutiboa. Legeak aurreikusitako zio jakin batzuetan oinarrituta soilik jar daiteke. Ezagutzeko eskumena Auzitegi Gorenari edo JustiziaAuzitegi Nagusiari dagokio, esleitutako arauaren izaera-ren arabera.

Ez da komeni definizio zirkularrak ematea:

a. ↓ ERREKUSAZIO. Errekusatzea

b. ↓ ERREKUSATU. Errekusazioa eskatzea

ERREKUSATZE [RECUSACIÓN] ♦Admin. Zuz. Administrazio-prozedura baten izapidetzearen edo-zein unetan interesdunek langile edo agintari publiko baten arbuiatzea edo errefusatzea galdatzeko duten eskubidea, gai bat ebatz ez dezan edo arazo batean esku har ez dezan bere inpartzialtasuna zalantzan jar dezaketen arrazoiak daudelako tartean (40/2015 Legeko 23.2. artikuluan zerrenda-tzen direnak, alegia). Oh. 40/2015 Legea, 24. art. Ik. abstentzio.

ERREKUSATU [RECUSAR] ♦Zuz. Proz. Ik. *errefusatu*.

ERREFUSATU [RECUSAR] ♦Zuz. Proz. Modu legitimoan epaimahai, epaile, peritu edo denadelako batek prozedura edo epaiketa batean jardutea galarazi. Sin. *errekusatu*.

Definigaiaren eta deskriptorearen gramatika-baliokidetasunaren baldintza zaindu behar da definizioan (esaterako, kategoriari eta numeroari dagokienez):

↓INTERPRETATU [INTERPRETAR] ♦Zuz. T. 1. Zerbaiti esanahi zehatza *atxiki edo eranstea*. 2. Zerbaiten esanahia *argitzea edo azaltzea*, bereziki argi ez dagoen testua. Legeak, adibidez.

INTERPRETATU [INTERPRETAR] ♦Zuz. T. 1. Zerbaiti esanahi zehatza *atxiki edo erantsi*. 2. Zerbaiten esanahia *argitu edo azaldu*, bereziki argi ez dagoen testua. Legeak, adibidez.

Definizioetan, beharra eta abar adierazteko egokiagoa da, oro har, orainaldi burutugabea erabiltzea geroaldia baino:

KUDEAKETA PROZESAL ETA ADMINISTRATIBOKO KIDEGO [CUERPO DE GESTIÓN PROCESAL Y ADMINISTRATIVA] ♦Zuz. Proz. Justizia-administrazioaren zerbitzura dauden funtzionarioen kidego orokorra: goi-mailako titulua *eskatzen da* (? *eskatuko da*) kidegoan sartzeko eta goi-mailako jarduera prozesalean laguntzen dute.

Azken oharra hiztegi juridikoaren garrantziaz eta zabalkundeaz

Espainiako Hizkera Juridikoa Modernizatzeko Batzordeak bere txosten orokorrean dioenez¹³, zuzenbide-estatu batean beharrezkoa da ekintza eta arau juridikoak ulergarriak izatea herritarrentzat, eta, beraz, zuzenbideko eragile juridikoek erantzukizun berezia dute arlo horretan; eta are erantzukizun handiagoa dute erakundeek herritarren ulertzeko eskubidea bermatzeko orduan.

UPV/EHUK, hiztegi hau argitaratuta, pauso garrantzitsu bat ematen du herritarren ulertzeko eskubidea bermatzeko ahaleginean: izan ere, hiztegi hau guztiz lagungarria izan daiteke zuzenbideko ikasle eta profesional euskaldunen komunikazio-prestakuntzan aurrera egiteko.

Orain UPV/EHUK euskarazko hiztegi juridiko hau liburu-formatuan eta euskarri digitalan argitaratu du; baina onuragarria litzateke unibertsitateak —Euskara Institutuaren laguntzaz, agian— gizarte osoarentzat etorkizun hurbilean hiztegia Internet bidez zabalduko balu, RAEk *Diccionario del español jurídico* hiztegiarekin egin duen moduan. Hala izango ahal da!

Erreferentzia bibliografikoak

Alberdi, X. (2007): «Zenbait gogoeta euskal definizioen idazkeraz» in Iñaki Ugarteburu, Pello Salaburu: *Espezialitateko Hizkerak eta Terminologia II. Euskara estandarra eta espezialitate hizkerak*, Euskal Herriko Unibertsitateko Argitalpen Zerbitzua: Bilbo, 145-170.

Alberdi, X. (2008): «Definizioen idazkera: adierazmoldeak eta diskurtso-estrategiak azalpenezko testu espezializatueta» in Iñaki Ugarteburu, Pello Salaburu: *Espezialitateko Hizkerak eta Terminologia III*, Bilbo: Euskal Herriko Unibertsitateko Argitalpen Zerbitzua, 79-93.

Alberdi, X.; Azkarate, M. eta beste (2019): *UPV/EHUKo Zuzenbide Fakultatearen hiztegi juridikoa*. Bilbo: UPV/EHUKo Argitalpen Zerbitzua.

Arco Torres, M.A.; Villa-Real, R. (2006): *Diccionario de términos jurídicos*. Granada: Comares.

Arrieta, A. eta beste (2003): *GALDEZKA. Euskarazko zalantzei erantzunez*. Donostia: IVAP: http://www.ivap.euskadi.eus/contenidos/informacion/euskalduntze_argit/eu_def/adjuntos/Galdezka1.pdf

Bosque, I. (1982): «Sobre la teoría de la definición lexicográfica», *VERBA*, 9: 105-123.

Cabré, M.T. (1993): *La terminología: teoría, metodología, aplicaciones*. Barcelona: Antártida/Empúries.

¹³ <http://www.upv.es/entidades/VRSC/info/U0711342.pdf>

- Estopá, R. (2001) «Elementos lingüísticos de las unidades terminológicas para su extracción automática». In Ma T. Cabré & J. Feliu (ed.) *La terminología científico- técnica*. Barcelona: Institut Universitari de Lingüística Aplicada: 67-80.
- Fernández Martínez, J.M. (2012): *Diccionario jurídico*. Pamplona: Aranzadi.
- Fonseca Herrero, J.I.; Iglesias Sánchez, M.J. (2010): *Diccionario jurídico*. Madrid: Colex.
- Fundación Tomás Moro (2004): *Diccionario jurídico Espasa*. Madrid: Espasa.
- Gómez de Liaño, F. (1979): *Diccionario jurídico*. Oviedo: Forum.
- Guerrero Ramos, G.; Pérez Lagos, M.F. (2017): «La definición en el diccionario desde la teoría lingüística». *Pragmalingüística*, 25, 286-310.
- Lorente, M. (2001): «Teoría e innovación en terminografía: la definición terminográfica», Cabré, M.T. y Feliu, J. (eds.): *La terminología científico-técnica: Reconocimiento, análisis y extracción de información formal y semántica* (DGES PB96-0293), Barcelona: Institut Universitari de Lingüística Aplicada / Universitat Pompeu Fabra, 81-112.
- Ortiz Sánchez, M.; Pérez Pino, V. (2012): *Diccionario jurídico básico*. Tecnos.
- Rey-Debove, J. (1971): *Étude linguistique et sémiotique des dictionnaires français contemporains*. The Hague-Paris: Mouton.
- SECO, M. (1987): «Problemas formales de la definición lexicográfica», *Estudios ofrecidos a Emilio Alarcos Llorach* (con motivo de sus XXV años de docencia en la Universidad de Oviedo), II., Oviedo: Universidad de Oviedo, 217-239.
- Villa-Real, R.; del Arco Torres, M.A. (2014): *Diccionario jurídico elemental*. Granada: Comares, 7. argitaraldia.
- Wüster, E. (1998), *Introducción a la Teoría general terminológica y a la lexicografía terminológica*. Barcelona: IULA (1979koa da jatorrizkoa).

Interneteko erreferentziak

- Ainhoa Altzuguren-en «Hizkera juridiko argia» web-orria: <https://ainhoaaltzuguren.wixsite.com/hizkerajuridikoargia>
- Cercaterm: <https://www.termcat.cat/ca/cercaterm>
- Diccionario del español jurídico: <https://dej.rae.es/>
- Euskalterm: <http://www.euskadi.net/euskalterm>
- Herritarrek Justiziaren aurrean dituzten eskubideen Gutuna (2002): https://www.justizia.eus/servlet/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadername1=Content-Disposition&blobheadervalue1=filename%3Ddocumentos-303-carta-derechos-EUSK_0.pdf&blobkey=id&blobtable=MungoBlobs&blobwhere=1290510252341&ssbinary=true&miVar=1524343183302
- Hizkera Juridikoa Modernizatzeko Batzordearen txostena (Comisión para la Modernización del Lenguaje Jurídico): <http://www.upv.es/entidades/VRSC/info/U0711342.pdf>
- Terminologia Batzordeak onartutako hiztegiak: <http://www.euskadi.eus/hiztegiak-banku-terminologikoak-eta-entziklopediak/web01-a2eutres/eu/#9588>
- UPV/EHUko «Legeak/Leyes» bilduma: <https://www.ehu.eus/es/web/euskara/legeak-leyes-bilduma>
- UPV/EHUko «Terminologia Sareak Ehunduz (TSE)» programa: <https://www.ehu.eus/eu/web/euskara/terminologia-sareak-ehunduz>
- Zuzenbide Corpusa: <http://www.ehu.es/ehg/zuzenbidea/>

Lege-hizkera euskaraz: begiratu bat historian zeharrekoari eta zenbait gogoeta XXI. mendekoaz

MIREN AZKARATE¹, ALBERTO LOPEZ BASAGUREN²

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

Laburpena

Artikulu honetan, izenburuak berak dioen bezala, zenbait gogoeta plazaratu nahi ditugu euskarazko lege-hizkeraren gainean. Hasi XVII. mendeko testu xumeetatik eta gaurko corpus berezi, terminologia banku, ikasmaterial eta testu juridiko sorta mardulera iristeko egindako ibilbidearen nondik norakoez arituko gara. Garbizalekeria eta itzulpenaren morroi izatea ezaugarri (izan) dituen hizkera juridiko horren erronkak eta aurrera egiteko oinarriak zein diren zehazten saiatuko gara, gure inguruko hizkuntzalari eta idazleen idatziak lagun hartuta.

1. Sarrera: lehen gogoeta

Gauza jakina da euskara, mendeetan zehar, baztertua —eta auto-baztertua— izan dela botere publikoen eremutik, salbuespen oso urriak kenduta. Euskarak ez du aukerarik izan legeen munduan hazteko, bilakatzeko, egokitzeke, lege-hizkera bere egiteko. Ez bakarrik aukerarik eman ez digutelako; baita, euskaldunek eremu hori landu ez dugulako ere. Gogora dezagun Axularrek «iracurtçailleari» esaten diona *Geroren* hitzaurrean: «Baldin eguin baliz euscaraz hanbat liburu...».

Koldo Mitxelenak (1978: 26) aspaldi ohartarazi zigun gisa, «mende luzeen lasaikeriak honenbeste eginkizun bizkarreratu digun garaian» bizitzea egokitu zaigu. Hainbeste dira halako eginkizun bereziki zorrotza eskatzen duten alorrak; baina ez da dudarik zuzenbidekoak egin behar diela aurre eremu horien guztien artean, mendeetako lasaikeriak bizkarreratu dizkigun eginkizunik zorrotzenetakoei, astunenetakoei eta zailenetakoei. Zuzenbidearen eremua eta, ezinbestez, berarekin lotuta dauden eremuak; besteak beste, eta guztien gainetik, administrazioarena, edo hobeto esanda, administrazioena. Bereziki zaila baita eremu horretan egokitu zaigun erronka, aldi berean elkartzen direlako lexikoari dagozkion erronkak eta, bere osotasunean, *diskurtsoaren* hizkerari dagozkionak. Lege-hizkerarena ez da, bakarrik —agian, oso inportantea izan arren—, lexiko teknikoaren arazo hutsa.

Legelariok —eta hori nik diot, Alberto Lopez Basaguren; Miren Azkarate ez da honetan bekatari— harroegiak izan gara uste izan baitugu nahikoak ginela geu, zuzenbidean adituak iza-

¹ Egileak lan hau finantzatzeko izan duen dirulaguntza eskertu nahi du: Eusko Jaurlaritza, IT1169-19 (2019-2021).

² Grupo de Investigación del Sistema Vasco de Ciencia y Tecnología, Gobierno Vasco, IT1380-19.

tearren, behar genuen lege-hizkera eta lexikoa sortzeko eta egokitzeko. Gauza jakina da euskaldun bakoitzaren buruan hizkuntzalari bat ezkutatzen dela. Baina gaztelania hain hizkuntza finkatua eta estandarizatuta izanda ere, legelariak hizkuntza horretan sortu duten —eta darabilgun— hizkera, hiztun arruntentzat sarritan higuinagarria, ulergaitza eta okerra baldin bada, umilago izan behar genuen geuk geure hizkuntzaren ahultasunak ezagututa. Aspaldi ohartarazi zigun Koldo Mitxelenak (1954: 39): «aberatsek galera handirik gabe egin dezaketena, ordea, ez zaie behartsuei egoki. Eta nor gu baino behartsuago gai honetan?» Mitxelenak hori esan zuen literaturako klasikoen irakurtzearen eta argitaratzearen beharraz ari zela. Baina bada uste duenik Mitxelenak ez lukeela berdin esango lege-hizkerak gurean pairatzen duen egoeraren aurrean?

2. Euskara Batuaren sorrera

Euskararen eta lege-hizkeraren beharrik handienetakoa aldaera estandarrik eza izan da. Berrogeita hamar urte besterik ez dira igaro Euskaltzaindiak euskara batuaren abiapuntutzat hartzen den Kongresua antolatu zuenetik. Izan ere, 1968an Mitxelenak Arantzazun aurkeztu zuen txostena, *Ortografia*, ondoren egindako lanaren bide-orritzat har daiteke: idazkera edo ortografia, antzinako euskal hitzen formak (ondare lexikoa), hitz berrien eraketa eta besteetatik hartuen forma (eratorpena eta hitz-elkarketa batetik eta maileguak nola egokitu, bestetik), morfologia eta joskera. Hein handi batean, Arantzazun bertan egin zen *hautatzea*, hizkuntza bat estandarizatzeko lehen eginkizuna³ (Haugen 1987): «Ez da nor Euskaltzaindia, oraingoz behintzat, oinarritzat zein euskalki hautatuko duen erabakitzeak. Badirudi, halarik ere, euskara idatziaren beharretarako egokia goak direla, besteak beste (eta Bilbo euskaldun ez delarik), erdialdeko dialektoak bazterretakoak baino» (Mitxelena 1968)⁴. Erdi aldeko euskalkietara jotzea defendatu zuen Villasantek ere (1968).

Ondoko urteetan hasi zen bigarren eginkizuna, *kodetze* lanari dagokiona. Baturako hautatutako forma haiek hedatu ziren garaiko eragileen artean, ebaluatu zen nolako harrera izan zuten 1970eko hamarkadako erabakiek. Hau da, kodetzearen ondoren *implementazioa* etorri zen, aldaera estandarra hedatu zen eta komunikazio formalaren esparrura eraman. Baina hizkuntza bateko aldaera estandarra eratzea amaierarik gabeko lana da; zentzurik estuenean, beti geratzen da *elaborazio* lana, hau da, aldaera estandarra eguneratzea, aberastea eta komunikazio-funtzio guztietarako baliagarri egitea (besteak beste, terminologia eguneratzeko lana edo modu bateko eta besteko estiloak garatzea).

Inguruko hizkuntzetan ere, estandarizazioa aspaldi amaitutzat ematen denetan, beti izaten dira terminologia beharrak, hiztegia eguneratu beharrak. Bestelako kezkarik ere ez da falta izaten, ordea: «Dena dela, hizkuntza guztiek egin behar dute egoera berrietara egokitzeko ahalegina. Hizkuntza txiki nahiz handi izan, denek egin behar zenbait aldakuntza biziko badira. Eta aldakuntzak beti ekarri ohi du egonezina, eztabaida eta zalantza. Hizkuntza handien jauregietan ere sarri

³ Ezagunak dira Haugenek (1987: 59), hizkuntza-plangintzaz ari dela, jasotzen dituen lau eginkizunak. Horietatik bi gizar-teari dagozkio, eta beste biak hizkuntzaren barrukoak dira:

	Norm	Function
Society	(1) Selection	(3) Implementation
Language	(2) Codification	(4) Elaboration

⁴ Euskara estandarriera egokitu dugu aipamena. Bestalde, morfologian erdialdeko euskalkiak hautatu izan arren, lexikoan edo joskeran euskalki guztietako ekarriak integratu dira, neurri batean edo bestean.

ikusten dira horrelako egoerak. Baita frantsesaren eta gaztelaniaren jauregietan ere. Ingelesaren eragina gehiegizkoa dela, gazteriaren hizkera behar bezain jaso eta aberatsa ez dela, ortografian akats larregi egiten direla... horiexek dira, besteak beste, frantsesaren eta gaztelaniaren jauregietan azken boladan entzuten diren kezkak eta aldarriak» (Zuazo, 2000: 13). Hizkuntza gutxituetan askoz nabarmenagoa da behar hori: «Eta hizkuntza handien jauregietan horrelako erabakiak hartzen badira⁵, akademikiek eta gobernuek horrelako astinduak eragiten badituzte gizartean, zer ez ote da gertatuko hizkuntza txikien bordetan! Astinduak ez ezik, benetako lurrikarak, benetako iraultzak ere izaten dira hauetan» (Zuazo, 2000: 14).

Euskararen kasuan, *hautatze*, *kodetze* eta *inplementazio* ahaleginek izandako harrera iraultzat har daitezke, nolako arrakasta izan duten kontuan hartuta: «Ikastola, gaueskola, Euskara Batua, euskaldun berriak, euskara alderdi erdaldunetara hedatu izana, eta euskaldunon arteko batasuna estutu izana dira (...) 60ko hamarkadan abian jarri zen Iraultzak ekarri zituen ondorio nagusiak» (Zuazo, 2000: 17). Iraultza horretaz gain, euskarak, *elaborazio* aldian, esparru bakoitzeko diskurtso mota egokia landu eta garatu beharra ere izan du, terminologia egokia adostu eta hedatu beharra. Edo, Xabier Lizardiren hitzak erabiliaz, larreko euskara «noranahiko» bihurtzeko gaitzea; eta, horretarako, hainbat jakintza-eremutarako prestatzeko, «jakite hegoek igoa» izan behar. Euskara noranahiko bihurtze eta jakite hegoek igotze horretan, Zuzenbidearen alorrerako ere prestatu behar izan da euskara.

Nolanahi ere, euskara batuaren sorrera 1968an izan zela esan arren, Arantzazu aurretik ere izan ziren saioak eta ekarpenak, hala nola ortografiari buruzko 1901eko eta 1902ko Hendaiako eta Hondarribiko biltzarrak⁶ edo 1964an Ipar eta Hegoaldeko hainbat idazlek ortografiaren edo morfologiaren adostutako neurriak (1964ko Baionako Biltzarraren erabakiak⁷).

3. Administrazio- eta zuzenbide-testuak 1937 arte

Euskararen batasunerako ahalegina Arantzazuko Kongresuaren aurretik hasi zen bezala, 1968a baino ondotxo lehenagokoak dira Zuzenbidearen edo Administrazioaren alorreko testuak ere (Elosegi, 2002). XVII. eta XVIII. mendeetakoak batik bat Iparraldeko testuak dira, baina baita Hegoaldekoak ere: administrazioen arteko gutun ofizialak, lehen udal ordenantza eta bestelako udal agiriak, epaiketa agiriak; zuzenbide pribatuan, erlazio juridiko eta merkataritzakoetako euskarazko agiriak, eta abar. XIX. mendean ugaltu egiten dira Hegoaldeko euskarazko testuak (elkarrekin araudiak, maizter-kontratuaren ordainketa agiriak...), eta 1850 ingurutik aurrera goi-mailako administrazioetako testuak ere aurki daitezke (Bizkaiko Batzar Nagusietako aktak, erdaraz idatzi eta euskarara itzuliak)⁸.

⁵ 1994ko Toubon legeaz ari da, kongresuetan, unibertsitatean, iragarki eta publizitate-mezuetan, eta baita erakunde, elkarte eta marka izenetan ere, frantsesa erabiltzea babestu eta bultzatu zuen legeaz. Lege honek oinarri hartzen du bi urte lehenago, 1992an, Konstituzioa eraberritzean, 2. artikuluan ezarri zen xedapenean: «Frantsesa da Errepublikako hizkuntza». Kezka beretik sortu zen Fernando Lázaro Carreter akademikorekin *El dardo en la palabra* ere (1997, Galaxia Gutenberg-Círculo de Lectores) 1975etik aurrera idatzitako artikuluen bilduma.

⁶ Zalvide, M. (2003): «Hendai-Hondarribietako biltzarrak: XX. mendeko hizkuntza plangintzaren iturburu», *Euskaltzaleen Biltzarraren Mendeurrena*. Bilbo: Sabino Arana Elkargoa, 1-134.

⁷ Davant, J.L. (2005): «Baionako biltzarra (1964)», *Nerekin yaio nun. Txillardegiri omenaldia*. Iker-17. Bilbo: Euskaltzaindia, 207-216.

⁸ Gallastegiren (2005) lanean ikus daitezke horietako zenbait.

Mende horietako testu batzuk euskaraz sortuak dira; beste batzuk, itzulpenak. Iparraldekoetan «nabarmena da, itzuliak ez direnetan ere, frantsesezko hitz eta formula burokratikoak euskaraturik agertzen direla. Terminologia juridikoa, hein handian maileguzkoa da» (Elosegi, 2002: 197). XX. mendearen hasieran, 1917an eta 1918an, hurrenez hurren, Bizkaiko Aldundia zirkularrak bi hizkuntzetan zabaltzen hasi zen eta Gipuzkoako Komisio Probintziala, berriz, udaletara bidalitako zirkularrak bi hizkuntzetan bidaltzen. Horiek ere, itzulpenari zor zaizkion testuak dira. Beraz, benetako testu juridikotzat eman daitezkeen artean, nabarmena da testu itzulien pisua.

XVII. mendetik aurrerako testu horiei dagokienez, bestalde, kontuan hartu behar da aldaera estandarrik gabeko garaiko ahaleginak direla; Haugenek (1987) aurreikusitako lau eginkizuneko lanari inork ekin aurrekoak. Ez da, beraz, harritzekoa Elosegik (2002: 214) esatea «gainerako euskal testuetan bezala, testu juridiko-administratiboetan ageri da batasun falta eta estandarrik eza». Hala ere, «Lapurdiko udaletan, gutunerian, notario-eskrituretan, garatutako hizkeran, frantses administratiboaren eragin itzela zegoen, dudarik gabe, baina era berean, eredu nahiko bateratu bat bazegoen, bi mende luzetan erabili zena eta euskal hizkera juridiko-administratiboaren hasikintzat jo zitekeena» (Elosegi, 2002: 214). Hasikin horrek ez zuen izan aukerarik hazteko eta garatzeko. Berehala ikusiko dugun bezala, bestelako helburua zuten Hegoaldeko euskarazko testuek. Eta Zuzenbideko hizkeraren ezaugarriek ere ez zion laguntzen mugaren bi aldeetarako balio zezakeen hizkera eta terminologiari.

Aldaera estandarra landu gabe, bat-batean ekin zitzaion 1936an *Euzkadi'ko Agintaritzaren Egunerokoa / Diario Oficial del País Vasco* bi hizkuntzetan argitaratzeari; horixe da ordura arte inoiz argitaratutako testu juridikoen bildumarik garrantzitsuenak. Baina testu hauetan eta, oro har, XX. mendearen hasierakoetan, nabarmena da hizkuntzaren funtzio sinbolikoa: «hainbat testu ez da egiten (aurreko garaietan gertatzen zen bezala) euskaldun jendeak hobeto uler dezan, hizkuntza agerrarazteagatik baizik» (Elosegi, 2002: 212).

Laburbilduz, aldi bakoitzeko testu-kopurua gorabehera, ezaugarri hauek dituzte XVII. mendean hasi eta 1937 arteko administrazio eta zuzenbide alorreko idatziek: itzuliak dira gehienak, ez jatorrizkoak (eta, berehala esan dezagun, ezaugarri berbera dute 1980tik aurrerakoek ere); funtzio desberdina dute Iparraldekoek eta Hegoaldekoek (XX. mendekoek). Iparraldean komunikazio-funtzioa lehenesten da eta horrek dakar nabarmena izatea testu horietako lexiko eta esapideetan frantses administrazioeko hizkeraren eragina. Hegoaldean, aldiz, batetik funtzio sinbolikoa zuen euskara erabiltzeak, komunikaziokoa baino areago; eta, bestetik, agerikoa zen Larramendiz geroztik indarra hartu zuen garbizaletasunaren eragina.

4. Larramendiren eta Hegoaldeko eredu garbizalearen eragina lexikoan eta (zuzenbideko) terminologian

Euskara Batuaren aurreko saiotzat hartu ditugu XVII. mendetik aurrera euskaraz idatzitako testu juridikoak. Hala da, ia XX. mendearen amaiera arte ez duelako izan euskarak aldaera estandarrik. Eta, hala ere, ezin esan daiteke aldaera estandar bati buruzko ikuspegirik (nolabaiteko ikuspegirik) ez zegoenik Larramendiren edo Aranaren eta bere jarraitzaileen artean (Azkarate 2015, Azkarate eta Lindemann 2018). XVIII. mendean, Larramendiren (1745) hiztegia, *Diccionario Trilingüe del Castellano, Bascuence y Latín*, hiztegi soil bat baino askoz gehiago izan zen. Inguruko hizkuntzak bezala, euskara ere kultura-hizkuntza izan zitekeela erakutsi nahi zuen Larramendik; horregatik ditu bi helburu nagusi bere hiztegiak. Batetik, «kanpoko» erdaldunei, arrotzei erakutsi

euskara hizkuntza bizia eta jakintzagaietarako egokia zela (Mitxelena, 1961), eta, bestetik, euskara eredu jaso eta garbiago bat eskaini nahi zien garai hartako predikariei. Bi helburu horiek kontuan hartuta, Mitxelenak (1984: 33) dio: «Larramendi izan da, lehenbiziko aldiz eta erabat, hizkuntza larrekia noranahiko bihurtu nahi izan duena». Terminologia beharrekin topo egin zuen Larramendik ere jakintzagaietako hitzak sartu nahi izan zituenean; sartu baino areago, asmatu, esan beharko genuke, *voces facultativas* deitzen dituenak. Behar zituen termino horiek euskaratik beretik sortzera jo zuen; latinetik edo grekotik mailegatu ordez, nahiago izan zuen euskarak dituen baliabideetara jotzea. Ez da hau lekua Larramendiren sorkariak zehatz-mehatz aztertzeke, nahiz merezi lukeen norbaitek Larramendiren termino juridikoak biltzea. Hauexek jaso ditugu, guregana heldu ez direnen erakuskari: *iritarau* ‘política’, *bialkinde* ‘embajada’, *bialkin* ‘embajador’, *inolazkogo* ‘nulidad’, *bedoi* ‘prueba (con la que se justifica algo)’, *atzipe* ‘dolo, engaño’, *ekadoiko* ‘jurídico, judicial’, *araujarraide*, *araujakintza* ‘jurisprudencia’, *kotiza* ‘tributo’, *legaraude* ‘legalidad’, *legarrauko* ‘legítimo’ edo *legoki* ‘legítima, parte de la herencia’. Askok dira ordain bateko baino gehiagoko sarrerak ere: ‘ministro’ *ekidari*, *ministro*, ‘regente’ *erregentea*, *erondaria*, *gobernataria*, *bazkazalea*, ‘aduanak’ *aduanak*, *peajelekua*, *kostamaga*, ‘pena, castigo’ *pena*, *minaria*, *gaitzondorea*. Larramendik sorturiko hitzetako batzuk, oster, guztiz onartuak daude gaurko hizkera juridikoan: *besterendu*⁹ ‘enajenar’, *ordezte* ‘subrogación’, *urube* (orube) ‘solar’, *gaieztua* (zegaitua) ‘inhabilitado’ guztiz forma horretan izan ez arren, *salaketa* ‘acusación’...

Larramendiren hiztegiko proposamenek eman zuten bide, adibidez, 1820. urtean *Jaquin-bide Iritarautia* argitaratzeko¹⁰. Eta Larramendiren hiztegiko terminoak ageri dira *Lore-jokoetan* ere: *ekauta* ‘diputación, comisión’, *goarpelari* ‘secretario’, *neurkida* ‘regla, ordenanza’, *otsaunditu* (*hotsanditu*) ‘solemnizar’ edo *salkindari* ‘traidor’, gutxi batzuk baino ez ekartzearren hona¹¹.

XIX. mendearen amaieran, Arana Goirik, garai hartan beste hizkuntza askok eginiko bidetik, jatorri arrotzeko hitz guztiak baztertu eta haien ordez euskal hitz jatorrak sortu nahi izan zituen¹², askotan euskarazko eratorbidea edo hitz-elkarketa arauak kontuan hartu gabe. Garai horretan nagusitu zen purismoaren aldeko joerak nabarmen baldintzatu zuen orduan sortutako testuetako jostakera, bai eta terminologia ere. Hala ere, Aranaren neologismo askok iraun dute eta gaur euskara estandarreko lexikoaren zati dira (Azkarate eta Lindemann, 2018): *alderdi*, *jaurларитза*, *abertzale*, *hauteskunde*, *hautetsi*, *ordezkari*, *batzorde*, *idazkari*, *espetxe*, *aldundi*, *aberri*, *bazkide* edo *lehen-dakari*.

Testuinguru horretan hasi zen argitaratzen bi hizkuntzetan *Euzkadi’ko Agintaritzaren Egunerokoa / Boletín Oficial del País Vasco*¹³. Ez da zaila Sarriegirekin (1985: 11) bat etortzea eta laudagarri eta txalogarritzat jotzea euskarari ordura arte inon eman gabeko hedadura ematea, «ez denbora, ez oinarri, ez lasaitasun eta agian epe laburrerako asmorik (euskara oinarria izanik) ez bazen ere» eta «nahiz eta berori ezinbestean itzulpen huts-hutsa izan eta beharbada ez hizkuntz eredurik proposagarrien eta aukerakoenean egina ere, orduko **Agintaritzaren Aldizkarian** ikus

⁹ Larramendi eredu harturik sortu bide zituen Azkuek *inorendu* adiera berarekin (‘enajenar’) eta *norberarendu*. Azken honek ez du agerraldi bakar bat ere ez *Zuzenbide corpusean* eta ez *Egungo Testuen Corpusean* ere.

¹⁰ Cadizko Konstituzioa euskaraz, Jose Felix Amundarainek izenburu horrekin itzulua. Zumalakarregi Museoaren webgunean kontsulta daiteke eta, nahi izanez gero, testu osoa eskuratu ere bai pdf formatuan http://www.zumalakarregimuseoa.eus/eu/ekintzak-eta-zerbitzuak/ikerketak-eta-dokumentazioa/argitalpen-digitalak/1812ko-dotrina?set_language=eu

¹¹ Ikus Elozegi (2002) Eranskinak, 2.6.5. Lore-jokoetako testuetako hitz juridiko-administratiboak.

¹² Ikus Pagola (2005).

¹³ Azterkizun daude *Eguna* egunkarian eta *Ereintza* aldizkarian 1937 urtean argitaratutako artikuluetako gaiak eta terminologia (Elozegi, 2002).

daitekeenez». Baina nabarmena da terminologia¹⁴ garai hartako joera puristaren mende dagoela: *uri sari-artzalle* ‘becario’, *uri-sari* ‘beca’, *landola* ‘dependencia’, *Zaingoa* ‘departamento’, *oroagindu* ‘decreto’, *Goi Banakaritza* ‘Administración Central’, *erki-banakaritza* ‘administración provincial’, *lanari* ‘funcionario’, *latziñaldari* ‘notario’, *Epailletza argibidelari* ‘Juzgado de Instrucción’, *txindi* ‘dinero metálico’, ‘público’ esateko *herri-*, hitz elkartuko lehen osagai... Erabiltzen diren terminoak ere ez dira beti osaeraz egokiak: *ibilkor zubi* (adjektiboa eta izena, alderantzizko hurrenkera behar lukeena), adverbio + izen egiturako terminoak (*ia-injeniero* ‘perito’, *ia-osagille* ‘practicante’, *ia diru artu-emangarri* ‘semibancable’¹⁵). Beste askotan, ostera, guztiz egokiak dira osaeraz: *diru-etxeen alkartze eta banatze arazo*, *Diru-etxeak arautzeari buruz emandako Legea*, *Diru-Zaintza Araudia*, *Ettxerik gabeko lurren zergari*, *zerga bereziei*, *eta plus valiari dagokien Araudia*, *zergaren oñarri*. Edo bete-betean gaur erabiltzen direnak ez izanda ere, oso gertukoak dira: *erantzina* ‘anexos’.

Badira bi ezaugarri azpimarratzekoak iruditzen zaizkigunak, biak ere gaurko terminologian ikus daitezkeenak: batetik, terminologiaren aldakortasuna eta bestetik, eremu askotako terminoetan ohikoa dena, alegia unitate konplexuen (sintagma egiturako unitateen) ehunekoa euskal segmentu baliokideetan (goraxeago emandako adibideez gain, *lan eta alogera gabe uzteko zigor*, *gai-xo-susmopean dagoen abel-gorri*, *diruaz ber-jabetu al izateko argibide* edo *legez biztanleturik aurkitu* ‘tener residencia legal’). Aldakortasunaren adibide (nabarmenenak) dira: ‘contratista’ *lanartulari*, *itunbenlari*, *lanartuari*, *lan-artzalle*, *lanartuzale*, *lanartutzeale*, *lanartzelari*, *lanartzale*, *lanartzalle*, *lan-artulari*; ‘sanción’ *ezitu-bear*, *jipoi*, *zigor*, *zigorkada*, *zigortze*, *nekaltze*, *neke* edo ‘Junta provincial de Protección de Menores, Junta provincial de Protección a la infancia’ *Bizkaiko Aur Babes-etxea*, *Aur laguntza Erki-Batzordea*, *Aur Laguntzalle Erki-Batzordea*, *Gaztetxoan Erki Zainketa-Bazkuna*, *Aurrendako Babes Erki-Batzordea*, *Erkietako Aurrendako Babes Batzordea*, *Gaztetxoan Babeserako Erki-Batzordea*, *Aurraren Laguntasun Erki-Batzordea*. Guztiz ulergarria da aldakortasun hori jakinik zein egoeratan egin zen ahalegin hura (ikus goraxeago Sarriegiren hitzak). Gaur asko murriztu da aldakortasun hori, baina badira oraindik ere sinonimoak: *UPV/EHuko Zuzenbide Fakultatearen Hiztegi Juridikoan*, esaterako, ‘ámbito de la acción protectora’ *babes-jardunaren hedadura*, *babes-jardunaren irismen*; ‘albacea’ *albazea*, *testamentu-betearazle*; ‘audiencia previa al juicio’ *epaiketa aurreko entzunaldi*, *judizio aurreko entzunaldi* edo ‘auditoría’ *auditorretza*, *auditoria*, *ikuskaritza*. Terminologiaren teoria berrietan onartzen da komunikazio espezializatuan erabiltzen den terminologia errealean aldakortasuna gertatzen dela¹⁶; areago, erdal termino batek adiera bat baino gehiago izan dezake eta, horrenbestez, ordain bat baino gehiago eska dezake (ikus azken puntuan Alvarez Mingoren aipamena).

Ezin dugu ahaztu, bestalde, euskaltzainen artean ere zenbaterainoko lekua zuen maileguren auziak. Agerikoak ziren Altuberen eta Azkueren iritzi kontrajarriak. Gerra ondoren ere iraungo du eztabaida horrek (Azkarate, 1997).

Atal honen laburpen modura esan dezakegu Elosegik dioena oso ondo egokitzen zaiola XIX. mendeko hizkera juridikoari eta XX. mendearen hasierakoari: «Jakina da zein pisu handia izan duten Hegoaldeko literatura osoan, baita itzulpengintza juridikoan ere, euskararen zuzentasunari

¹⁴ IVAPek argitaratu zuen 1997an *Euzkadi’ko Agintaritzaren Egunerokoaren* 252 aleetako hitz-bilduma. Lexiko teknikoa edo berezitua ez ezik, lexiko arrunta ere jasotzen duena.

¹⁵ Semantikaren ikuspegitik ere hain bitxia den *ia* adverbioaz gain, euskaraz ezinezkoa den *diru artu-emangarri* adjektibo elkartu sintetiko itxurakoa.

¹⁶ Ikus, esaterako, Freixáren lanak, bereziki (2002, 2005, 2013): autoaldakortasuna (autore beraren testu bakarrean edo testu batean baino gehiagotan gertatzen dena) / heteroaldakortasuna (hainbat autoreren testuetan gertatzen dena. Edo Lorente (2013).

buruzko aurreiritziek. Jakina denez, hainbat hizkuntza baliabide, bai lexiko, bai morfosintaktiko, bai estilistiko, debeku izan dira garai batzuetan. Horrela, esan daiteke Iparraldeko tradizioa ez zela kontuan hartzen —hartzekotan, eredu negatibo gisa— euskarak 1936ko urrian ofizialtasuna eskuratzeko duen unean» (Elosegi, 2002: 215).

5. 1980tik aurrerakoak

1978an *Euskal Kontseilu Nagusia* sortu zenerako, euskara batuak finkatuak zituen morfologiako funtsezko kontuak. Baina oraindik bizirik zeuden gerraurreko irizpide garbizaleak. Eta berehala hasi ziren tirabirak 1936ko ereduari jarraitu nahi ziotenen eta Euskaltzaindiaren gidaritzapean aurrera zihoan aldaera estandarren aldekoen artekoak. Eztabaidak bizi-bizirik jarraitu zuen eta eztabaida horretako hainbat alderdi agerian geratu ziren HAEE/IVAPek urte haietan antolatu zituen administrazioa hizkeraren inguruko jardunaldietan¹⁷. Berasategik (1989) bi eredu bereizten ditu garai haietan: 1936-37 garaiko tradizioari atxikitzen zaiona bata, hizkuntza modernizatu eta garai berrietara egokitu nahi zuena bestea. Tradizioari atxikizaleen jardunak aurreko garaiko hizkeraren ezaugarri berberak ditu: garbizalea da, mailegutzatik ihes egin nahi du; garbizaletasunak, euskal hiztunen ohiko moldeetatik urruntzera daramatza; maiz jotzen du hitz-elkarketara eta postposizio-atzizkiak metatzera; euskal testuari autonomia handiagoa eman nahi diete. Hizkuntza modernizatzearen aldekoek, berriz, espainolezko terminologiaren zehaztasunak eta konnotazioak gorde nahi dituztenez, erraz onartzen dituzte maileguak; inguruko alorretan garatzen ari den hizkerarekin (irakaskuntzakoarekin edo komunikabideetakoarekin) bat datorren moldetik joan nahi dute; maileguz beteriko hiri-hizkerara hurbildu nahi dute. Aurreko tradizioa ez ezagutzeak ere ez zion inolako mesederik egin hasierako urte haietako jardunari: «Orduko ez jakitearen kariaz edo, hobeto esanda, falta zen ikerketaren ondorioz, aurki dezakegu orduko testuetan euskaraz berez zeuden terminoak gaztelaniara itzultzen zirela, eta, horiek berriro euskarara ematean, euskaraz inolako tradizioarik gabeko ordainak asmatu zirela, adibidez «herrialde» > «territorio histórico» > «kondaira lurralde», «lurralde historiko», «herrialde historiko» eta «historia lurralde». Orduko joera biek aukeran zuten horietako bat edo beste, baina ez zuten jatorrizkoa proposatu» (Lobera 2005: 116). Garai hartatik geratu zaigu *Herri-Arduralaritzaren Euskal Erakundea* izena edo ‘público’ erdal erreferentzia-adjektiboaren ordain *herri* ematea (askotan behintzat): *herri-ikastetxeak*, *herri-lan(ak)*, *herri-zerbitzu* (nahiz *publiko* mailegua nagusitu).

Ondorengo urteetan, euskara batuaren oinarriak finkatzen diren heinean, testu juridiko gehien-gehienak haiek jarraituz moldatzen dira. Beste arloetan bezala, euskara batuarekin batera, lexikoan ere aldaketa nabarmenak ageri dira (Elosegi, 2002: 220). Ekarpen handia egin du Terminologia Batzordeak, han ezarritako irizpideek eta han bildu den lantaldearen jardunak (<https://www.euskadi.eus/euskararen-corpus-terminologia/web01-a2corpus/eu/>). Bestalde, HAEE/IVAPek eginitako lana aitortu behar da, bai terminologian eta bai hizkera argi eta ulergarriaren alorrean; administrazioaren alorretarako prosa eredu estandar zaindua lantzeko argitaratutako lanak, bereziki *Estilo liburua*, *Administrazio Euskaraz* aldizkaria, *Euskara Juridikoaren Ataria*, *IUSPLAZA Zuzenbidearen webgunea euskaraz* edo *Legeak-Leyes bilduma*¹⁸. Itzulpenek oraindik ere pisu handia

¹⁷ Bi daude argitaratuak, 1982koa eta 1989koa. Herri-Arduralaritzaren Euskal Erakundea / Instituto Vasco de Administración Pública (1982): *Euskara eta administrazioari buruzko I. Ihardunaldiak*, HAEE/IVAP; Herri-Arduralaritzaren Euskal Erakundea / Instituto Vasco de Administración Pública (1989): *Administrazioa hizkera eta terminologiaz jardunaldiak*. HAEE/IVAP. Ikus Elosegi (2002: 274-276).

¹⁸ <https://www.ehu.eus/eu/web/euskara/legeak-leyes-bilduma>

izan arren, badira euskaraz sortzen diren testuak ere. *Eleria* aldizkariak (1996-2013)¹⁹ aurrerabide nabarmena ekarri zuen euskara hutseko lanak argitaratzen ziren neurrian, komunikazio-funtzioa zaindu behar zuelako. Hamazazpi urtean zehar unibertsitateko irakasleen eta IVAPeko teknikarien artikulua eta lanak jaso zituen. Gaur merezi du *Zuzenbidea ikasten. Irakaskuntzarako aldizkaria* aipatzea. 2014an sortua, *Zuzenbideko graduari* aritzeko eta *zuzenbidea ikasteko* erabiltzen diren ikasmateriak argitaratzea du helburu, bai eta ikasleen ekarpenak eta ikasleentzat interesgarriak izan daitezkeen berri juridikoak ere. Eta, nola ez, unibertsitateetan sortu diren irakaskuntzarako testuliburuak (EHUkoak <https://ikasmateriak.ehu.es/zuzenbidea>, edo Deustuko Unibertsitateko Zuzenbide alorreko bilduma). Ibilbide horri esker egin ahal izan da Zuzenbide Corpusa, gaur egun 696 dokumentuz (lege-testuak eta testu akademikoak) eta 7,9 milioi testu-hitzez osatua, eta Euskara Institutuan kontsultagai dagoena.

Horrek guztiak errotik aldatu du terminologiaren egoera. Orainsu arte hiztegien bidez sortzen ziren terminoak; hiztegiek egiten zituzten proposamenak zein termino erabili erdarazkoaren ordain, bai euskara batuaren aurreko aldian eta bai ondoren ere. Lantalde baten edo pertsona jakin baten proposamenak jasotzen ziren (terminologia) hiztegieta. Gaur, aldiz, baditugu euskaraz sortutako testuak, corpusak, horiek aztertu eta hortik terminoak erauzteko: corpus orokorrak (*Egungo Testuen corpusa*), nahiz corpus berezituak (*Zuzenbideko corpusa*)²⁰, Horrek eman digu aukera terminologia planifikatutik terminologia naturalerako pausoa emateko.²¹

Ez da dudarik XXI. mendearen lehen laurdenera iristen ari garenean, lege-hizkeraren eta, oro har, zuzenbideko eta administrazioko hizkeraren egoerak ez duela zerikusirik orain dela mende batekoarekin. Oro har, esan daiteke euskara estandarrek berak egin duen ibilbidea jarraitu duela; ibilbide hori, Mitxelenak (1968) Arantzazun esandakoaren arabera egin dela esan dezakegu: «Euskaldunok ez dugu, herri gaitan izan ezik, kultura berezirik: mendeko gera, kultura-barruti horretan bizi gera. Ezin izan dezakegu, beraz, askatasunik hizkuntzaren aldetik: askatasun hori erdiets zitekean VIIIgarren mendean, XIIIgarrenean eta XVIgarrenean ere, bai oraindik, agian, ez, ordea, oraingoan eta are gutxiago hurrengoan. Latinaren, eta erabat sortaldeko hizkuntzen, kutsua geure bekokian daramagu. Aitor dezagun, onar dezagun, eta goazen aurrera bide horretarik. (...). «Euskal-usaia», «euskal-kutsua» deritzatena ongi dago, noski, euskal-gaiak darabiltzagunean, Euskal herriko mendi-giroa irakurleari sumarazi nahi diogunean. Hortik kanpora, ordea, badirudi ur garbiaren pare behar duela izan behar bezalako prosak: inongo usairik, kolorerik eta zaporerik gabea. Euskal-kutsua badario, barren-muinetarik dario, ez axaleko edergarrietarik».

6. Zenbait gogoeta: euskarazko lege-hizkeraren erronkak eta aurrera joateko oinarria

Gai horretaz arduratu direnek ia aho batez aitortu dute itzulpenaren eskutik etorri dela lege-hizkera eta, oro har, administrazioko hizkera. Terminologian ere bada oraindik zer gogoeta egin. Beraz, bi puntu horien inguruan harilduko ditugu aurrera begirako gogoetak, beste idazle batzuenak euskarri hartuta.

¹⁹ Euskal Herriko Legelarien Aldizkaria, <http://www.eusko-ikaskuntza.eus/eu/argitalpenak/eleria-euskal-herriko-legelarien-aldizkaria/bi-109/>

²⁰ <https://www.ehu.es/eu/web/eins>

²¹ Alberdi X. eta Salaburu, P. (arg.) (2013): *Terminologia naturala eta terminologia planifikatua euskararen normalizazioari begira*. Ugarteburu Terminologia Jardunaldiak V <https://www.ehu.es/eu/web/eins/jardunaldiak-2013>

Itzulpenez

Ingurumari zabalagoa eman nahi diogu itzulpenari buruzko gogoetari, azken batean itzulpen lanaren gaineko gogoeta orokorrek berdin-berdin balio dezaketelako esparru zehatzagoetarako ere; eta, askotan, gainera, argi egin dezakete zein bide hautatu. Lege-hizkera egokitzeo eginkizunean itzulpena izan dugu lanabes eta bide ia bakarra. Hori, nekearen nekez ikasi dugu, ez da, berez, ez ona ez txarra. Are gehiago P. Esterházyk diosku, hungarieraren kasuan, itzulitako literaturak oso tradizio sendoa eta oparoa izan duela beti, bere ustez herri txikiek bere buruarekiko duten begirunearen bereizgarri dena²². Baina eskarmentu handiko —eta kalitate literario goreneko— idazleek badakite itzultzeak zer zailtasun eta arrisku dituen. Bi adibide hona ekartzeagatik: Jorge Luis Borgesek diosku ez dagoela bestelako arazorik letrek —idazkuntzak— hain bere-berezko duenik, eta haien misterio apalarekin lotuta dagoenik, itzulpenak proposatzen duena baino²³. Eta, gurean, Koldo Mitxelenak ohartarazi zigun «itzulpenak irristakorak eta zailak direnez gero»²⁴ trebezia berezia eskatzen dutela; bereziki, uste dugu, hizkuntza maisuki menderatzea eta gaitasuna izatea ez-titasun guztiarekin erabiltzeko.

Hizkuntza bakoitzak badu bere unibertsoa, bere berezitasunak, bere-berezkoak diren bi-deak. Axular izan dugu gurean, duda gabe, itzultzailerik trebeenetakoa. Anjel Lertxundik²⁵ berriki ekarri dizkigu Axularrek erabiltzen zituen metodoak; era berean, hitzez-hitzezko itzulpena eta, gehienetan, horrekin batera, jarraian, itzulingurukoa, sarritan sinonimo ugari erabiltzearen bidez. Axularrek berak, gardentasun guztiarekin, biluzten du bere burua: «...eztitut bethiere... hitzez hitz euscarara bihurtzen. Ceren euscarra eta bertce hitzcuntçac different baitira. Ordea ezta ez handic seguitcen gaiztoago dela euscarra. Aitcitic badirudi ecen bertce hitzcuntça eta lengoaya comun guztiac bata bertcearequin nahasiac direla, Baiña euscarra bere lehenbicico hastean eta garbitasunean dagoela»²⁶. Ez da dudarik, neurri batean behintzat, azken xehetasunak, euskararen garbitasunari buruzkoak, beste xehetasun eta ñabarduraren bat behar duela. Baina Axularrek erakusten digu nola egokitu euskararen ezaugarrietara euskarara ekarri behar dugun hori. Mitxelenaren hitz hauetan ere aurki dezakegu Axularren mezuaren oihartzuna; eta honela uler dezakegu mezu hori: euskarazko lege-hizkera egokitzean ez dugu hizkuntza berria asmatu behar, euskara egina baitago, beharrak izan arren: «Euskarak, egia esan, badu beharrik; zertxobait leuntzeko, hornitzeko eta egokitzeo premia ederrean dagoela guziok dakigu. Baina egin beharrik ez, aspaldidanik egin-egina baitago. Euskara egiten —edo egin ustez desegiten— hasi baino hobe genuke euskaldunok euskara ongi ikasten hasi» (Mitxelena, 1961)²⁷.

Anjel Lertxundik behin eta berriro errepikatu du itzulpen literarioak emaitza oparoak ekarri dizkiola gure hizkuntza literarioari. Itzultzaile benetan trebeak sortu baitira gurean azken hamarkadetan, agian hizkuntzaren inguruan izan dugun emaitzarik aberatsenetakoa. Maisu izan dira kanpoko hizkuntzetako bideak gure hizkuntzaren berezko ezaugarrietan egoki txertatzen. Hizkuntza aberastu dute hizkuntza bortxatu gabe.

²² Esterházy, Peter, Dezső Kosztolányi-ren *Alondra* nobelari egindako hitzaurrean, Ediciones B, Barcelona, 2002, 10. or.

²³ Borges, Jorge Luis: «Las versiones homéricas», Discusión, *Obras Completas*, I, Círculo de Lectores, Barcelona, 1992, 267. or.

²⁴ Mitxelena, K. (1965) EGAN, Manuel de Lekuonaren *Literatura oral* vascaren argitalpen berriaz [*Euskal idazlan guztiak*, III, 149. or.]

²⁵ Lertxundi, A. (2019): *Itzuliz usu begiak*, Alberdania.

²⁶ Axular, P.: *Gero*, 19-20. or.

²⁷ Mitxelena, K. (1956) EGAN Barandiaran'dar Salbadorren *Ilias'enaz*, Iliadaren euskarazko itzulpenaz [*Euskal idazlan guztiak*, II, 69. or.]

Esan da gurean sortu den lege-hizkera edo hizkera burokratikoa, nahi badugu, gaztelaniazko hizkera burokratikoaren maizter dela, *jerga* horrek dituen gaitz guztiak gurera ekarri ditugula (A. Lertxundi, *Eskarmentuaren paperak zein Itzuliz usu begiak* saiakeretan). «Hizkera burokratikoak gaztelania du, eta ez euskara, iturri eta ispilu, eta itzulpena du oinarri» diosku azken saiakeran (2019, 110). Joseba Lozanok (1997) ere aipatu ditu euskara juridikoaren ajeak eta urratuak —literaturaren eta hizkera juridikoaren arteko eremutik, esango genuke—. Erdararen morrontzatik urrundu, erdara «jaso» horren mezuak deskodetu, testuan ezkutatzen diren nagusienak zerrendan jarri, ideia horiek euskarazko oinarritzko esaldi bihurtu, eta esaldiok euskarazko diskurtso modura egituratu, euskarak berak gure esku jartzen dituen tresnen laguntzaz: horiexek Lozanoren errezetaren osagaiak.

Arazoa ez da, beraz, itzulpena oinarri izatea. Bide gehiagorik ez genuen eskura. Arazoa da gaztelania izatea, eta ez euskara, ez bakarrik iturri —hori saihetsezina baita—, baizik eta ispilu bakarra. Eta, zilegi bekigu berriro aipatzea legelarien —eta burokraten— harrokeria, uste baitute hala izan behar dela eta ez dela, horrekin batera, bestelako ispiluren bila ibili beharrik. Batzuen eta besteen artean egokitu nahi izan dute lege-hizkera. Askok pentsatzen zuten ez genuela hizkuntzarien beharrik; ez eta euskara literarioa ederto menderatzen duten idazleen beharrik ere. Zuzenbide Fakultatean, aspalditik, ohikoa da entzutea «ba gu, juristok, legelariok, horrela esaten dugu». Maiteminduta daude haiek beraiek sortu duten hizkerarekin, lexikoarekin, bereziki.

Kontuan izan behar dugu mota desberdineko arazoei egin behar diegula aurre; eta batzuen eta besteen ezaugarriak ez direla mota —eta munta— bere-berekoak. Kepa Altonagak (2015, 77) arazoa egoki ezartzen du esaten duenean «bertsio [beraz, itzulpen] *onak* eta *fidelak* behar ditugu(la) euskaraz» (nabarmentzea gurea da). Bereziki inportanteak dira testu juridikoetan Altonagak eskatutako bi betebeharrak. Ez dago inolako dudarik testu juridiko baten euskal ordainak, aldez aurretik eta ezinbestean, jatorrizkoarekin erabat *fidela* izan behar duela erabiltzaile juridikoak baliagarri —erabilgarri— izango badu. Baina, era berean, *ona* izan beharko luke hizkuntzaren aldetik ere. Berehala ikusiko dugun gisa, gure hizkuntzaren ezaugarri —ahulezia— soziolinguistiko eta ez-soziolinguistiko bereziengatik, lege-hizkuntzak —hizkuntza burokratikoak— eragin berezia du hitzunenengan, erabat kutsakorra izaten ari dena, hizkuntza jasoaren eredu gisa hartua izan den neurrian. Beraz, gurean arazoa ez da bakarrik lege-hizkera adituen *jerga* bihurtzeko arriskua izatea, baizik eta, orokorrean, hizkera burokratikoa euskara jasoaren eredu bihurtuta, hitzunen *jerga* —are gehiago, hitzun ‘hezi’ena— bihurtzeko arriskua izatea ...dagoeneko bihurtu ez bada.

Terminologiaz

Asko izan dira aldarrikatu dutenak gaztelaniazko lexiko eta hizkera juridikoaren pareko —ispilu— izango zen lege-hizkera behar genuela euskaraz. Eta halakoa sortu dute. Gaitz sendoa dugu lege-hizkeran gaztelaniaren ispilu diren ordain zuzenak sortzea. Askotan esaten dute lege-antolamendu bakoitza mundu guztiz berezia dela eta, beraz, Espainiako lege-antolamendua dela erabili behar duguna, halako legeek ezarri dutena ekarri behar dugula euskarara. Hori egia da neurri nahiko txiki batean. Gutxitan dira beste sistemetan inolako ordainik ez duten «berezitasunak». Horrek ez du arazorik sortzen; salbuespenak diren neurrian, salbuespen arrotzenek ere —hizkuntzaren ikuspegitik— ez diote kalterik egingo hizkuntzari. Baina zerk eragozten digu gurearekin antz handiago duten hizkuntzek gauzak adierazteko eta lexikoa sortzeko erabiltzen dituzten bideak aztertzea eta kontuan izatea? Ez dira onuragarriagoak guretzat ingelesak edo alemanak dituzten

erak? Ingelesean nabaria da hizkera juridikoan hitz arrunten erabilera, hitzaren eremu semantikoa zabalduta. Berdin erabiltzen da *claim* hitza hitzun arrunten arteko solasetan eta juristen artekoetan. Baina esangura ez da bera batean eta bestean.

Bestelakoa da egoera, zientziaren —edo zientzien— eremuan gertatzen den gisa, hitz edo termino zientifikoak erabat hedatuta daudenean. Ezin da dudarik izan halako kasuetan hedatutako lexikoa gureganatu behar dugula, inolako mugarik gabe, esaterako Altonagak aldarrikatzen duen gisa. Ez du ez hankarik eta ez bururik hori aldatu nahian ibiltzeak, purismoaren bidetik euskal kutsu handiago duten hitzen bila ibili beharrak. Baina Zuzenbidean lexikoa, gehienetan, ez da nazioartekoa, kasu batzuetan hala izanik ere. Gehienetan, hitz arruntak hartzen ditu oinarri eta iturri. Beraz, hizkera arrunta dute iturburu nagusia. Euskararen ere, hizkera arrunta iturburu duela sortu behar dugu lege-hizkera. Eta, horregatik, tentuz ibili behar dugu, euskara egokitzean gure hizkuntza jo duten gaitz guztiez ohartuta, lege-hizkera ez dadila gerta izurriteak jota.

Terminologia auzietara heldu garelarik, gogoan izan behar dugu zein helburu duten gurean, batez ere, terminologia lanek. Izan ere, hizkuntza gutxituen kasuan, hizkuntza plangintzaren helburua nagusitzen da; hizkuntza plangintzari, hizkuntzaren normalizazioari loturik egon ohi dira terminologia hizkuntza horretara egokitzeko ahaleginak (Cabr  1992). Leku berezia hartzen dute horrelako lanek estandarizazioaren azken fasean, *elaborazio* aldian, nahitaezkoa baita lan hori hizkuntza edozein arlotan erabilgarri egiteko. Hala izan da eta da euskararen kasuan ere. Ikuspegi horretatik, terminoak komunikaziorako elementuak baino ez dira, balio espezializatuko unitate linguistikoak, baina betiere hizkuntza jakin bateko arau eta erregelak oinarri dituztenak²⁸. Beste hitz batzuetan esanda, hitz berriak sortzeko (edo maileguz hartuak egokitzeko) hizkuntza batek dituen arauak eta erregelak bete behar dira terminoak sortzerakoan ere. Euskarari dagokionez, Euskaltzaindiaren Hiztegia izango da oinarri, batetik, hiztegi horretan baliatutako irizpideak; eta, bestetik, euskaraz hitz eratorriak eta hitz elkartuak sortzeko bideak. Hizkera berezituko hitzak sortzeak ez ditu eskubiderik ematen lexiko orokorreko hitzen arauak hausteko.

Baina nola jakin zein bide hautatu? Mitxelenak Arantzazuko txostenean eta beste idatzi batzuetan esandakoek argi egin dezakete kontu honetan. Bi ideia azpimarratu zituen Mitxelenak (1968): «Hizkuntza bizia hitz berrien premian da beti: bere baitarik sortzen ditu batzuetan, beste-rengandik hartu besteetan. Bi bideak zilegi eta beharrezko dira eta nork taxutuko du alde zurretik bakoitzaren neurria eta mugak? Euskaltzaindiak behinola esanak honetarako bakarrik balio zuen: ez dagoela hitz berririk sortu beharrik adierazteko euskal hitz, zahar-jatorrak, erdaratikakoak direlakoak, baztertzeko asmoz». Eta bigarren gogoeta, «Euskal hitzak sortzean, eta nornahik sortzen ditu mintzatzean, gogoan eduki dezagun beti, premiaz gainera, gure hizkuntzaren moldeak, beti izan baitugu horretarako erraztasuna, hitz-elkartzearen bitartez, batez ere. Noraezekoa du hizkuntzak sor-iturri horrek bizirik iraun dezan, bere izatean iraungo baldin badu».

Egia da, lege-hizkeraren alorrean ere sarri hautatu ditugula euskaren moldeko baliabideak, lexiko orokorrean egin den bezala: «Lexikoaren eguneratzeak hautamen bat egitera behartzen gintuen beste hizkuntzen morrontzaren eta euskarak berak ematen zizkigun aukeren artean. Zorionez, asko tiratu dugu bigarren aukeratik, euskara bera izan dugu iturri nagusia. Bada, ordea, tamalik ere: ez dugu beti asmatu»²⁹.

²⁸ Ikus artikulu-sorta honetan Alberdiren lana, «UPV/EHUko Zuzenbide Fakultatearen Hiztegi Juridikoa: hiztegi-sarreraren egitura eta definizioen idazkera».

²⁹ Lertxundi, A. (2009): *Eskarmentuaren paperak*, Alberdania.

Beti izaten da tirabira barrutik sortutako eta maileguz hartutako hitzen eta terminoen artean. XIX. eta XX. mendearen zatirik handienean maileguetatik ihesi ibili ondoren, badirudi eskuzabalegi jokatu dugula maileguak hartzean. Hala iradokitzen digute Mitxelenaren (1985) hitz hauek: «mailegatzeko eskubidea ezker-eskuin eta zuzen-bihurri mugarik gabe geureganatua dugunez gero, ez ote dugu zilegi, emendailuz bezala, geure baitarik zer-edo-zer sortzea?»³⁰. Baina ezin gara purismoaren atzaparretan erori: «Genetikari buruzko unibertsitate liburu batean ez dela, kasu, «herentzia» hitza behin ere azaltzen. «Ondoretasuna» erabiltzen dela ordeztu. Betiko purismo antzuaren ildoak da hori. Eta ez da kasu bakana euskal idazkietan: lirdinga bat bezala erantsi zaio joera purista euskara idatziari.» (Atxaga, B. 2015). Irtenbidea ez da erraza, erdibidea aurkitzea ez da inoiz eta inon erraza. 1980ko hamarkadaren hasieran, hizkera juridikoaren mundutik bertatik egindako gogoetek lagun dezakete erdibide hori aurkitzen. Alor horretako euskarazko testuak (eta, esango genuke, gainerako alorretakoak berdin) baliozko eta eraginkorrak izatea nahi badugu, eta ez erabilerarik gabeko apaindura hutsak: a) edukiari begirunea gorde behar zaio; b) ezin ditugu setaz eta mendetasunez erdaratik maileguak hartu; c) kontzeptuak zehazki emateak ezin du ekarri erdaraz hitz bat erabiltzen den bakoitzean euskaraz ere hitz bat eta bera erabiltzea; gaztelaniazko terminoaren balio semantikoa aztertu behar da, euskal ordain egokia hautatzeko; d) mugak jartzea komeni da maileguak hartzeko; muga hori hizkuntza bera da, jatorrizko testua aztertu eta ulertu ondoren, beste biderik aurkitzen ez dugunean, orduan jo mailegura (Alvarez Mingo, 1982). Azken batean gorago jaso dugun Lozanoren (1997) irtenbide bera, Mitxelenaren hitzekin oso bat datorrena.

Termino eta egitura jakin batzuez

Artikulu honen lehen partean 1980 arteko hizkeraren garbizaletasunaz jardun dugu. Orain jaso ditugun gogoeta orokor horien ondoren, beha diezaiegun azken urteetako hizkera juridikoan erabili diren zenbait hitzi. Batzuek arrakasta izan dute; beste batzuk hor dabilta, baina, oraingoz, ez dute onarpen ofizialik.

- *Jarauntsi* ‘ondarea’. Ez da termino arrotza gaurko testuetan, baina Olabidek Bibliaren itzulpenean erabiltzeak Mitxelenaren (1959) haserrea piztu zuen: «...[halako hitzak] dakidanez, *hápax legómena* [literaturan behin bakarrik agertu direnak] dira. 1556ko *Refranes y sentencias*-en bakarrik ageri direnak. Ba ote zuten horrelako zaharren zaharrez aspaldi usteldu ziren hitzek oraingo liburu batean sartzeko eskubiderik? Ez, ene ustez, ezinbestean izan ezik, eta hemen ezinbestetik ez zegoela bistan dago»³¹.
- *Lerrun*. Lekukotasun ugari ditu *Egungo Testuen Corpusean* (ETC), eta agerraldi multzoa *Zuzenbide Corpusean* ere. Baina *Orotariko Euskal Hiztegira* jotzen badugu, ikusiko dugu Azkueren irakurketa okerraren ondorio dela. *Lerrunkan* (Pouvreauk dakarren forma), *lerrunean* zela uste izan zuen. Hortik sortu zuen berak *lerrun*, Pouvreauren *lerrunka* bazter utzita.

³⁰ Mitxelena, Koldo (1985): «Esaera zahar baten bigarren agerraldia» [*Euskal idazlan guztiak*, VI, Euskal Editoreen Elkar-
tea, 186. or.].

³¹ Mitxelena, K. (1959): «Itun Zar eta Berria. Olabide'tar Erraimun, S.J.», [*Euskal idazlan guztiak*, II, Euskal Editoreen El-
kartea, 112. or. eh.].

- *Egikaritu, egikaritze*. Testuetan lekukotasun handirik ez duen Larramendiren neologismo bat oinarri hartuta, indar handia hartu dute, hurrenez hurren, ‘ejercer, ejercitar’ eta ‘ejecución’ adieraz testu juridikoetan.
- *Tarteratu*. Erdara ispilu harturik sortu den adibide gordinen adibidetzat hartuko dugu aditz hau, lege-testuetan ez, baina testu akademikoetan aurki dezakeguna. Euskarazko esapidea kendu eta gaztelaniaren morrontzako kalkoa sortu da, inolako beharrik gabe. *Errekurtsua jarri, helegitea jarri* edo *aurkeztu*, baditugu euskal predikatuak kalkora jo beharrik gabe.
Zaila da hemen ere erdibidea, baina lege-testuetako eta testu akademikoetako terminoak moldatzean, on da Mitxelenaren (1959, 114) hitzak begien aurrean izatea: «Ez dago ezer txarrik [euskalki guztietatik hitzak hartzean], berriz ere, neurrian eginez gero. [Baina] (...) ez du hitz batek beste edozeinek adinbat esku (...) maizegi [hartzen ohi da] Euskal Herriko xokorik galduenetan jaso izan diren hitzak, baita aspaldidanik nonahi ahaztu zirenak ere, literaturan nahiz herri hizkuntzan zabalduago eta ezagunagoak direnen kaltetan». Izan ere, batzuetan badirudi lege-hizkera moldatzean badela joera tradizio urriena duten terminoak hautatzeko; arrotzenak diruditen aukerak hartzeko, euskal hiztun ikasiari zailen egokitzen zaizkienak. Nolanahi ere, hitz horiek sartu dira eta bidea egin dute gaurko hizkeran. Beude, beraz, bakean.
- *Erdal ‘izen + erreferentzia-adjektibo’ egiturako sintagmez*. Hitz elkartuen osaera eta idazkerari buruzko araua eman zuenean, Euskaltzaindiak (25. araua) ohartarazi zigun izen-elkarketa dela euskaraz tradiziozko bidea mota horretako sintagmen ordainak emateko, izena izaten delako adjektibo horien oinarri, nahiz batzuetan latinetik gertuagoko forman ageri adjektiboa (*mano/manual, leche/lechero, lácteo, economía/económico*, eta abar). «Bide honi eutsi beharko litzaioke orain ere, ahal den guztian, erdal hitz-moldaera arrotzen mendekoegi bihurtu nahi ez badugu» zioen Akademiak. Garrantzi handiko aukeria da, ugarriak baitira erreferentzia-adjektibodun sintagmak terminologian oro har, bai eta lege-hizkeran nahiz administraziokoan. Garai garbizaleetan ihes egin zieten *público* edo *jurídico* zuten erdal sintagmen ordainetan *publiko* edo *juridiko* maileguak hartzeari. Baina egia da zenbat eta urrunago egon oinarriko izena eta adjektiboa, zailagoa gertatzen dela izen elkartua erabiltzea. *Normal* adjektiboak, adibidez, galdua du *norma* izenarekiko lotura semantikoa. Nolanahi ere, asko ikertu da gai hau gure artean Euskaltzaindiak 1992ko araua eman ondoren. Corpusetako datuen argitan, badirudi erraz lerratzen garela azken urteetan erdal egitura mailegatzera, berez horretarako beharrik ez dugunean ere (Azkarate, 2019). *Konstituzio Auzitegi* eta *Herri Administrazio* defendatu zituen sutsuki Alvarez Mingok (1982). Hala erabiltzen dira gaur ere. Baina ‘konstituzio + izen’ edo ‘herri + izen’ egituren ondoren bizi dira, etxe erosoetan, *printzipio konstituzional, oinarri konstituzional* edo *zor publiko, jabari publiko, segurtasun publiko, administrazio publiko* eta abar. Beste batzuetan, bi formak eskaintzen dira: *atestatu polizial, polizia-atestatu*. Izen elkartuaren aldeko hautuak ere ez dira falta: *administrazio-espediente, administrazio-prozedura*, baina *organo administratibo*. Ertz ugari ditu auzi honek ere, orain arteko guztiek bezala. Baina aurreko puntuak bezala, gogoetatxo bat ekarri nahi izan dugu lerro hauek irakurtzen dituenari baliagarri izango zaiolakoan.

Eta, amaiera gisa, apokalipsia hurbil dagoela pentsaraz dezaketen hitzak eta aipamenak ekarri ditugula hona usteko du norbaitek, baina ez da hori izan, inola ere, gure asmoa. Pizgarri izan nahi dute, ikaragarri baino areago, azken parte honetako gogoeta guztiek. Pizgarri eta oroigarri nolako erantzukizuna dugun administrazio- eta lege-hizkera, modu batean edo bestean, esku artean du-

gunok; oroigarri, hizkuntzalariek eta adituek eskuz esku egin behar dutela bidea; oroigarri, administrazio- eta lege-hizkera ez dela adituei —administrazioko teknikariei edo legelariei— bakarrik dagokien arazoa, gauzak nola joan diren ikusita, hizkuntzaren osasuna bera kolokan jartzen duen hizkera baita.

Erreferentzia bibliografikoak

- Altonaga, K. (2015): *Back to Leizarraga*, Pamiela.
- Alvarez Mingo, P. (1982): «E.H.A.A.ko itzulpen-lana», *Revista Vasca de Administración Pública-Herri-Arduralaritzako Euskal Aldizkaria* 4, 121-142.
- Atxaga, B. (2015): Adierazpenak Felix Ibargutxiren «Zeinismoa dela eta» artikuluan, *Diario Vascon*, 2015.05.23 argitaratuta, Kepa Altonagaren *Back to Leizarraga* liburua dela-eta.
- Azkarate, M. (1997): «Hiztegi Batua», *Federiko Krutwig-i omenaldia*, Iker-10, Bilbo: Euskaltzaindia, 79-104.
- Azkarate, M. (2015): «Euskal lexikografia: kodetzetik garatzera», Fernandez B. eta P. Salaburu (arg.) *Ibon Sarasola, Gorazarre. Homenatge, Homenaje*, Bilbo: Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua, 119-136.
- Azkarate, M. (2019): «Adjektibo erreferentzial ala erreferentzi-adjektibo? Euskaltzaindiaren gomendio baten hariari tiraka», Epelde, I. eta O. Jauregi, *Bihotz ahots. M. L. Oñederra irakaslearen omenez*, Bilbo: Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua, 99-113.
- Azkarate, M. eta Lindemann, D. (2018): «Basque Lexicography and Purism», *International Journal of Lexicography*, Volume 31, Issue 2, 132-150 <https://doi.org/10.1093/ijl/ecy003>
- Berasategi, J. I. (1989): «La traducción jurídico-administrativa en el País Vasco». *Revista de Llengua i Dret* 13, 29-45.
- Cabré, T. (1992): *La terminologia. La teoria, els mètodes, les aplicacions*. Barcelona: Empúries.
- Elosegi, A. (2002): *Zuzenbideko euskal hizkera berezia. Lege-corpus itzuli bateko terminologiaren azterketa lingüistikoa eta terminologikoa*, Bilbo: Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua.
- Euzkadi'ko Agintaritzaren Egunerokoa (1936ko urritik 1937ko ekainera)*. *Hiztegia*, 1997, HAEE/IVAP.
- Freixá, J. (2002): *La variación terminológica; anàlisi de la variació denominativa en textos de diferent grau d'especialització de l'àrea de medi ambient*, Sèrie Tesis 3, Barcelona: Universitat Pompeu Fabra.
- Freixá, J. (2005): «La variación terminológica: ¿por qué y para qué?», *Meta: journal des traducteurs. Translator's Journal* 50 (4) <https://www.erudit.org/fr/revues/meta/2005-v50-n4-meta1024/019917ar/>
- Freixà, J. (2013): «Otra vez sobre las causas de la variación denominativa», *Debate terminológico* (9), 38-46.
- Gallastegi, C. (2005): «Itzulpen juridikoa eta euskararen historia», Larrazabl S. eta E. Oregi (koord.), *Hizkera juridikoa eta itzulpengintza: euskararen norabideak*. Bilbo: Deustuko Unibertsitatea-HAEE, 87-109.
- Haugen, E. (1987): *Blessings of Babel: Bilingualism and Language Planning: Problems and Pleasures*, Mouton de Gruyter.
- Lobera, G. (2005): «Gerra-ondoko itzulpen juridikoaren bilakaera», Larrazabl S. eta E. Oregi (koord.), *Hizkera juridikoa eta itzulpengintza: euskararen norabideak*. Bilbo: Deustuko Unibertsitatea-HAEE, 111-123.
- Lorente, M. (2013): «Terminología in vivo y variación funcional» Alberdi, X. eta P. Salaburu (arg.), *Terminologia naturala eta terminologia planifikatua euskararen normalizazioari begira*, Euskara Institutua, Euskal Herriko Unibertsitatea: Euskal Herriko Unibertsitateko Argitalpen Zerbitzua, 2-18.
- Lozano, J. (1997): «Zenbait argibide euskara juridikoaren gainean!», *Eleria* 1, 55-75.

- Mitxelena, K. (1954): EGAN, 1954 (2-4) 39 (Euskaltzaindiak Axularren *Geroren* argitaratzeari buruz) [*Euskal idazlan guztiak*, II, Euskal Editoreen Elkarte, 39].
- Mitxelena, K. (1961): «Euskal hiztegiak XVII-XVIIIgarren mendeetan», *Euskera* 2. Aldia, VI, 7-22 [Hemen berrargitaratua: Koldo Mitxelena. *Euskal idazlan guztiak VII*, Euskal Editoreen Elkarte. Erein, 47-66].
- Mitxelena, K. (1968): «Ortografia», *Euskera* XIII, 203-219.
- Mitxelena, K. (1978): «Sarrera gisa» Etxepare: *Olerkiak*, Txertoa, 1978, argitalpenari egindakoa [1988 *Euskal idazlan guztiak*, V, Euskal Editoreen Elkarte, 26].
- Mitxelena, K. (1984): «*Hauta-Lanerako Euskal Hiztegia*-ren aurkezpena», I. Sarasola, *Hauta-Lanerako Euskal Hiztegia* [Hemen berrargitaratua: Koldo Mitxelena. *Euskal idazlan guztiak VII*, Euskal Editoreen Elkarte. Erein, 25-45].
- Pagola, I. (2005): *Neologismos en la obra de Sabino Arana Goiri*, Iker-18, Bilbo: Euskaltzaindia.
- Sarriegi, A. (1985): «Administrazioa eta euskara: itzultzaileak eta arduradunak», *SENEZ: Itzulpen eta terminologiazko aldizkaria*, 4, 11-29 <https://eizie.eus/eu/argitalpenak/senez/19851201/sarriegi>
- Villasante, L. (1968): «Antziñako euskal hitzen formaz», *Euskera* XIII, 175-201.
- Zuazo, K. (2000): *Euskararen sendabelarrak*, Irun: Alberdania.

(Euskal) Bozak: erdikoak, inkoatiboak eta arazleak¹

ANE BERRO

Deustuko Unibertsitatea

BEATRIZ FERNÁNDEZ

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

Atarikoak

Argumentuak gaineratzeko mekanismoak, batez ere egitura aplikatiboak, euskal sintaxilarien jomuga izan dira laurogeita hamarreko hamarkadaz geroztik, baina argumentuak gutxitzeko mekanismoek, hots, *iragangaiztearekin* (ingelesez, *detransitivization*) zerikusia duten egiturek, salbuespenak salbuespen, ez dute horrenbesteko arretarik jaso euskaraz. Eremu ezezagunago honetan sakondu nahian, lan honetan, iragangaiztearekin zerikusia duten egitura *inpersonal* deitutarikoak aztertuko ditugu, hots, predikatu iragankor baten gainean eraiki arren, iragangaitzak diren egiturak, hala nola, *liburuak erraz saltzen dira* bezalakoak (Ortiz de Urbina 1989, 1991, 2003, 2011-2019; Albizu 1997, 2001). Honelako egiturak berriki aztertu ditu Urrestarazuk (2019), eta Ahn & Sailor-en (2014) lanean oinarrituz, *erdiko boz*eko egituratzat hartu ditu, *boz aktiboko* egituratik bereiziz (e.g. *liburuak erraz saltzen dituzte*). Lan honetan, Urrestarazuren lanari segida emanez erakutsiko dugunez, euskararen erdiko egiturak ez zaizkie bete-betea lotzen ingelesez edo alemanez erdikoei egotzi izan zaizkien ezaugarriei, eta antzekotasun handiagoa dute hizkuntza erromantzeetako kideko egiturekin. Alde horretatik, Ackema & Schoorllemmer-ek (2005) erdiko egituren artean bereizten dituzten bi motetatik, hots, 1-motako erdikoak eta 2-motako erdikoak, euskararenak 2-motakoak direla dirudi. Erdikoak, bestalde, iragangaitzak diren aldetik *inkoatibo*en oso antzekoak dira (*hil da*), nahiz eta ez diren egitura bera, eta *arazle*etatik (*hil du*) bereizten dira —inkoatibo/arazle alternantziarako, ikus Oihartzabal (2003), Ormazabal (2008), eta Berro *et al.* (2018)—. Hori dela eta, lan honetan hiru egiturak jarriko ditugu harremanetan eta aztertuko dugu zein diren inkoatibo/arazle alternantzianez ezik, erdikoetan ere ager daitezkeen predikatuak. Ikusiko dugunez, aldaera arazlea duten predikatuak beti ager daitezke erdikoetan; aldaera inkoatiboa baino ez dutenak, berriz, inoiz ez.

Lanak ondoko atal hauek izango ditu: 1. atalean, erdiko egiturak zedarritzatuko ditugu; 2. atalean, ingelesezko erdikoak eta euskararenak konparatuko ditugu beste hizkuntza batzuen artean; 3. atalean, erdikoak, inkoatiboak eta arazleak jarriko ditugu harremanetan predikatu jakinak aztertuz; eta 4. atalak bilduko ditu ondorio nagusiak.

¹ Lan honen egileek Eusko Jaurlaritzaren diru-laguntza jaso dute IT1169-19 ikerketa proiektuaren bidez eta Espainiako Gobernuko Zientziaren, Berrikuntzaren eta Unibertsitateen Ministerioarena PGC2018-096380-B-I00 ikerketa proiektuaren bidez.

Lan hau Pello Salaburu adiskideari eskaini nahi diogu. Salaburu dago euskaldunok eskura ditugun bi gramatika egitasmo handiren atzean. Batetik, 1997az geroztik Euskaltzaindiaren Gramatika Batzordearen burua den aldetik, bere zuzendaritzapean idatzi da azken bi hamarkadetan *Euskal Gramatika: Lehen Urratsak*(*EGLU*) (Euskaltzaindia 1991-2011). Bestetik, bere ekimenez eman zen argitara 2011an *Sareko Euskal Gramatika* (*SEG*) (Salaburu, Goenaga & Sarasola 2011-gaur arte), sarean euskaraz dagoen gramatika bakarra. Euskal sintaxilariok zordun gatzazkio, beraz, eta lan hau gure esker onaren adierazgarria da.

1. Erdiko boza zedarriztatzen

Euskaraz badira zenbait egitura, predikatu iragankor baten gainean eraikitzen diren arren, iragangaitzat har daitezkeenak. Konpara bitez (1a) eta (1b).

- (1) a. Argialetxeek liburuak erraz saltzen dituzte.
- b. Liburuak erruz saltzen dira.

(1b) egitura *inpersonaltzat* hartu izan da euskal sintaxi modernoan (Ortiz de Urbina 1989, 1991, 2003, 2011-2019; Albizu 1997, 2001). Alabaina, deitura beraren azpian askotariko egiturak jaso direnez, guztien artean sintaktikoki *iragangaitzearekin* (ing. *detransitivization*) zerikusia dutenak, (1b) bezalakoak alegia, gainerakoetatik bereizi nahi izan ditu Urrestarazuk (2019). Hala, *bozarekin* lotu eta *erdiko bozeko* egiturak direla aldarrikatu du, Sailor & Ahn-ek (2010) eta Ahn & Sailor-ek (2014) ingeleseko egindako proposamenetan oinarrituz. Egile hauek boz aktiboa bereizten dute boz *ez-aktibotik*, eta *ez-aktiboan* bere aldetik boz pasiboa erdiko bozetik. Ingeleseko erdiko bozeko egitura (2b) konpara daiteke boz aktibokoarekin (2a) (Keyser & Roeper 1984: 381) eta boz pasibokoarekin (2c) (Ahn & Sailor 2014) —.

- | | |
|------------------------------------|---------------|
| (2) a. Someone bribed bureaucrats. | (boz aktiboa) |
| b. Bureaucrats bribe easily. | (erdiko boza) |
| c. Bureaucrats are bribed easily. | (boz pasiboa) |

Erdiko bozeko egiturak zedarriztatzeko, bost ezaugarri nagusi aipatzen dituzte Ahn & Sailor-ek (2014: 4). Lehenik eta behin, erdiko bozeko egiturako subjektua (*bureaucrats* ‘burokratak’) predikatuaren (*bribe* ‘erosi’) barne argumentua da. Bigarrenik, erdiko bozeko egiturak ezin daitezke pasibo bihurtu. Hirugarrenik, kanpo argumentua sintaxitik kanpo dago. Laugarrenik, kanpo argumentu inplizitu bat dute, eta bosgarren eta azkenik, predikatu arazle, *ez-estatibo* baten gainean eraikitzen dira.

Euskarari dagokionez, Urrestarazuk (2019) dio boz aktiboko zein erdiko bozeko egiturak daudela, (3a) eta (3b) hurrenez hurren —baita (1a) eta (1b)—, eta (3b, 1b) bezalakoek Ahn & Sailor-ek erdiko egiturei egozten dizkieten lehen lau ezaugarriak betetzen dituztela —itzuliko gara geroago (2. atala) bosgarren ezaugarriara, Urrestarazuk aztertzeke uzten duena—.

- | | |
|---|---------------|
| (3) a. Baten batek burokratak erosten ditu. | (boz aktiboa) |
| b. Burokratak erraz erosten dira. | (erdiko boza) |

Bestalde, ingelesez ez bezala zeinetan (2c) bezalako egiturak dauden, euskaraz boz pasiborikez dagoela dirudi, (3b) erdikoa ez, baizik eta pasiboa ez bada behintzat. Izatez, euskal sintaxi

modernoan pasibotzat hartu diren egiturak, (4) bezalakoak esate baterako (Euskaltzaindia 1991, 1993; Salaburu & Lakar 2005 beste batzuen artean), ez zaizkie lotzen hizkuntza tipologiak zein gramatika teoriak pasibotzat jotzen dituzten egiturei. Halakoak, beraz, alde batera utziko ditugu —ikus Ortiz de Urbina & Uribe-Etxebarria (1991) azterbide baterako—.

- (4) a. Orduan cumplitua izan cen Jeremiaz profetaz errana izan cena
b. Eta an presentatu cioten perlesiaz artu bat goatze batean etzana

(Etxenike, *San Mateoren Ebanjelioa*, Salaburu & Lakar 2005: 146)

Aise ikus daiteke ingelesezko (2a) eta euskarazko (3a) egiturak harremanetan daudela. Bietan predikatu iragankor bat dago (*bribe/erosi*), eta espero bezala, bina argumentu ageri dira, hots, subjektu eta osagarria, egile eta gai rol tematikoa jaso dutenak, hurrenez hurren. Beste alde batetik, (3b) egitura printzipioz ingelesezko boz ez-aktiboko bi egituretako edozeinekin jar liteke harremanetan: edo (2b) erdiko bozeko egiturarekin eta (2c) boz pasibokoarekin. Izan ere, bata zein besteak iragangaiztearekin dute zerikusia: ez dira berez iragangaitzak, predikatua iragankor baten gainean eraikitzen direlako, baina sintaktikoki iragangaiztu egin dira, argumentu bakarreko egiturak sortaraziz. Hori bera esan daiteke baita (3b) egiturari dagokionez: argumentu bakarra dago (*burokratak*), nahiz eta predikatua (*erosi*) iragankorra den. Egitura iragangaiztu hauetan —(2b,c) eta (3b)—, bestalde, argumentuetako bat *goratu* (ing. *promote*) dela dirudi, izatez barne argumentua delako subjektu gunean ageri dena, eta beste argumentuetako bat, zehazki kanpo argumentua, berriz, *beheratu* (ing. *demote*) edo are gehiago, ezabatu. Boz pasiboko (2c) egituran, beheratu den subjektu logikoa ez da esplizituki ageri, baina adjuntu baten bitartez berreskuratu liteke,² *by* preposizioa buru duen sintagma batekin (*by managers*) (5a). Erdiko bozeko egituran, aldiz, badago kanpo argumentu inplizitu bat, baina boz pasiboko egituran ez bezala ezin daiteke sintaktikoki *by*-sintagma batez burutu (5b) (Keyser & Roeper 1984: 406). Hori bera esan daiteke (3b) egituraz den bezainbatean, Urrestarazuk dioenez: kanpo argumentua ezin daiteke sintaktikoki berreskuratu (5c). Alde honetatik behintzat (3b) egitura erdiko bozetik gertuago dago boz pasibotik baino.

- (5) a. Bureaucrats are bribed easily by managers. (boz pasiboa)
b. *Bureaucrats bribe easily by managers. (erdiko boza)
c. *Burokratak erraz erosten dira managerrez/managerrekin. (erdiko boza)

Gainerakoan, Urrestarazuren ustez (2019), Ahn & Sailor-ek (2014: 4) erdiko bozeko egiturak zedarritzatzeko aipatzen dituzten ezaugarri nagusiak betetzen dituzte. Hasteko, esan berri dugun bezala, erdiko bozeko egiturako subjektua predikatuaren barne argumentua da; beste alde batetik, kanpo argumenturik ezin daiteke sintaktikoki inola berreskuratu, esate baterako, komitatibo edo instrumental baten bidez, eta sintaxitik kanpo dago beraz; azkenik, kanpo argumentu inplizitu bat dago (ikus Keyser & Roeper 1984 eta Bhatt & Pancheva 2005 eztabaidarako). Beheko taulan ikus daitezke bi hizkuntzen ezaugarri nagusiak erdiko bozeko egiturei dagokienez.

² Ikus, alabaina, Siewierska (2013).

1. taula

Erdiko egituren ezaugarri nagusiak:
ingeleza (Anh & Sailor 2014) vs. euskara (Urrestarazu 2019)

Erdikoegiturak	ingeleza	euskara
subjektua, barne argumentua	✓	✓
pasiborik ezin osatu	✓	—
kanpo argumentua sintaxitik at	✓	✓
kanpo argumentu inplizitua	✓	✓

2. Erdiko boza: ingeleza, euskara eta beste

Ezaugarriok ikusirik, ez dirudi printzipioz eragozpenik dagoenik (1b) eta (3b) bezalako boz ez-aktiboko egiturak erdikoak direla irizteko. Alabaina, ingelesez bederen erdiko bozeko egiturek badituzte beste zenbait ezaugarri euskaraz itxuraz betetzen ez direnak. Lehenik eta behin, Keyser & Roeper-en (1984) arabera —aurrerantzean, K&R (1984)—, erdiko bozeko egiturak generikoak dira eta ez dute, beraz, gertaera zehatzik adierazten (*ez-ebentiboak* dira); ondorioz, ezin daitezke denbora zehatz bateko gertaerekin lotu —ikus, alabaina, Sailor & Ahn (2010: 6-7)—. Hala, (6a) ezgramatikala da (K&R 1984: 384). Oso bestela, euskaraz kideko egiturak gramatikalak direla dirudi (6b). Kontuan izan, dena dela, (6b) adibideak aspektu burutua duela, berebiziko garrantzia duena egitura ebentiboa edo gertaerazkoa izateko.

- (6) a. ?Yesterday, the mayor bribed easily, according to the newspaper.
b. Atzo udalen osaketan, alkateak erosi ziren, egunkariaren arabera.

Bigarrenik, gertaerarik adierazten ez dutenez, ezin daitezke aginterazko perpausetan ageri (7a) ezta egitura progresiboetan ere (7b). Alde horretatik, predikatu estatiboak bezalakoak dira. Erdiko bozeko adibideak dira (7a,b) eta *know* ‘jakin’ aditz estatibokoak (7c,d) (K&R 1984: 385).

- (7) a. *Bribe easily, bureaucrat!
b. *Bureaucrats are bribing easily.
c. *Know the answer, John!
d. *John is knowing the answer.

Honetan, ez dirudi oso ezberdin jokatzten dutenik ingelesezko erdikoak eta hemen aztergai ditugun egiturak: murriztapen berak dituzte bai aginterari (8a) bai progresiboari (8b) dagokionez, estatiboek bezalaxe —progresiboaren azterketarako, ikus, Laka (2003, 2006)—. Jakina, (8a) eta (8b) gramatikalak dira, erdiko egitura gisa interpretatzen ez badira.

- (8) a. (*)Eros(i) erraz, burokrata!
b. (*)Burokratak erraz erosten ari dira.
c. *Jakin ezazu erantzuna, Jon!
d. *Jon erantzuna jakiten ari da.

Hirugarrenik, ingelesez, erdiko bozeko egiturak adberbioekin batera agertu behar dira gramatikalak izateko (K&R 1984: 385) —ikus baita Sailor & Byron (2010: 9, 10 oin-oharra)—. Alde horretatik ere, euskaraz ez dirudi halako murriztapenik dagoenik.

- (9) a. *Bureaucrats bribe.
b. Burokratak erosten dira.

Ezaugarriok laburbilduta daude beheko taulan, Urrestarazuren (2019) tipologia osatzera datozenak:

2. taula

Erdiko egituren ezaugarriak:
ingeleza (K&R 1984) vs. euskara

Erdikoegiturak	ingeleza	euskara
ez-ebentiboak	✓	✗ / ✓
aginterarik ez	✓	✓
progresiborik ez	✓	✓
adberbioek lagunduta	✓	✗ / ✓
oharmenezko aditza + perpaus txikia	✓	—

Laburbilduz, euskaraz, Ahn & Sailor-ek (2014) ingelesezko erdiko egiturez aipatzen dituzten ezaugarri nagusiak betetzen badira ere, badira beste zenbait ezaugarri (K&R 1984) betetzen ez direnak. Zorionez, ez da euskara bakarrik ingelesezko erdikoek tipologia horretatik aldentzen dena. Beste hizkuntza batzuetan, hala nola, hizkuntza erromantzeetan, ez dirudi erdikoek ingelesezko kideko egituren ezaugarri berak dituztenik. Adibidez, Ahn & Sailor-ek (2014) lanean, erdikoei egozten dizkieten ezaugarrien artean diote egitura hauek predikatu arazle, ez-estatibo baten gainean eraikitzen direla (ikus 1. atala), eta nagusiki hala dela dirudi. Alabaina, hizkuntza erromantzeetan —eta euskaraz ere hala da—, erdikoak ez dira predikatu arazleetara mugatzen, Hale & Keyser (1988: 51-52) aspaldi samar konturatu ziren bezala. Adibidez, *see* oharmenezko aditza ezin daiteke ingelesezko erdiko bozan erabili (10a), baina gaztelaniaz bai (10b) —eta euskaraz ere bai (10c)—.

- (10) a. *Planets see easily.
b. Los planetas se ven facilmente. (Hale & Keyser 1988: 51)
c. Planetak erraz ikusten dira.

Eta frantsesak ere gaztelaniaz eta euskaraz bezalaxe jokatzen duela dirudi:

- (11) La Tour Eiffel se voit de loin. (Ackema & Schoorllemmer 2005)

Beharbada, Hale & Keyser-ek (1988: 51) dioten bezala, hizkuntza erromantzeetako erdiko egituretan ager daitezkeen predikatu sorta handiagoa da ingelesezko predikatuena baino, eta antzekotasuna, erdiko egituren adibide kanonikoetan baino ezin daiteke ikusi.

- (12) a. John cut the bread.
 b. The bread cuts easily.
 c. Juan cortó el pan.
 d. El pan se corta fácilmente. (Hale & Keyser 1988: 51)

Euskaraz ere hala dela gaineratuko genuke guk:

- (13) a. Jonek ogia zatitu zuen.
 b. Ogia erraz zatitzen da.

Beste alde batetik, erdiko egituretan modifikatzaile adberbialak agertu beharra dagoela eta, ez dirudi hizkuntza guztietan hala denik. Frantsesez, adibidez, ez dago halako murriztapenik:

- (14) a. Ce papier se lave.
 b. Cette racine se mange.
 c. Cette solution se discute.

(Zribi-Hertz 1982: 348, apud Ackema & Schoorllemmer 2005)

Ezta gaztelaniaz eta euskaraz ere:

- (15) a. Este papel se lava.
 b. Esta raíz se come.
 c. Esta solución se discute.
 (16) a. Paper hau garbitzen da.
 b. Sustrai hau jaten da.
 c. Irtenbide hori eztabaidatzen da.

Honi dagokionez, beraz, hiru hizkuntzok ingelesetik bereizten dira:

- (17) a. *This paper washes.
 b. *This root eats/consumes.
 c. *That solution debates. (Ackema & Schoorllemmer 2005)

Honek guztiak eta orain aipatuko ez ditugun beste ezberdintasun batzuek eraman dituzte Ackema & Schoorllemmer (2005) erdiko egituren artean bi mota bereiztera, hots, 1-motako erdikoak, ingelesaren tipologiari lotzen zaizkionak, hala nola, alemana eta nederlandera, eta 2-motako erdikoak, hizkuntza erromantzeetan jasotzen direnak adibidez —baita grezieraz, norvegieraz, errusieraz eta serbiera-kroazieraz ere—, eta itxura guztiaren arabera, euskaraz ere baditugunak.

Gainerakoan, argitzeke gelditzen da nola daitekeen euskaraz erdikoaren egitura itxuraz bera baliatuz, adieraz daitezkeen ebentiboak diren adibideak, hala nola, (6b). Erdiko egiturak estatiboak dira, ez-ebentiboak alegia. Alde horretatik, (6b) harrigarria da, gertaera bat adierazten duelako (atzoko udal osaketan, alkateak erosi zirela alegia). Momentuz eta geroko lanaren zain, gure susmoa da euskaraz erdikoak adierazteko erabiltzen dugun estrategia bera eskura dagoela, beharbada erdikoak ez diren egiturak ere adierazteko.

Bestetik, erdiko egituren ezaugarri nagusietako bat bada kanpo argumentu inplizitu bat ego-
tea, orduan erraz uler daiteke zergatik ezin daitekeen (19a) inkoatiboa izan: inkoatiboez ez dute
kanpo argumentu inpliziturik onartzen.

Gainerakoan, ez da harrigarria aldaera arazlea dugun kasuetan, erdiko egitura ere eskura iza-
tea. Izan ere, goian (1. atala) erakutsi dugunez, erdiko egiturak predikatu arazle baten gainean
eraiki daitezke (Ahn & Sailor-ek 2014), eta 1-motako erdikoak dituzten hizkuntzetan predikatu
arazle baten gainean baino ez (Ackema & Schoorllemmer 2005). Hortaz, euskaraz arazleak dituz-
ten predikatuetatik espero izatekoa da erdiko egituretan agertzea.

Dena dela, oinarrian predikatu arazle bat izatea ez dirudi nahikoa denik erdiko egitura bat
sortzeko. Adibidez, *atxilotu* ere, *legeztatu* bezala, aldaera arazlea duen predikatua da (20b), nahiz
eta aldaera inkoatiborik (20a) ez duen:

- (20) a. *Jon atxilotu da. (inkoatiboa)
b. Poliziak Jon atxilotu du. (arazlea) (Berro *et al.*, 2018: 18)

Eta printzipioz esango genuke *atxilotu* predikatuarekin erdiko egiturak sor daitezkeela:

- (21) Gazteak erraz atxilotzen dira.

Alabaina, corpusetan ez dugu halakorik aurkitu eta ez dakigu arrazoia zein den. Susmoa
dugu, nahiz susmoa besterik ez den, (18a,b,c) eta (19a) adibideetan barne argumentua bizigabea
izateak —*sexu berekoen ezkontzak*, *bikotekide homosexualen arteko ezkontzak* eta abar— errazten
duela erdiko bozaren irakurketa, nekezago aurkitzen dena, nahiz eta ezgramatikala ez den, barne
argumentua biziduna denean —*gazteak*— (21) adibideetan den bezala. 3.3 atalean ikusiko dugu-
nez, barne argumentuaren biziduntasunak badu eragina erdiko egituraren interpretazioan.

Beste alde batetik, biziduntasunarekin batera, barne osagarriaren zehaztasunak ere eragina du
atxilotu predikatuaren erdiko egituraren zilegiztatzean. Izan ere, barne argumentua entitate zeha-
tza denean, izenordain zehatza (22) adibidez, erdiko egitura ezgramatikala da; gramatikala da, or-
dea, barne osagarria zehaztugabea denean (23).

- (22) *Zuek erraz atxilotzen zarete.

- (23) Gazte asko erraz atxilotzen dira.

Badirudi, beraz, erdikoetan murriztapenen bat edo dagoela subjektuari dagokionez. Hori ez
da, noski, euskarara mugatzen, italieraren ondoko adibide honek erakusten duenez:

- (24) *Io mi trasporto fácilmente.
ni BIHUR garraiatu erraz
'Ni garraiatzen erraza naiz'. (Cinque 1988: 565)

Ackema & Schoorllemmer-ek (2005) euren buruari galdetzen diote ea erdiko egituretako
subjektuak hirugarren pertsonakoak izatea ote den subjektuari eragiten dion *biziduntasunaren*
kontrako baldintza bat (*anti-animacy condition* diote, hitzez hitz). Jakina, halako baldintzarik edo
murriztapenik ez dago inkoatiboetan. Arretaz arakatu beharrekoa da ustezko baldintza hori —er-
diko egituretako subjektua hirugarren pertsonakoa delako eztabaidarako, ikus Urrestarazu (2019)
eta bere aipamenak—.

3.2. *Erdikoak eta aldaera inkoatiboa baino ez duten predikatuak*

Erdikoek kanpo argumentu inplizitu bat behar dutela erakusten duen beste ebidentzia bat da aldaera inkoatiboa besterik ez duten predikatuak erdikoetan ezin agertzea. Argi ikusten da hori *loratu* bezalako predikatuetan, hau da, *barrutik arazitako gertakariak* (ing. *internally caused event*) adierazten dituzten horietan (Oihartzabal 2003; Berro *et al.*, 2018).

- | | |
|---|--------------|
| (25) a. Gereziondoak loratu dira maiatzeko eguzkiarekin | (inkoatiboa) |
| b. % Maiatzeko eguzkiak gereziondoak loratu ditu | (arazlea) |
| c. *Maiatzean hainbat gereziondo loratzen dira | (erdikoa) |

Beste hainbat hizkuntzatan gertatzen den bezala, barrutik arazitako gertakariak adierazten dituzten predikatuak nekez onar dezakete aldaera arazlea —zentzu metaforikoan ez bada behintzat (Berro *et al.*, 2018)—. Orobat, ezin dute erdiko egiturarik sortu, kanpo argumentu inpliziturik ez dutelako. Horregatik, (25c) ezgramatikala da erdiko interpretazioa emanaz gero.

Beraz, erdikoak eta inkoatiboak iragangaitzak izan arren, ez dirudi sintaktikoki egitura bera izan behar dutenik. Erdikoek kanpo argumentu inplizitua dute; inkoatiboek, aldiz, ez. Besterik da kanpo argumentu inplizitu hori nola inplementatzen den, sintaxitik kanpo badago. Nolanahi ere den, inkoatiboetan ez bezala, kanpo argumentu hori egon badago.

3.3. *Erdikoak eta aldaera inkoatibo zein arazlea dituzten predikatuak*

Azkenik, aldaera arazlea zein inkoatiboa duten predikatuak ere erdiko bozeko egiturak sor ditzakete. 3.1 atalean ikusi dugunez, aldaera arazlea duten aditzak erdiko bozeko egituretan ager daitezke oro har; hortaz, ez da harritzekoa aldaera arazlea zein inkoatiboa dituzten predikatuak ere zilegi izatea erdiko bozan.

- | | |
|-----------------------------------|--------------|
| (26) a. Kristal asko apurtu dira. | (inkoatiboa) |
| b. Kristal asko apurtu dituzte. | (arazlea) |
| c. Kristalak erraz apurtzen dira. | (erdikoa) |

Dena dela, erdikoak eta inkoatiboak ez dira ausaz hautatzen. Izan ere, inkoatibo/arazle alternantzia parte hartzen duten predikatu batzuetan, bederen, berriro ere barne argumentuaren biziduntasunak zerikusia du erdikoaren eta inkoatiboaren arteko hautaketan. Har dezagun *ezkutatu* aditza adibide gisa.

- | | |
|--|--------------|
| (27) a. Umeak eskailera azpian ezkutatu dira. | (inkoatiboa) |
| b. Andereñok umeak ezkutatu ditu eskailera azpian. | (arazlea) |

Bi adibideotan, barne argumentua biziduna da (*umeak*), baina bizigabea denean (28), aldaera inkoatiboa ez dago eskura, eta bere ordeaz, erdiko egitura baino ezin daiteke eskuratu.

- | | |
|--|-----------------------|
| (28) a. Dokumentuak erraz ezkututzen dira. | (*inkoatiboa/erdikoa) |
| b. Dokumentuak erraz ezkutatu dituzte. | (arazlea) |

Hala, (28a) ezin daiteke inkoatibotzat hartu, dokumentuak berez ezkutatu ezin daitezkeelako, eta kanpo argumentu inplizitu bat (politikariak edo beste) dagoela ulertu beharra dagoelako. Hau da, egitura iragangaitza delarik, predikatua zein den, biziduntasunik ezak zokoratu dezake aldaera inkoatiboa eta erdiko boza bideratu.

Dena dela, ez du horrek esan nahi erdiko boza bideratzeko, barne argumentua nahitaez bizigabea izan behar denik.

- (29) a. Gaitz psikologikoak erraz ezkututzen dira. (erdikoa)
- b. Gaitz psikologikoak dituzten gaixok erraz ezkututzen dira. (erdikoa)

Nolanahi ere den, corpuseko datuetara begira berriro ere, bizigabeak dira barne argumentuak erdiko bozeko egitura askotan, EPGko adibideok erakusten duten bezala.

- (30) a. Jaunak beti ditu aurrean gizakien jokabideak, ez dira ezkututzen haren begietatik (*Elizzen arteko biblia*, Askoren artean, Si 17,15)
- b. Gure sukaldean zeuden familiako gainerako guztiak, mahaiaren inguruan, marrubiei hostoak eta txortanak kentzen, hori ari ziren, behintzat, ustez egiten, zeren marrubi gehienak ahoan barrena ezkututzen baitziren, ontzira joan ordez (*Anne Franken egunkaria*, A. Frank/J. Zabaleta)
- c. Ez, gorde egin beharko genukeen, ezkutatu egin beharko zatekeen: toki sekretu eta ilunen batean ezkutatu (*Eraztunen Jauna III*, J.R.R. Tolkien/A. Otxoa, 101 orr.)

Badirudi, beraz, barne argumentuaren biziduntasunik ezak (edo nahiago bada, bizigabetsunak) erdiko bozaren irakurketa errazten duela (eta zenbait kasutan behintzat inkoatiboa oztopatzen), nahiz eta, ikusi dugunez, ez diren halabeharrez bizigabeak izan beharerdiko egituretako barne argumentuak.

4. Ondorioak

Lan honetan, boza duten hiru egitura aztertu ditugu: erdikoak, inkoatiboak eta arazleak. Urrestarazuk (2019) erakutsi duenez, euskal syntaxian egitura inbertsonalak deiturikoen artean, bereziki iragangaizteari esker sortutakoek erdiko egiturei egotzi izan zaizkien ezaugarri berberak dituzte (Ahn & Sailor 2014). Nolanahi ere, euskaraz erdiko bozeko egiturak ez datoz beti bat ingelesezko erdiko egituren ezaugarriekin; izan ere, euskaraz ez dira halabeharrez predikatu arazleak erdiko boza sortarazten dutenak; erdiko bozak ez du nahitaez modifikatzaile adberbialik behar, eta azkenik, aspektu markapena aldatzen bada, interpretazio episodikoa ere izan dezake, ingelesez ezinezkoa dena. Badirudi, beraz, euskal erdiko bozak hizkuntza erromantzeetan aurkitzen diren erdikoen antz handiagoa duela ingelesarena baino, eta Ackema & Schoorllemmer-en (2005) tipologian oinarrituz, 2-motako erdiko delatzen (frantsesa, gaztelania...), eta ez 1-motakoa (ingelese, alemana...). Azkenik, erakutsi dugu euskarazko erdiko egitura aldaera arazlea duten predikatuen gainean eraikitzen dela, izan inkoatibogabea (e.g. *legeztatu*) izan inkoatiboduna (e.g. *apurtu*), eta ez, ordea, inkoatiboa besterik ez dutenen gainean (e.g. *loratu*). Horrekin lotuta ikusi dugunez, predikatu batzuetan barne osagarriaren biziduntasunak eragina du bozaren interpretazioan; adibidez, *ezkutatu* predikatuan, barne osagarria bizigabea bada, erdiko interpretazioa nagusitzen da, interpretazio inkoatiboaren kaltetan.

Aipamenak

- Ackema, Peter & Maaike Schoorllemer. 2005. Middles. In Martin Everaert & Henk Van Riemsdijk (ed.), *The Blackwell Companion to Syntax*. vol. 3. Blackwell Publishing. 131-203.
- Ahn, Byron & Craig Sailor. 2014. The Emerging Middle Class. *Proceedings from the 46th Annual Meeting of the Chicago Linguistic Society (CLS 46)*. 357-371.
- Albizu, Pablo. 1997. En torno a la inacusatividad sintáctica de las construcciones impersonales del euskera. *Interlingüística* 8. 13-18.
- Albizu, Pablo. 2001. Sobre la distribución sintáctica de las formas finitas del verbo vasco: condicionamiento léxico y sintáctico. *Julio Urkixo Euskal Filologi Mintegiaren Urtekariaren Gehigarriak (ASJU)* 35 (1). 65-106.
- Bhatt, Rajesh & Roumyana Pancheva. 2005. Implicit arguments. In Martin Everaert & Henk Van Riemsdijk (ed.), *The Blackwell Companion to Syntax*. vol. 3. Blackwell Publishing. 558-588.
- Berro, Ane, Beñat Oihartzabal & Beatriz Fernández. 2018. Inkoatibo/arazle alternantzia euskaraz eratorzen. In Beatriz Fernández & Ane Berro (ed.), *Euskaratik giza hizkuntzara. Jon Ortiz de Urbina gora-zarre. Fontes Linguae Vasconum* 126. Nafarroako Gobernua. 99-122.
- Cinque, Guglielmo. 1988. On *si* constructions and the theory of arb. *Linguistic Inquiry*. 19: 521-582.
- Euskaltzaindia. 1991. *Euskal Gramatika. Lehen Urratsak I (EGLU I)*. Bilbo: Euskaltzaindia.
- Euskaltzaindia. 1993. *Euskal Gramatika Laburra: Perpaus Bakuna (EGLA)*. Bilbo: Euskaltzaindia.
- Hale, Kenneth & Samuel J. Keyser. 1988. Explaining and constraining the English middle. In Carol L. Tenny (ed.), *Studies in Generative Approaches to Aspect. [Lexicon Project Working Papers 24]*. Cambridge, MA: Center for Cognitive Science, MIT. 41-57.
- Keyser, Samuel Jay & Thomas Roeper. 1984. On the middle and ergative constructions in English. *Linguistic Inquiry* 15. 381-416.
- Laka, Itziar. 2003. Ari progresiboaz: euskararen kasu markak. In Pablo Albizu & Beatriz Fernández (ed.), *Euskal gramatika xxi. mendearen hasieran: arazo zaharrak, azterbide berriak*. UPV/EHU. 111-132.
- Laka Itziar. 2006. Deriving Split-ergativity in the progressive: the case of Basque. In Alana Johns, Diane Massam, & Juvenal Ndayiragije (ed.), *Ergativity: Emerging Issues*. Dordrecht/Berlin: Springer. 173-195.
- Oihartzabal, Beñat. 2003. Lexical causatives and causative alternation in Basque. *ASJuren Gehigarriak* 46. 223-253.
- Ormazabal, Javier. 2008. Kausatibo aldizkatzeak euskaraz eta inguruko hizkuntzetan. In Xabier Artiagoitia & Joseba A. Lakarra (ed.), *Gramatika jaietan Patxi Goenagaren omenez. ASJuren Gehigarriak* 51. 643-663.
- Ortiz de Urbina, Jon. 1989. *Parameters in the Grammar of Basque*. Dordrecht: Foris.
- Ortiz de Urbina, Jon. 1991. Deskribapenetik azalpenetara euskal gramatikan. In Adolfo Arejita (ed.), *Euskal Sintaxiaz*. Bilbo: Labayru. 11-27.
- Ortiz de Urbina, Jon. 2003. Impersonal clauses. In José Ignacio Hualde & Jon Ortiz de Urbina (arg.), *A Grammar of Basque*. Berlin: Mouton de Gruyter. 572-591.
- Ortiz de Urbina, Jon. 2011-2019. Aditz inbertsonalak. In Pello Salaburu, Ibon Sarasola & Patxi Goenaga (ed.), *Sareko Euskal Gramatika(SEG)*. [Interneten eskuragarri: http://www.ehu.eus/seg/morf/5/12/2/2#predikatu_inbertsonalak; azkenbisita: 2019-06-21].
- Ortiz de Urbina, Jon & Myriam Uribe-Etxebarria. 1991. Participial predication in Basque. *Julio Urkixo Euskal Filologi Mintegiaren Urtekariaren Gehigarriak (ASJU) [Memoriae L. Mitxelena Magistri Sacrum]* 14 (2).993-1012.

- Sailor, Craig & Byron Ahn. 2010. The Voices in our Heads: the VoiceP in English. In *Morphological Voice and its Grammatical Interfaces*. Esku-orriak, University of Vienna.
- Salaburu, Pello & Maite Lakar. 2005. *Baztangomintzoa: gramatika eta hiztegia*. Nafarroako Gobernua & Euskaltzaindia.
- Salaburu, Pello, Patxi Goenaga & Ibon Sarasola. 2011-gaur arte. *Sareko Euskal Gramatika*. [Interneten eskuragarri: <https://www.ehu.eus/seg/>; azkenbisita: 2019-06-24].
- Sarasola, Ibon, Pello Salaburu & Josu Landa. 2001-2007. *Ereduzko Prosa Gaur* (EPG), UPV/EHU, Euskara Institutua. [Interneteneskuragarri: <http://www.ehu.eus/euskara-orria/euskara/ereduzkoa/>].
- Siewierska, Anna. 2013. Passive constructions. In Matthew Dryer & Martin Haspelmath (ed.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. [Interneten eskuragarri: <https://wals.info/chapter/107>, azkenbisita 2019-06-24].
- Urrestarazu, Iñigo. 2019. *Nire barruko Setatsua: (euskal) boza eta erdiko egiturak*. Gradu Amaierako Lana. UPV/EHU.
- Zribi-Hertz, Anne. 1982. La construction ‘se-moyen’ du français et son statut dans le triangle moyen-passif-réfléchi. *Linguisticae Investigationes. International Journal of Linguistics and Language Resources* 6. Amsterdam/Philadelphia : John Benjamins. 345-401.

Coreference and antecedent frequency effects in Spanish: An eye-tracking study¹

NEREA EGUSQUIZA, KEPA ERDOZIA, ADAM ZAWISZEWSKI

Abstract

In this paper, we explore the nature of the antecedent representations involved during anaphoric pronoun resolution in Spanish. We assume a serial, anaphoric lexical access account derived from the speech production model of Jescheniak and Levelt (1994); Levelt *et al.* (1999) and use word-frequency effects in order to investigate whether antecedent representations vary across languages during coreference resolution by testing Spanish native speakers and comparing our results to those obtained previously in English and German. We report and discuss the results from an eye-tracking while reading experiment in light of the current hypotheses.

Background

Sentence comprehension requires immediate access to the mental lexicon, where conceptual, grammatical and lexical information is stored. We say that speakers understand a sentence, a text or a discourse when they have an accurate mental representation of it. Needless to say, the longer the text or the speech, the more likely will be to find expressions referring to previous entities in the discourse. Anaphoric pronouns belong to this class of referential expressions which depend on the lexical properties of their referent or *antecedent* for interpretation. Evidence from agreement studies in Spanish, Italian, French and German shows that anaphoric pronoun resolution involves retrieving grammatical gender information from its antecedent representation in memory (Garnham *et al.*, 1995; Di Domenico and De Vicenzi, 1995; Cacciari *et al.*, 1997; Irmen and Knoll, 1999; Arnold *et al.*, 2000; Rigalleau *et al.*, 2004; Hammer *et al.*, 2007). Nevertheless, as suggested by Meyer and Bock (1999), antecedent representations could, in principle, differ qualitatively across languages based on the kind of information needed to mediate coreference. It might be the case that in languages like English, for instance, discourse/conceptual information will suffice to interpret anaphoric pronouns (Hankamer and Sag, 1976; Sag and Hankamer, 1984; Cloitrew and Bever, 1988; Lucas *et al.*, 1990). Hence, it is unclear what kind of antecedent representation is retrieved during anaphoric pronoun resolution and whether it is language specific or not.

¹ We acknowledge funding by the Basque Government (IT1169-19 (2019-2021); BFI-2012-219), MINECO (FFI2015-64183-P) and MICINN (PGC2018-097970-B-I00).

Here, we adopt a serial, two-stage, activation-based, lexical access model derived from the speech production model of Jescheniak and Levelt (1994); Levelt *et al.* (1999), which suggests that lexical entries store information at two separate levels: grammatical and semantic information at the lemma level and orthographic and phonological information at the lexeme level. Evidence for a separate access to each informational layer during processing comes from a translation latency task where Jescheniak and Levelt (1994) found that homophones (words that only share the lexeme) inherited the lexical frequency of their control pairs –a word frequency effect which suggests that the locus of the effect lies at the lexeme level (see also Laubstein, 1999, Laubstein, 2002; for a dissenting view, see Bonin and Fayol, 2002; Finocchiaro and Caramazza, 2006; Navarrete *et al.*, 2006).

In order to investigate what kind of antecedent representation is retrieved during pronoun processing in Spanish, we use the *word-frequency effect*, which is one of the most robust phenomena in experimental psycholinguistics. It refers to the fact that frequently used words are processed faster and more accurately than infrequent ones. High-frequency words yield faster responses in reading (Rayner and Duffy, 1986), lexical decision (Schilling *et al.*, 1998) and picture naming tasks (Oldfield and Wingfield, 1965; Almeida *et al.*, 2007). Furthermore, data from brain damaged patients suggest that high-frequency words are preserved more often than low-frequency words (Dell, 1990; Colangelo *et al.*, 2004; Knobel *et al.*, 2008) and that speakers experience less tip-of-the-tongue states with high-frequency words (Brown, 2012; Gianico, 2010; Vitevitch and Sommers, 2003). Although all these observations suggest a processing advantage for high-frequency words compared to low-frequency words during lexical access, it remains an open question whether anaphoric pronoun resolution depends on the lexical frequency of the antecedent.

Hypotheses

In order to account for the results found in previous literature, three hypotheses have been proposed: the *full re-access account*, the *lemma re-access hypothesis* and the *saliency account*.

The *full reaccess account* states that pronoun comprehension is similar to lexical access for non-anaphoric words: readers will retrieve the same information from the antecedent as when they first accessed to it. Hence, it predicts that the lexical frequency of the antecedent should be transferred to the anaphoric pronoun. This effect has been observed in language production tasks. Navarrete *et al.* (2006), for instance, reported faster naming latencies for sentences containing a pronoun with a high-frequency antecedent noun than with a low-frequency antecedent noun. Similarly, participants tested by Finocchiaro and Caramazza (2006) elicited faster pronominal clitic production latencies when the replaced noun was highly frequent.

The *lemma-reaccess account*, however, assumes that pronoun processing differs from non-anaphoric lexical access in that there is only partial re-access to the antecedent representation in memory; more specifically, to its lemma, where grammatical and semantic information is stored (Simner and Smyth, 1998). Based on evidence from speech production showing that the word frequency effect lies at the lexeme (Jescheniak and Levelt, 1994), the lemma reaccess account predicts that pronouns referring to low-frequency nouns should not be harder or easier to process than those linked to high-frequency nouns (Simner and Smyth, 1999). In a series of experiments carried out by Simner and Smyth (1999), participants read English sentences containing a pronoun

whose antecedent had been previously presented in picture form. The lexical frequency of the depicted noun was either high or low. Simmer and Smyth found that reading times for sentences containing a pronoun were unaffected by the lexical frequency of the antecedent. Further evidence supporting this account comes from Lago (2014), where she manipulated the lexical frequency of English common and proper noun antecedents in two eye-tracking experiments and concluded that low-frequency antecedents had no additional cost compared to high-frequency antecedents during pronoun resolution.

The *saliency account* challenges any null antecedent frequency effect. Van Gompel and Majid (2004), for instance, ran an eye-tracking study in which participants read sentence pairs in English containing a pronoun referring to a high or low-frequency antecedent. Pronouns with low-frequency antecedents elicited shorter reading times at the post-anaphoric region in first-fixation, first-pass and total reading times measures. The authors explained their reverse word-frequency effect in terms of saliency (Pynte and Colonna, 2000): infrequent words tend to have rather unusual —salient— features and their processing is thus more costly compared to that of high-frequency words, as evidenced by the word-frequency effect in nouns. The investment of additional attentional resources on infrequent words increases their probability of being better encoded in long-term memory (Garnham *et al.*, 1996; Malmberg and Nelson, 2003; Foraker and McElree, 2007) and, consequently, low-frequency words are recognized more quickly than high-frequency words (O'Brien and Myers, 1985; Glanzer and Adams, 1990; Shiffrin and Steyvers, 1997; Garnham, 2001; Malmberg and Nelson, 2003; Nicol and Swinney, 2003). The *saliency account* predicts that word saliency determines faster processing when pronouns refer to low-frequency antecedent nouns. Further evidence supporting this account comes from the EEG study in Heine *et al.* (2006b), where pronouns referring to high-frequency antecedents elicited a larger P300 component than those referring to low-frequency antecedents in German, suggesting a higher processing cost for the former (see also Heine *et al.*, 2006a).

The three accounts described above make different predictions. The *full reaccess account* predicts the same effects in nouns as in pronouns, i.e. faster reaction times with pronouns referring to high-frequency words than with pronouns referring to low-frequency words. The *lemma-reaccess account* predicts no frequency effects, whereas the *saliency account* predicts a reverse frequency effect, with faster processing for pronouns referring to low-frequency words.

Exploratory research

Previously, in Egusquiza *et al.* (2016), we reported two self-paced reading tasks showing that pronoun resolution in Spanish was affected by the syntactic role of the antecedent noun, but unaffected by its lexical frequency, in support of the lemma-reaccess account. The reliability of our results was evidenced by the fact that the frequency manipulation was strong enough to replicate the word-frequency effect in the antecedent regions as well as in the control conditions. Nevertheless, the self-paced reading method might not have been sensitive enough to detect any antecedent frequency effects during pronoun processing. Therefore, in the present study, we opted for eye-tracking, a methodology with a higher temporal resolution and which makes reading more natural than self-paced.

The present study

The aim of the present study is to examine the nature of the antecedent representations involved during anaphoric pronoun resolution in Spanish by manipulating the word frequency of the antecedent noun in an eye-tracking experiment. Two factors were crossed: lexical frequency of the antecedent (high vs. low) and anaphor type (pronoun vs. repeated noun). *Repeated noun* refers to the noun phrase control condition where the antecedent was repeated in order to pre-empt a null frequency effect across conditions. Based on Rayner *et al.* (1995), frequency effects should persist at repeated noun phrases. We considered three hypotheses in our study: (i) faster reading times for pronouns referring to high-frequency antecedents (*full reaccess account*); (ii) no antecedent frequency effects for pronouns (*lemma-reaccess account*); (iii) faster processing for pronouns referring to low-frequency words (*saliency account*).

The eye-tracking experiment

The underlying initial assumption in eye-tracking studies is that eye movements are strongly linked to cognitive processing (Just and Carpenter, 1980; cf. Anderson *et al.*, 2004). Reading times are interpreted as an index of language processing effort; with longer reading times (and more fixations) revealing processing difficulty compared to faster reading times (and less fixations), much as in self-paced reading tasks. Eye-movement data, however, offer a more detailed map of what happens during online sentence comprehension. Regressions (i.e. re-reading earlier parts of a text), for instance, cannot be captured in self-paced reading tasks and are interpreted as a sign of language processing difficulty at the discourse/semantic integration levels. Despite its higher temporal resolution compared to self-paced reading, eye-tracking while reading shows higher data loss rates due to word skipping and blinking (Rayner *et al.*, 2011), among other factors. The method provides highly accurate information about where readers look, for how long they gaze and how many eye fixations they make in a region.

Method

Participants

A set of twenty-two native speakers of Spanish (twelve women), ages 18-24 ($SD = 1.96$), who had not participated in Egusquiza *et al.* (2016), were recruited from the University of the Basque Country (Vitoria-Gasteiz campus). All participants had normal or corrected-to-normal vision. Two of them were excluded from further analysis because their data accuracy was below 70%.

Materials and Design

The materials were adapted from those used in Experiment 2 in Egusquiza *et al.* (2016). Two factors were crossed: frequency of the antecedent noun (high vs. Low) and anaphor type (repeated noun vs. pronoun). 160 sentence pairs were distributed in four lists (10 items per condition per list) so that each participant read only one version of the same item. 80 filler sentences were randomly intermixed with the experimental sentence pairs of each list and 40 yes/

no comprehension questions (half of them about the experimental sentences) were included. Both filler and experimental sentences shared a similar structure and length.

Our materials differed from Experiment 2 (Egusquiza *et al.*, 2016) in two aspects: (i) a three-word spillover region was included after the post-anaphoric one to capture any wrap-up, sentence-final effect; (ii) an additional discourse entity was inserted in the prepositional phrase following the verb in order to improve the naturalness of the materials in the repeated noun conditions. It mismatched in gender with the first element in the prepositional phrase (the retrieval target.) The subject was feminine and singular across conditions, omitted in the second sentence of each pair and it always mismatched the retrieval target in gender. Table 1 contains a sample of our materials.

Table 1

Sample of the materials used in the eye-tracking experiment

EYE-TRACKING EXPERIMENT	
Repeated noun phrase	01. La senadora criticó [a un ministro] y a una parlamentaria durante el discurso de ayer. The senator_{fem} criticized [a minister_{masc}] and a member of parliament_{fem} during yesterday's speech. HF Posteriormente arremetió [contra el ministro] en los medios nacionales y europeos. Later, (she) attacked [the minister] on the national and European media. <i>Meaning: The senator criticized a minister and a member of parliament during yesterday's speech. Later, (she) attacked the minister on the national and European media.</i>
	02. La senadora criticó [a un banquero] y a una alcaldesa durante el discurso de ayer. The senator_{fem} criticized [a banker_{masc}] and a mayoress during yesterday's speech. LF Posteriormente arremetió [contra el banquero] en los medios nacionales y europeos. Later, (she) attacked [the banker] on the national and European media. <i>Meaning: The senator criticized a banker and a mayoress during yesterday's speech. Later, (she) attacked the banker on the national and European media.</i>
	03. La senadora criticó [a un ministro] y a una parlamentaria durante el discurso de ayer. The senator_{fem} criticized a minister_{masc} and a member of parliament_{fem} during yesterday's speech. HF Posteriormente arremetió [contra él] en los medios nacionales y europeos. Later, (she) attacked [him] on the national and European media. <i>Meaning: The senator criticized a minister and a member of parliament during yesterday's speech. Later, (she) attacked him on the national and European media.</i>
	04. La senadora criticó [a un banquero] y a una alcaldesa durante el discurso de ayer. The senator_{fem} criticized [a banker_{masc}] and a mayoress during yesterday's speech. LF Posteriormente arremetió [contra él] en los medios nacionales y europeos. Later, (she) attacked [him] on the national and European media. <i>Meaning: The senator criticized a banker and a mayoress during yesterday's speech. Later, (she) attacked him on the national and European media.</i>

The discourse entities were taken from Es-Pal Spanish Lexical database (Duchon *et al.*, 2013), and pairs in the conjoined prepositional phrase were matched for relative frequency within each item to keep the prepositional-phrase saliency level constant. The Es-Pal Spanish Lexical database showed that the frequency type of the high-frequency antecedents (mean: 93.06 per million words, range: 18.25-912.26) was much higher than that of the low-frequency antecedents (mean: 3.10 per million words, range: 0.08-5.83).

Norming study. We ran a norming study where sixteen native speakers of Spanish judged the acceptability of our materials on a 7-point Likert scale, where 7 meant perfectly acceptable and 1 was completely unacceptable. Each participant read ten experimental sentences intermixed with thirty fillers. None of them took part in the following eye-tracking experiment. Target sentences with a rate higher than 3.5 were considered for the follow-up experiment. None of them had to be excluded. The mean score value for the four lists was 5.58 (range: 5.15-6.09).

Procedure

The experiment involved an eye-tracking reading task. Participants were tested individually in a dimed-light, sound-proof booth, and eye-movements were recorded using a remote Tobii X120 Eye Tracker interfaced with a PC. The sampling rate for recordings was 120 Hz. The maximum gaze angle was 36 degrees and viewing was binocular. Participants were seated 67 cm from a 19-inch LCD screen. The forty experimental sentence pairs were randomly intermixed with the eighty fillers and displayed on a single line in 13.5 pt. fixed width Arial white font on a black screen. Each participant was randomly assigned one of the four lists of stimuli. The experiment was implemented using the *Tobii Studio 3.2* software. A calibration check was run at the beginning of each individual session. Previous to the task, participants were instructed to read carefully at their normal rate and completed a short practice session with 5 sentence pairs and their corresponding yes/no comprehension questions. A fixation cross on the left edge of the computer screen was displayed for 800 milliseconds before each item. The position of the fixation cross coincided with the beginning of every sentence to prevent participants from random looking and to make sure that the starting point of every sentence was the same for all conditions and items. Forty yes/no comprehension questions were included to ensure participants were attending to the stimuli. The entire experimental session lasted approximately 45 minutes and six optional breaks were included along the task. All participants had at least one break.

Analysis

Only subjects whose data accuracy was 70% or above were taken into consideration for further analysis. Two subjects did not meet this criterion and thus, were excluded. Skips of a region in any particular measure were treated as missing data points. Following Van Gompel and Majid (2004), three regions of analysis were delimited: the antecedent region, the anaphor region and the post-anaphor region. The antecedent region was the first prepositional phrase in the conjoined object position (e.g. *a un ministro* ‘to a minister’). The anaphor region was the prepositional phrase immediately after the verb in the second sentence (e.g. *contra el ministro/ contra él* ‘against the minister/against him’). The post-anaphor region consisted of the three-word prepositional phrase after the anaphor region. We report three eye-tracking measures: first-fixation

duration, fixation duration and total fixation duration. *First-fixation duration* is the duration of the reader's first fixation in a region. *Fixation duration* corresponds to the sum of all fixations on a critical region before the reader leaves it for the first time (only first-pass reading is taken into account; no regressions). *Total fixation duration* is the sum of all fixation durations in a region (possible regressions included). During the recordings, Tobii Studio 3.2 software's IV-T *Fixation Filter* (Olsen, 2012) was set on: adjacent fixations were automatically merged provided the fixation was shorter than 75 ms (see Salojärvi *et al.*, 2005) and the maximum angle between both fixations was 0.5 degrees - to filter out micro-saccades (Komogortsev *et al.*, 2010), which usually have an amplitude of 0.5 or less (Yarbus, 1967). Fixations shorter than 120 ms or longer than 890 ms were deleted (for similar approaches, see Drieghe *et al.*, 2010; Folk and Morris, 2003; Johnson *et al.*, 2007; Lowder *et al.*, 2013; Rayner *et al.*, 2010; Van Gompel and Majid, 2004). The remaining data points that exceeded a threshold of 2.5 standard deviations from the mean by region and condition were excluded from further analysis (Ratcliff, 1993). This process affected 1.6% of the data in first-fixation duration, 2.99% in fixation duration and 1.84% in total fixation duration. Two ANOVAs were conducted for each eye-tracking measure and region, one with participants (F1) and one with items (F2) as the random factors, with Frequency (High Frequency / Low Frequency) and Anaphor type (Noun Phrase / Pronoun) as within-participants and within-items factors. Only frequency effects and its interaction with anaphor type will be reported. In order to directly compare frequency effects, paired T-test analyses were carried out whenever appropriate.

Results

Antecedent region. Low-frequency antecedents took longer to read than high-frequency antecedents in all measures but total fixation duration (first fixation: $F(1, 19) = 8.11$, $p = .01$; $F(2, 39) = 7.01$; $p = .01$; fixation duration: $F(1, 19) = 6.03$; $p = .02$; $F(2, 39) = 3.99$; $p = .05$; total fixation duration: $F(1, 19) = .27$; $p = .61$; $F(2, 39) = .49$; $p = .48$).

	Antecedent	Anaphor	Post-anaphor
<i>First-Fixation times</i>			
Repeated NP-HF	237 (9)	269 (16)	248 (13)
Repeated NP-LF	275 (15)	303 (20)	255 (10)
Pronoun-HF	247 (13)	252 (14)	284 (17)
Pronoun-LF	279 (14)	242 (11)	280 (17)
<i>Fixation times</i>			
Repeated NP-HF	241 (9)	255 (14)	248 (13)
Repeated NP-LF	269 (14)	286 (17)	270 (10)
Pronoun-HF	245 (9)	254 (16)	268 (16)
Pronoun-LF	274 (13)	241 (11)	276 (12)
<i>Total Fixation times</i>			
Repeated NP-HF	507 (51)	449 (37)	589 (48)
Repeated NP-LF	559 (50)	513 (44)	620 (51)
Pronoun-HF	550 (56)	384 (34)	817 (65)
Pronoun-LF	529 (39)	334 (27)	725 (49)

Anaphor region. In first-fixation duration, no frequency effect ($F(1, 19) = .7; p = .41$; $F(2, 39) = .01; p = .91$) or interaction with anaphor type ($F(1, 19) = 2.53; p = .13$; $F(2, 39) = 3.89; p = .056$) was found. However, in fixation duration, there was an interaction between both factors ($F(1, 19) = 4.94; p = .04$; $F(2, 39) = 11.65; p = .002$). Paired T-tests showed that repeated low-frequency nouns took longer to read than repeated high-frequency nouns ($t(19) = -2.31; p = .03$; $t(39) = -2.89; p = .006$), but pronoun conditions did not elicit any significant difference ($t(19) = 1; p = .32$; $t(39) = 1.54; p = .13$). In total fixation duration, the interaction between frequency and type of anaphor turned out to be statistically significant ($F(1, 19) = 7.33; p = .01$; $F(2, 39) = 6.58; p = .01$). Pronouns referring to low-frequency antecedents were read faster than those referring to high-frequency ones in the analysis by subjects ($t(19) = 2.47; p = .02$; 334.49 ms. vs. 384.80 ms.); but not by items ($t(39) = 1.34; p = .18$; 343.44 ms. vs. 394.79 ms.). Regarding the control condition, the opposite pattern was observed: repeated high-frequency nouns were read faster than repeated low-frequency nouns in the analysis by items ($t(39) = -2.31; p = .02$; 449 ms vs. 522.48 ms.), but not by subjects (Repeated HF NP vs. Repeated LF NP: $t(19) = -1.67; p = .11$; 449.39 ms. vs. 513.27 ms.).

Post-anaphor region. No frequency effects or interaction with anaphor type occurred in this region in any measure (first-fixation duration: ($F(1, 19) = 0; p = .93$; $F(2, 39) = .09; p = .76$); fixation duration: ($F(1, 19) = .31; p = .58$; $F(2, 39) = .77; p = .38$) and total fixation duration: ($F(1, 19) = 2.24; p = .15$; $F(2, 39) = 2.94; p = .09$).

Discussion

The aim of this eye-tracking-while-reading study was to determine how the lexical frequency of the antecedent affects anaphoric pronoun resolution. Similarly to the eye-tracking results obtained by Van Gompel and Majid (2004) and Lago (2014), we found a lexical frequency effect at the antecedent region in first-fixation and fixation duration; however, we do not observe any in total fixation duration. This difference suggests that high and low-frequency antecedents in our materials elicited a similar amount of regressions.

In the anaphor region, there was an interaction between frequency and anaphor type in fixation duration and total fixation duration, such as repeated noun phrases referring to high-frequency antecedents were read faster than those linked to low-frequency nouns. Nevertheless, this frequency effect was fully significant in fixation duration only.

Results elicited in total fixation duration should be interpreted with caution since they are partially significant: pronouns linked to low-frequency antecedents were read faster than the same pronouns referring to high-frequency nouns, but this saliency effect was only significant in the analysis by subjects². The fact that there was no effect in fixation duration for pronouns compared to repeated noun phrases suggests that the saliency effect in total fixation duration was caused by readers looking back to the anaphor region. Similarly to Van Gompel and Majid (2004) and Lago (2014), the pronoun was a short word at the edge of the region, so we expected participants' eyes to move onto the next without having fully processed it. However, there were no spill-over effects of antecedent lexical frequency at the post-anaphor region (cf. Van Gompel and Majid, 2004; Lago, 2014), which led us to think that the saliency effect detected in the previous region could be a Type I error.

² Despite the fact that we ran a norming study to evaluate the grammatical acceptability of our materials, a plausibility effect might be responsible for the lack of significance in the analysis by items. It is likely that high-frequency words fitted better in the context than low-frequency words.

Conclusion

The aim of this study was to investigate the nature of the antecedent representations involved during anaphoric pronoun resolution in Spanish after Meyer and Bock (1999) suggested that pronoun processing could, in principle, differ across languages based on the kind of information needed to mediate coreference. Whereas we replicated the word-frequency effect in the case of repeated nouns in the control condition, no reliable effect was found for pronouns: infrequent antecedents did not cause any processing cost during pronoun comprehension. We interpret this finding as evidence in favor of the lemma-reaccess account by Simner and Smyth (1998, 1999), which postulates that anaphoric pronoun resolution involves lemma retrieval (i.e. retrieval of the syntactic and semantic properties of the antecedent) whereas lexical (re)access to nouns additionally requires orthographic or phonological information. Simner and Smyth (1999) based their proposal on the serial, two-stage, activation-based speech production model proposed by Jescheniak and Levelt (1994), which locates the word-frequency effect at the lexeme level. The fact that this effect was elicited at the antecedent region as well as with repeated nouns in the control condition suggests thus that the lexeme must have been retrieved during sentence comprehension. Therefore, we argue that the absence of antecedent frequency effects during anaphoric pronoun resolution is a reliable null effect indeed rather than evidence for a difference between the production and comprehension systems. The current research extends Simner and Smyth and Lago's conclusion in English to Spanish. Likewise, it contradicts the results reported in English and German by Van Gompel and Majid (2004) and Heine *et al.* (2006a,b), respectively, in support of the saliency account: i.e. the idea that since infrequent words require more attention and effort during language processing, they are better encoded in memory and, therefore, reaccessed faster than frequent antecedents during anaphoric pronoun resolution. Van Gompel and Majid (2004) suggests that their results can be accommodated within the lemma re-access account by positing two processing stages: one where the memory access mechanism finds the antecedent and retrieves the infrequent word faster, and a second stage where the processor re-accesses the lemma information of the antecedent only. The question which follows then is if discourse information would be enough to target and retrieve the antecedent during stage one of pronoun resolution when the grammatical properties of the antecedent are not yet available and there is not enough context.

Given our results, further research on pronoun processing and lexical access in general needs to be carried out in order to provide data from a larger pool of participants and typologically different languages before drawing any strong conclusions on the nature of the antecedent representations retrieved across languages.

Summary

This article provides empirical evidence in favor of a lexical access model where nouns and pronouns retrieve qualitatively different antecedent representations from memory, in support of a theory of anaphoric lexical access in comprehension (Simner and Smyth, 1998, 1999). We show that coreference resolution in Spanish only involves lemma retrieval —i.e. retrieval of the syntactic and semantic properties of the antecedent—, whereas lexical access to nouns and repeated nouns additionally targets orthographic-phonological information.

References

- Almeida, J., Knobel, M., Finkbeiner, M., and Caramazza, A. (2007). The locus of the frequency effect in picture naming: When recognizing is not enough. *Psychonomic Bulletin & Review*, 14(6):1177-1182.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4):1036-1060.
- Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., and Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eye-tracking. *Cognition*, 76(1):B13-B26.
- Bonin, P. and Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, 14(3):289-313.
- Brown, A. S. (2012). *The tip of the tongue state*. Taylor & Francis.
- Cacciari, C., Carreiras, M., and Cionini, C. B. (1997). When words have two genders: Anaphor resolution for Italian functionally ambiguous words. *Journal of Memory and Language*, 37(4):517-532.
- Cloitre, M. and Bever, T. G. (1988). Linguistic anaphors, levels of representation, and discourse. *Language and Cognitive Processes*, 3(4):293-322.
- Colangelo, A., Holden, J. G., Buchanan, L., and Van Orden, G. C. (2004). Speculation about behavior, brain damage, and self-organization: The other way to herd a cat. *Brain and Language*, 90(1):151-159.
- Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5(4):313-349.
- Di Domenico, E. and De Vicenzi, M. (1995). Gender and number in the retrieval of pronoun antecedents: differences in use and representation. *Actes du deuxième colloque Langues et grammaire*, Paris 8 juin 1995, pages 95-109.
- Drieghe, D., Pollatsek, A., Juhasz, B. J., and Rayner, K. (2010). Parafoveal processing during reading is reduced across a morphological boundary. *Cognition*, 116(1):136-142.
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., and Carreiras, M. (2013). Espal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, 45(4):1246-1258.
- Egusquiza, N., Navarrete, E., and Zawiszeński, A. (2016). Antecedent frequency effects on anaphoric pronoun resolution: Evidence from Spanish. *Journal of Psycholinguistic Research*, 45(1):71-84.
- Finocchiaro, C. and Caramazza, A. (2006). The production of pronominal clitics: Implications for theories of lexical access. *Language and Cognitive Processes*, 21(1-3):141-180.
- Folk, J. R. and Morris, R. K. (2003). Effects of syntactic category assignment on lexical ambiguity resolution in reading: An eye movement analysis. *Memory & Cognition*, 31(1):87-99.
- Foraker, S. and McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language*, 56(3):357-383.
- Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Psychology Press.
- Garnham, A., Oakhill, J., Ehrlich, M.-F., and Carreiras, M. (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. *Journal of Memory and Language*, 34(1):41.
- Garnham, A., Traxler, M., Oakhill, J., and Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language*, 35(4):517-543.
- Gianico, J. L. (2010). *Word concreteness and word frequency as moderators of the tip-of-the-tongue effect*. State University of New York at Albany.
- Glanzer, M. and Adams, J. K. (1990). The mirror effect in recognition memory: data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(1):5.

- Hammer, A., Goebel, R., Schwarzbach, J., Münte, T. F., and Jansma, B. M. (2007). When sex meets syntactic gender on a neural basis during pronoun processing. *Brain Research*, 1146:185-198.
- Hankamer, J. and Sag, I. (1976). Deep and surface anaphora. *Linguistic inquiry*, 7(3):391-428.
- Heine, A., Tamm, S., Hofmann, M., Bösel, R. M., and Jacobs, A. M. (2006a). Event-related theta activity reflects memory processes in pronoun resolution. *Neuroreport*, 17(18):1835-1839.
- Heine, A., Tamm, S., Hofmann, M., Hutzler, F., and Jacobs, A. M. (2006b). Does the frequency of the antecedent noun affect the resolution of pronominal anaphors?: An ERP study. *Neuroscience Letters*, 400(1):7-12.
- Irmen, L. and Knoll, J. (1999). On the use of the grammatical gender of anaphoric pronouns in German. A comparison between Finns and Germans. *Sprache & Kognition*, 18(3-4):123-135.
- Jescheniak, J. D. and Levelt, W. J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4):824.
- Johnson, R. L., Perea, M., and Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance*, 33(1):209.
- Just, M. A. and Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4):329.
- Knobel, M., Finkbeiner, M., and Caramazza, A. (2008). The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production. *Cognitive Neuropsychology*, 25(2):256-286.
- Komogortsev, O. V., Jayarathna, S., Koh, D. H., and Gowda, S. M. (2010). Qualitative and quantitative scoring and evaluation of the eye movement classification algorithms. In *Proceedings of the 2010 Symposium on eye-tracking research & applications*, pages 65-68. ACM.
- Lago, M. S. (2014). *Memory and Prediction in Cross-Linguistic Sentence Comprehension*. PhD thesis, University of Maryland-College Park.
- Laubstein, A. S. (1999). Lemmas and lexemes: The evidence from blends. *Brain and Language*, 68(1):135-143.
- Laubstein, A. S. (2002). Experimental evidence for serial models of lexical access: A judgment task. *Brain and Language*, 81(1-3):424-431.
- Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1):1-38.
- Lowder, M. W., Choi, W., and Gordon, P. C. (2013). Word recognition during reading: The interaction between lexical repetition and frequency. *Memory & Cognition*, 41(5):738-751.
- Lucas, M. M., Tanenhaus, M. K., and Carlson, G. N. (1990). Levels of representation in the interpretation of anaphoric reference and instrument inference. *Memory & Cognition*, 18(6):611-631.
- Malmberg, K. J. and Nelson, T. O. (2003). The word frequency effect for recognition memory and the elevated-attention hypothesis. *Memory & Cognition*, 31(1):35-43.
- Meyer, A. S. and Bock, K. (1999). Representations and processes in the production of pronouns: Some perspectives from Dutch. *Journal of Memory and Language*, 41(2):281-301.
- Navarrete, E., Basagni, B., Alario, F.-X., and Costa, A. (2006). Does word frequency affect lexical selection in speech production? *The Quarterly Journal of Experimental Psychology*, 59(10):1681-1690.
- Nicol, J. L. and Swinney, D. (2003). The psycholinguistics of anaphora. *Anaphora: A reference guide*, pages 72-104.
- O'Brien, E. J. and Myers, J. L. (1985). When comprehension difficulty improves memory for text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(1):12.

- Oldfield, R. C. and Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17(4):273-281.
- Olsen, A. (2012). The tobii i-vt fixation filter. *Tobii Technology*.
- Pynte, J. and Colonna, S. (2000). Decoupling syntactic parsing from visual inspection: The case of relative clause attachment in french. *Reading as a Perceptual Process*, pages 529-547.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114(3):510.
- Rayner, K., Castelhana, M. S., and Yang, J. (2010). Preview benefit during eye fixations in reading for older and younger readers. *Psychology and Aging*, 25(3):714.
- Rayner, K. and Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14(3):191-201.
- Rayner, K., Raney, G. E., and Pollatsek, A. (1995). *Eye movements and discourse processing*. Lawrence Erlbaum Associates, Inc.
- Rayner, K., Slattery, T. J., Drieghe, D., and Liversedge, S. P. (2011). Eye movements and word skipping during reading: effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, 37(2):514.
- Rigalleau, F., Caplan, D., and Baudiffier, V. (2004). New arguments in favour of an automatic gender pronominal process. *The Quarterly Journal of Experimental Psychology Section A*, 57(5):893-933.
- Sag, I. A. and Hankamer, J. (1984). Toward a theory of anaphoric processing. *Linguistics and Philosophy*, 7(3):325-345.
- Salojärvi, J., Puolamäki, K., Simola, J., Kovanen, L., Kojo, I., and Kaski, S. (2005). Inferring relevance from eye movements: Feature extraction. In *Proceedings of the NIPS 2005 Workshop on Machine Learning for Implicit Feedback and User Modeling*. Publications in Computer and Information Science.
- Schilling, H. E., Rayner, K., and Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, 26(6):1270-1281.
- Shiffrin, R. M. and Steyvers, M. (1997). A model for recognition memory: Rem—retrieving effectively from memory. *Psychonomic Bulletin & Review*, 4(2):145-166.
- Simner, J. and Smyth, R. (1998). Anaphoric vs. stimulus-based lexical access. In *CUNY Conference on Human Sentence Processing*, Rutgers University.
- Simner, J. and Smyth, R. (1999). Phonological activation in anaphoric lexical access (ala). *Brain and Language*, 68(1):40-45.
- Van Gompel, R. P. and Majid, A. (2004). Antecedent frequency effects during the processing of pronouns. *Cognition*, 90(3):255-264.
- Vitevitch, M. S. and Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & Cognition*, 31(4):491-504.
- Yarbus, A. L. (1967). Eye movements during perception of complex objects. In *Eye movements and vision*, pages 171-211. Springer.

Eder baliz, on ez eiliz Euskaldunok hizkuntzarekin ditugun maitasun/gorroto harremanez

PATXI GOENAGA

1. Errefrau zaharra aitzakia

«Fuese hermosa y no buena» modura itzultzen du Garibaik izenburuan jarri dudana esaldia. Ezkontzan bertuteak baino areago edertasuna bilatzen dutenei egiten die erreferentzia, omen. Errefrau zahar honekin hasi nahi nuke nire jardun hau. Ongi datorkigukeelakoan. Baina ez gara euskara zaharraz ariko, gaurko kontuez baizik, euskaldunok geure hizkuntzarekin dugun maitasun/gorroto harremanaz nolabait. Gogora ditzagun gure poetaren hitzak:

Ezein ere lengoajerik
ez frantsesa ez berzerik
orai ez da erideiten
euskararen parerik.
Heuskara,
jalgi hadi dantzara!

Harro agertzen gara maiz euskaldunok geure hizkuntzaz, munduko hizkuntza ederrena dela iritzita. Harro, halaber, gure hizkuntza gaixoa Interneteko zeruan ez dakit zenbatgarren ‘gradora’ igoa aurkitzen delako eta abar. Batzuetan, ordea, esaten direnek ez digute asko laguntzen. Dantzarako gogo gutxi ematen digute.

2. Euskara konplexuegia delako ustea

Euskaldun askok ez diote ontasun handirik aurkitzen hizkuntzari. Hori bai, munduko hizkuntzarik ederrena dela aldarrikatuko dute, euskararen edertasunaz liluratuta edo. Ez dira fido, nonbait, hizkuntzak duen ontasunaz: hizkuntzaren ona eta ederra ezin uztarturik dabilta. Edertasuna zer datekeen definitzen ez guke berehalakoan asmatuko. Baina euskara ere, beste hizkuntza guztiak bezalaxe ‘artefakto’ ederra da eta denak liluratuta gauzka haren ederrak. Baina ematen du bost axola zaiela haren bertutea eta ontasuna. Ez dute aski on ikusten gaurko bizimodurako, hizkuntza lantzea, erabiltzea eta nork bere burua trebatzea, oraindik ere askorentzat gehiegi da. Kontenplatzeko zerbait bai, baina eguneroko bizitzan erabiltzeko egin behar den ahalegina gehiegizkoa nonbait eta askorentzat sinbolo bat besterik ez denez, hor gelditzen da: sinbolo bat. «Ongi

dago, haurrei erakutsi egin behar zaie, behar den dirua gastatu behar da eta abar...». Baina hori dena besteri dagokio.

Jarrera eroso horri aurre egin beharrean, batzuek aurkitu dute aitzakia polita, euskara zaila dela, konplikatuagia, gaztelania edo ingelesa dagoen lekuan, euskara bezalako hizkuntza zahar bat (Europako zaharrena omen) maite bai, baina gure bizitzaren parte eta partaide izatea... hori askotxo da. Mito batzuk eraiki ditugu eta koartada ederra eskaintzen diegu euskarari gure eguneroko bizitza arruntean eta profesionalean ere zirrikiturik ez eskaintzeko. Interneten ditugun webgune, wikipedia eta enparauetan sartzea aski da, serio-serio zabaltzen diren hainbat sinplekeria aurkitzeko. Oharkabean ontzat ematen ditugu, gainera.

3. Determinatzaile Sintagmen barneko egituraz

Behin honez gero, Larramendi handiarekin gogoratzea ezinbestekoa gertatzen da. *El imposible vencido* ospetsuan honela aurre egiten die arerioei:

«Otra objecion hacen contra el bascuete; porque en esta lengua todo se ha de construir al revés, por la posposición que tiene de artículos, v.gr. *ur garbi-á, burú uts-á* se han de construir *agua clara la, cabeza vana la*. Qué simpleza! Y de dónde te consta que esa construccin es al revés? Y mas si doy yo en decir que va al revés tu construccion castellana? Sabes? Los hebreos escriben el renglon tirándole de la mano derecha á la izquierda, y tú le tiras de la izquierda á la derecha: y quién escribe al revés? Según ese reparo se habrán de desterrar del latin el *que, ve, ne* y otros adverbios y aun toda su sintáxis. Y sino construye esta oracion: *semper enim hic homo leones inter, ursosque commoratus est*: mira que donosura, siempre por que este hombre leones entre osos y vivido ha. Pero yo hago mal en cansarme en estas puerilidades. *Goazén emendic*.» (Manuel de Larramendi, *El imposible vencido*, 1853, 7 orr.)

Hizkuntza baten konplexutasuna nola neur daitekeen jakitea ez da erraza. Gertakarien berri ematerakoan, azalpenak berak ere izan daitezke konplexuagoak azaldu nahi dena baino. Uste izatekoa da, hala ere, hizkuntza guztiak, giza pentsamentuaren eta beraz, burmuinaren fruitu diren aldetik, konplexutasun dezente izan dezaketela.

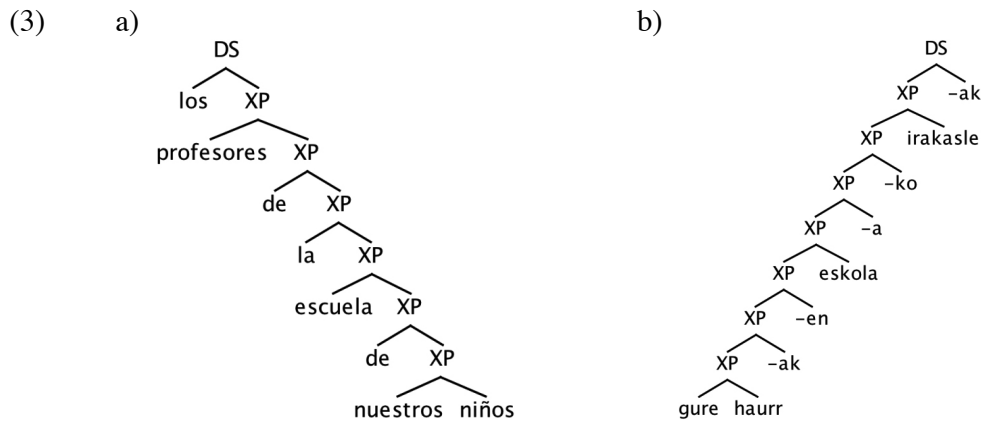
Determinatzaile sintagmaren egitura bera ere konplexua da, baina gure inguruko hizkuntzetan behintzat DS modu bertsuan osatua dagoela esango nuke. Egitura horren osagaiak berdintsuak izango dira, hitzak erabat diferenteak izanik ere. Osagai horien kokagunea, ordena, berezia izan liteke. Gure inguruan ditugun hizkuntza erromanikoak, gaztelania eta frantsesa adibidez, hartzen baditugu, elkarren antzekoak dira. Horiekin alderatuta, euskara 'gutziz desberdina' da jendearen ustetan. Batzuen arabera, gainera, indar komunikatibo desberdina dute era bateko hurrenkerak eta beste batekoak. Aldera ditzagun honako hitz multzoak:

- (1) a. gure haurren eskolako irakasleak
- b. los profesores de la escuela de nuestros niños...
- c. les professeurs de l'école de nos enfants...

Hitz segida hauek alderatuz gero, hizkuntza horiek ezagutzen dituenak erraz antzemango dio, frantsesa eta gaztelania bide batetik doazela eta euskara bestetik. Sintagmaren konplexutasunari dagokionez, ordea, elkarren paretsu joko ditu, seguru asko:

(2)	gaztelania	(1) <i>lo-s</i>	(2) <i>profesor-es</i>	(3) <i>de</i>	(4) <i>la</i>	(5) <i>escuela</i>	(6) <i>de</i>	(7) <i>nuestro-s</i>	(8) <i>niño-s</i>
	frantsesa	(1) <i>le-s</i>	(2) <i>professeur-s</i>	(3) <i>de</i>	(4) <i>l'</i>	(5) <i>école</i>	(6) <i>de</i>	(7) <i>no-s</i>	(8) <i>enfant-s</i>
	euskara	(7) <i>gu-re</i>	(8) <i>haur-</i>	(8') <i>(-ak)</i>	(3) <i>-en</i>	(5) <i>eskola</i>	(6) <i>-ko</i>	(2) <i>irakasle</i>	(1) <i>-a-k</i>

Arbola bidez irudikatzen badugu, are argiago oraindik:¹



Osagaien hurrenkera guztiz alderantzizkoa izan arren, elkarren ispilu-irudi bat funtsean, (3b)-koegitura (3a)-koa baino konplexuagoa denik ezin esan. Jakina, hizkuntza ikasi nahi duenak desberdintasun honen berri jasoko du eta horretaz ohartuko da. Agian sufritu ere egin beharko du pixka bat. Hala ere, garbi dago bi egituren konplexutasuna antzekoa dela. Hala behar du. Espainolez mintzo denari, hala ere, frantsesa hurbilago egingo zaio euskara baino, hizkuntzen arteko ‘alde’ desberdina baita.

Hizkuntzei kanpotik begiratzuz gero, alde handiak daude, bistan da, baina barnea arakutzen denean, gauzak bestela ikusten dira.

Gaur gure artean badira euskara ‘egokitzen’ saiatzen direnak. Adibidez, komunikazioa errazte aldera, aditzak azkenerako utzi gabe aurreratzea proposatzen du zenbaitek. Postposizioen kokagunea ere aldatzea proposatzen dute beste batzuek, postposizioak preposizio bihurtuz. Ez diet etorkizun handirik ikusten euskara ‘egokitu’ horren apostoluei. Jolas bezala ere ariketa horiek ez dute asko balio, nire ustez. Salabururen lanean aski argi gelditzen da kontua, Jesus Rubiori ematen dion errepasoan (Salaburu 2002, 183-187 orr.).

Perpaus erlatiboak ere arazotsu gertatzen dira batzuetan. Erlatibo murrizgarri deitzen direnak batez ere. Maiz, ordea, arazo faltsuak besterik ez dira. Amurizak (2010: 36) esate baterako ‘soluzio gabeko kasuak’ biltzen dituen atal batean holako zenbait erlatibo biltzen ditu. Adibidez, *Una persona con la que has vivido tantos años no puede ser ajena a tu vida* adierazteko modu egokia

¹ Ohar gaitzen, bide batez, sintagma hauen osagaietan zenbat aldiz markatzen den plurala (-s) artikularen, izenaren eta adjektiboen artean numero komunztadura markatuz. Euskaraz hainbeste bider hori markatu beharrik ez da.

honako hau izan omen liteke: *Hainbeste urtetan bizi izan zaren pertsonarekin ezin zara haserretu orain*. Nire ustez, balio izan dezake, Amurizak gustuko ez badu ere: «soluzio hori neketsu eta er-negugarri gertatu ohi da azkenerako» dio. Izenordain erresuntiboak erabiltzea ere izan liteke soluzioa. Nik neuk, hala ere, ez nituzke baztertuko gainerako bideak ere. Esate baterako, *Hay personas de quienes no consigues saber más de lo que preguntas* perpausari erraz eman dakioke buelta, er-negatu gabe ere. Adibidez: *pertsona batzuei buruz ezin da jakin galdetzen duzuna baino gehiago*, sintagmari buelta emanda.

Teorian erlatiboak adjektiboak bezalaxe, *ad infinitum* txerta liezke determinatzaile sintagmetan, baina edozeinek daki, adibidez, kopuru batetik gora, informazioa prozesatzea ezina gertatzen dela eta sintagma guztiek mugak izaten dituztela. Horretarako, saihezbideak beharrezko izaten dira askotan. Adibidez, euskaraz zilegi dira *nork idatzitako liburuak erre dituzte?* modukoak. Gaztelaniaz zuzen-zuzen ezin eman perpaus horiek, itzulunguru baten bidez ez bada. Adibidez, honela: *quién escribió los libros que han quemado*, atzekoz aurrera alegia. Eta zein da konplexuago, euskara ala erdara? Hizkuntzek baliabide desberdinak erabiltzen dituzte. Azken kasu honetan euskaraz bi modutara esan daiteke: a) *nork idatzitako liburuak erre dituzte?* / b) *nork idatzi zituen erre dituzten liburuak?* Gaztelaniaz modu bakarra dago, bigarrena.

4. Aditzaren konplexutasuna

Euskal aditzak ere konplexutasun pixka bat badu, ezin uka. Baina ez nuke esango beste hizkuntza batzuetakoa baino askoz konplexuagoa denik. Are gutxiago onartuko nuke euskal aditzaren ustezko konplexutasuna mamu gisa erabiltzea ikasleak izutzeko. Izan ere tresna konplexuak erabili beharrak, maiz, atzera botatzen ditu tresna horren balizko erabiltzaileak. Honek esan nahi du artefakto konplexuak erabiltzen irakasteko orduan pedagogia pixka bat behar dela. Ematen den informazioa neurtu beharra dago. Zer esan honako baieztapen hauetaz? Honela aurkeztzen du Wikipediak euskal aditza:

«El **verbo** es una de las **partes más complejas** de la **gramática del euskera**. A menudo se presenta como **uno de los retos más difíciles** para quienes aprenden la lengua, y muchos gramáticos vascos dedican la mayoría de sus páginas a listas o tablas de los paradigmas verbales.»

Beste honako hau ere aurkitu dut Interneteko web orriren batean:

«Dentro de la gramática euskera, **el verbo es una de las partes más complejas**. De hecho, quienes **se aventuran a aprender la lengua vasca**, encuentran en la utilización de sus verbos un verdadero **quebradero de cabeza**, y quienes se dedican **profesionalmente** a la investigación de este idioma, suelen centrar sus **estudios gramaticales y lingüísticos en el análisis de los paradigmas verbales**.»

Bi testu hauetan ‘moztu-itsatsi’ deitu ohi dugun eragiketa agerikoa gertatzen bada ere, esaldi bakoitzak eskatuko luke berariazko iruzkin bat. Baina iruzkin handirik gabe ere ongi ulertzen dira: euskal aditza hain da konplexua, non euskara ikastea abentura bat eta buruhauste itzela baita... Ez nago ados, jakina. Ideia horrek alferkeriara bultzatu nahi gintuzke, abenturetatik alde egitera, porrotera hitz batean esateko. Baieztapen biribil eta gehiegizko horiek egiteko ausardia dexente behar dela irizten diot. Zilegi bazait *pro domo* hitz egitea, euskara geure etxea den heinean, konplexutasuna ez da euskararen edo euskal gramatikaren akats bat, hizkuntza berez mekanismo kon-

plexua baita eta gramatikei buruz gehienez ere esan dezakeguna da hizkuntzak eta gramatikak, denak, mota bereko tresnak direla, funtzio bera duten tresnak. Baina denok ‘geurea’ den hizkuntzari ikusten diogu konplexutasun gutxien, geurea baita. Gehiegi kosta gabe ‘geuretu’ dugu gure hizkuntza. Arrazoi beragatik, hizkuntzarik ederrena norberarena izan ohi da. Hortaz, bi sentimentu kontrajarri izango ditugu. Nolanahi ere, geureaz gainera beste hizkuntzak ikasi nahi izan ditugunean azaleratzen dira hizkuntza ‘arrotzaren’ konplexutasunak eta ‘aldatz-gorak’. Euskararen kasuan, geurea izan arren eta txikitan ikasia, eskola faltaz, geure etxeko hizkuntza ere zail egiten zaigu maiz, eta txikitan ikasi dugun geure hizkuntzari ere erraz itsasten diogu hizkuntza konplikatuaren ‘sanbenitoa’: «euskara munduko hizkuntzarik ederrena da baina izugarri zaila».

Hemen ere, izenburuan ezarri dudana atsotitz horrek dioena aldarrikatzen dute euskararen zenbait maitalek: *eder baliz, on ez eiliz*. Baina euskaldunok beste modu batera eraiki behar genuke esaera: euskara bizikide edo ezkontide nahi badugu, izan bedi ederra (hizkuntza guztiak diren bezala) eta aldi berean, lanerako balio duena, lanean zaildua eta lanerako amorratzen dagoena. Horregatik, zenbaiten baieztapenak axolagabekeria izugarriak iruditzen zaizkit. Euskara irakasten saiatzen direnei konplexutasun horiek behar bezala neurtu eta gainditzen lagundu behar genieke euskararen normalkuntzarekin zerikusia izan genezakeen guztiok. Egia esan behar da eta ez sartu inori neurritz kanpoko beldurrik. Konplexutasunean indar eginez, azkenean konplexua geuk izango dugu geure lepoan.

Saia gaitezen ulertzen zertan datzan euskal aditzaren konplexutasun hori. Horretan hasi orduko, Patxi Altuna irakasle zorrotz eta finaz gogoratu nahi nuke. Hark eman baitzidan, ikasle nintzen garai haietan, euskara irakasle izateko aukera eta eginkizun horretarako gogoia ere. *Euskara hire laguna* izeneko liburuxkari zenbat buelta eman ote genizkion garai haietan ikasle-irakasleok. Kontua da neure bizitzan inoiz jaso dudana leziorik laburrena eta probetxugarriena berak eman zidala, aparte hartuta, ‘eskola partikular’ gisa edo. Euskal aditzaren paradigmek eta hauen erabilerak sortzen zidaten ezinegona agertu nion noizbait. Neu ere ilunpetan bezala bainenbilen ordu arte. Baina ordu laurden nahikoa izan zuen aditzaren ‘misterioak’ eta korapilo nagusiak argitzeko. Pare bat urte geroago edo, bere liburutxo didaktiko eta argian (Altuna, P. (1971)) azalduko zuen doktrina bera zen funtsean: ‘jokabide soila’ deitzen duenetik, indikatibotik alegia, abiatu eta «hau ikasiz gero, besteak arras erraz dira» liburuaren hasiera-hasieran dioen bezala.

Aditz guztiek eskema bera jarraitzen dute funtsean. Eredutzat DA/ DU adizkiak har ditzakegu, sinpleenak, konplikazio gutxien dutenak:

- (4) a. benetan ikusgarria [DA]; ez [DUZU] itxura txarrik...
b. berandu [etorri DA]; non [izango DA] bilkura?

Arestian aipatutako Altunaren irakats-saio hartan mapa batean zehar bidaltzen ninduela iruditzen zitzaidan. Lehenbizi eskatzen zidan ezaguna izan nezakeen adizki ‘erraz’ bat: *erosi duzu/du*, adibidez. Hau ondo zehaztu eta identifikatu ondoren, *duzu* laguntzailearekin erlazio morfologiko argia zuen beste adizki bat eskatzen zidan, konplexuagoa baina ‘arrunta’: ‘*zuk hura erosi d-u-zu*’-ren ordezt, ‘*guk hura erosi d-u-gu*’. Eta honen ordezt, iraganaldiko forma, bai batean bai bestean: *guk hura erosi genuen, nik hura erosi nuen*... eta abar. Gero hauetatik *diru aski ba-gen-u, diru aski ba-n-u*... *erosiko gen-u-ke, erosiko n-u-ke*.

Eta hauetatik, ‘ahal izatea’ adierazi nahi izanez gero?: *eros d-eza-ke -gu* (cf. *erosi d-u-gu*). Eta hemendik *eros d-eza-ke-t* (cf. *erosi d-u-t*), eta hemendik zergatik ez, *eros nezake* (cf. *erosiko nuke*)? Eman *diozu* adizkia ezagutuz gero, egokitzapen xume batzuk egin, eta prest oraingoan ere: eman

diezaiozun (*d-i-eza-io-gu-n*). Aginterazkoak txiki-txikitatik ikasten ditugu: *utzi zazu*, *emaiozu*, *al-txa zaitez* modukoak noiz ikasten ditugu? Bada, hauek jakinda zer kostatzen da *utz(i) ezazu*, *eman zaiozu* > *eman iezaiozu* eta forma berriak ikastea? Eta hemendik, antzara-jokoan bezala, adizki desberdinak osatuz joan gitezke.

Dena dela, ‘bisitatu’ dugun adizki bakoitzari bere ‘lekua’ bilatu behar zaio mapan. Baina mapak egin behar du izan, osatu beharra dago: batetik, *EDIN/ *EZAN laguntzaileak gehitu behar dizkiogu eta bestetik -(TE)KE morfemari ere leku egin beharko diogu. Hortaz, goiko adibideei-beste batzuk gehitzen ahal dizkiegu:

- (5) a. ondo [LEGOKE]; ez [DAKIGU]ezer...
b. bihar [joanLITEKE]; zerbait [eska DEZAKET]

(4)ko adibideei eta (5)ekoei leku egin behar genieke adiztegiko ‘mapan’.

(6)	orainaldia	lehenaldia	balizkoa
	DA/DU -DADI/-DEZA	ZEN/ZUEN ZEDIN/ZEZAN	Ba-LITZ/ ba-LU -LEDI/-LEZA
	DATEKE/DUKE DAITEKE/DEZAKE	ZATEKEEN/ZUKEEN ZITEKEEN/ZEZAKEEN	LITZATEKE/LUKE LITEKE/LEZAKE

Eskema honek, bere soiltasunean badu konplexutasun pixka bat ere, baina gaztelaniazko ordainek dutena baino handiagoa? Ez dut uste. Irregularitasunak ere bi hizkuntzetan gertatzen dira. Bestalde, gaztelaniak teoriarik beharko zituzkeen SER/HABER laguntzaileetatik bat bazter utzi zuen: *hemos llegado/hemos dicho*. Frantsesez ez: *nous sommes arrivés/nous avons fait*. Gaztelaniaz (frantsesez ez bezala, hain zuzen), aditza trantsitiboa izan nahiz izan ez, laguntzailea bakarra da: HABER alegia. Konplexutasunaren aldetik, gaztelania sinpleagoa litzateke, bi laguntzaile erabili beharrean bakarra baitu: *he venido* (intr.)/*he comido* (tr.) vs *etorri naiz* (intr.); *jan/bazkaldu dut* (tr.)... Puntu honetan euskal aditza konplexuagoa da, beharbada, laguntzailea aditz nagusiaren eta honen argumentu egituraren arabera hautatu behar baita. Gaztelaniaz horren premiarik ez da, baina, nahasiagoa ere gerta daiteke, hizkuntza askotan (urrutiago joan gabe, frantsesez) egiten den bereizketa ez baitu egiten. Kontua da batean irabazten dena bestean galdu egiten dela. Baina frantsesa bera baino konplexuago ere bada euskara, ‘subjuntiborako’ eta ‘potentzialerako’ ere beste aditz laguntzaile batzuk (*EZAN/ *EDIN) ere bai baititu euskara batuak. Hots: IZAN/EDUN alde batetik eta EDIN/EZAN bestetik.

Aditz batzuk laguntzaile gisa bakarrik erabil daitezke: *EDIN eta *EZAN (*isil GAITEZEN; ez GENEZAKE ezer esan...*). Dena dela, adizki guztiak eskema berean sar daitezke: ‘trokel’ bertsuaren arabera osatuak dira. (6)-ko eskema hori hirukoitza da. Baina beste zerbait ere gogoan izan behar da. Izan ere, adizki analitikoak osatzeko aditz ‘jokatugabe’ bati laguntzaile bat itsasten zaio: {*etorri/etortzen/etorriko*} + *da* > *etorri da / etortzen da / etorriko da*. Aditz nagusia erro soila ere izan daiteke: *etor*, *ekar*, *ibil*... Baina hauek *EDIN/*EZAN laguntzailea eskatzen dute: *etor dadin*, *ekar dezaten* (ez, **etorriko dadin/ *etortzen daiteke...*). *Etor dadin*, *eros deza-kegu* moduko adizkietan aspektu markarik ez da, beraz aspektugabeak dira. Baina adizki sintetikoak ere (*dator/dakar...*) aspektu markarik gabeak dira. Hortaz, *ETOR DADILA / DATORRELA, ETOR DAITEKE/DATORKE*...denak aspektu markarik gabeak dira, formalki behintzat. Diferentzia bakarra da adizki sintetikoetan aditz erroa adizkiaren barnean txertatua ageri dela: *da-TOR*,

da-RAMA-gu. Aldiz *EDIN/*EZAN laguntzaileak laguntzailearen ezker aldean, kanpoaldean kokatuak joango lirateke: *ETOR naiteke*, *ERAMAN dezaket*, *ESAN dezagun*... Gaztelaniarekin alderatuz gero, konplexutasunari begira, hemen ‘indikatioan’ ez bezala berriz ere forma sintetikoetara hurbiltzen dira gure euskal adizkiak: aspektu markarik gabeak baitira *daramagu*, *eraman dezaket*, *esan dezagun*. Eta euskararen eta gaztelaniaren arteko aldea oraingoan murriztu egiten dela esan genezake.

Hortaz, honela eman liteke aditz jokoaren oinarritzko mapa, nolabait esateko. Adizki jokatuak dira konplexutasun handiena dutenak. Bi modutakoak ditugu: a) *sintetikoak* hitz bakarrean ematen dira (*d-a//d-a-to-z-ki-t//d-u-te...*) eta b) *analitikoak*, laguntzaile gisa beste adizki jokatu bat hartzen dutenak (*etorriko + d-a; izan + z-ir-e-n; lagundu + d-i-egu...*). Hots:

(7)	Sintetikoak (= hitz bakarreko adizkiak)	da; datozkit; dute; zenuten,...
	Analitikoak (= aditz jokatuak + laguntzaile jokatuak)	joan nadin; ekar dezaket {ekarri/ ekartzen...} dut ekarriko nuke...

Adizki analitikoetan laguntzaile lanetan ari daitezkeen aditzak ez dira asko: *IZAN* (*etorri DA*, *ibili ZEN*, *ibili GINEN*, *joango LITZATEKE*...) eta *EDUN (*ikusi DU*, *jan GENUEN*, *jakin BALU*...). Adizki sintetiko gisa ere erabiltzen dira aditzok: *astirik ez DUT*; *nahikoa GENUEN*; *buru pixka bat BALU...ederra DA*, *politak ZIREN*... Adizki analitikoen eraketan ere parte har dezakete bai aditz jokatu bezala, bai laguntzaile gisa: *han izan nintzen*; *lagunak izango bagina*... *buru pixka bat izan balu*; *beti izaten zuen aitzakiaren bat*...

Badira, ordea, beste aditz batzuk laguntzaile gisa bakarrik erabiltzen ahal direnak: *EZAN, *EDIN dira nagusiak: *joan + DADIN*; *esan + DIEZAION*...

(8)	soilik laguntzaile	*EZAN, *EDIN
	Analitikoak (= ad. jokatuak + ad. laguntzaile jokatuak)	EKAR dezaket/ EKARRI dut/ EKARTZEN dut/ EKARRIKO nuke

Forma sintetikoak duten aditzak ere apur batzuk baino ez dira: *egon*, *etorri*, *ibili*, *joan* eta beste bakarren batzuk, tartean ‘defektibotzat’ edo har litezkeenak: *erion*, *etzan*, *jarraiki* bereziki. Iragankorrak gehiago dira baina hauek ere urriak: *eduki*, *ekarri*, *erabili*, *eraman*, *eroan*, **io*, *jakin* bereziki, eta *entzun*, *iritzi*, *erran*, *ezagutu*, *ihardun*, *iharduki*, *ikusi*, *iraun*, *irudi*, defektibotzat har litezkeenak seguru asko.

5. Moduaz, aldiaz, eta aspektuaz

Adizki jokatuetan biltzen den informazioa aditz *erroari* dagokiona da batetik. Aditz bakoitzaren eduki lexikoa honek ematen du: *dator/ etorri da* diogunean -TOR- erroa ageri da, *etorri* aditzaz ari baikara. Eta *dakar/ekarriko zuen* adizkietan berriz -KAR- erroa dugu. Oraingoan *ekarri* aditzaz ari gara, noski.

Erroaz gainera, *aspektuari*, *moduari* eta *aldiari* dagozkion *flexio* markak biltzen dira adizki bakoitzean. Eta perpauseko *subjektu* eta *objektu* gisa parte hartzen duten sintagmekiko *komunztadura* ere bai.

Adizki sintetikoetan ez dago aspektu markarik. Alegia, *dena DAKI* edo *berehala NATOR* diogunean, aspektu markarik gabeko adizkien aurrean gaude. Aldiz, *egunero etortzen da* edo *ongi ikasi duzu* diogunean, lehenbizikoak *aspektu burutugabea* adierazten du, bigarrenak, *aspektu burutua*. Baina *etortzen* eta *ikasi*-ren parean *ikasiko* (<*ikas-i-ko*)/*etorriko* (< *etorr-i-ko*) ere hor ditugu. Eta inolako aspektu markarik ez dutenak ere baditugu: *ikas* eta *etor* adibidez. Hortaz, *lau aukera* ditugu aspektuari dagokionez: a) aspektu gabea; b) burutua; c) ez-burutua eta d) burutzekoa.

(9)

aspektua	aditz erroa	modua	aldia		
			[+oraina, -iragana] ‘orainaldia’	[+iragan, -oraina] ‘iraganaldia’	[-oraina -iragana] ‘balizkoa’
aspekurik ez	-IZA- -TOR-	-KE	DA DATOR	ZEN ZETORREN	-LITZ -LETOR
		+KE	DATEKE DATORKE	ZATEKEEN ZETORKEEN	LITZATEKE LETORKE
	IZAN ETOR	-KE	DADI- DEZA-	ZEDIN ZEZAN	LEDI- LEZAN
		+KE	DAITEKE DEZAKE	ZITEKEEN ZEZAKEEN	LITEKE LEZAKE
[+burutua]	ETORRI IZAN EKARRI *EDUN	-KE	DA DU	ZEN ZUEN	ba-LITZ ba-LU
[-burutua]	ETORTZEN IZATEN ₁	+KE	DATEKE DUKE	ZATEKEEN ZUKEEN	LITZATEKE LUKE
	EKARTZEN IZATEN ₂				
‘burutzekoa’	ETORRIKO IZANGO EKARRIKO IZANGO				

Modu horretan aurkezten ditu, funtsean, adizkiok Euskaltzaindiaren *EGLU II*-k. NEN-TORREN/ba-NENTOR// ZETORREN/ba-LETOR... adizki ‘bixkiak’ dira, nolabait esateko. Ezagun dute aurpegian, morfologian, alegia: lehen pertsonaz (NI) ari garelarik N- agertuko zaigu bai *iraganean* bai *balizkoan*. Hirugarren pertsonaz ari bagara, berriz, Ø ‘agertuko’ da pertsona-komunztadura marka gisa. Baina batean -EN agertzen den bitartean, besteetan ez da holakorik ageri. Modu berean, ZEN/ba-LITZ// ZIREN/ba-LIRA adizkiak.

Zetorren, nentorren, zen, ziren iraganeko formak dira: *letor, banentor, balitz, balira...* ez dira orainaldikoak ez eta ere iraganaldikoak. Erraz onar daitezke honako esaldi hauek: *bihar baletor, serioagoak balira...* beste hauek, berriz, nekez: **joan den urtean baletor, *lehenago serioak balira*. Hauen aldi-balioa gorabehera, EGLU-n ‘alegiazkotzat’ ematen dira. ‘Supositibo’ deitura ere erabili izan da noizpait gure gramatiketan. Inoiz pentsatu dut ‘menturazko’ ere erabil litekeela. Edo, behar bada, ‘balizko’ balioa duela esan genezake. Horregatik, azkenean termino honen alde egingo dut (cf. *Balizko oleak burdiarik ez*). Balizko kontuez dihardugunean erabiltzen ditugu, izan ere. Uste dut azken batean *alegiazko* baino egokiago dela *balizko* hau. Hortaz, aspektuari dagokionez lau aukera eman diren bezala, aldiari buruz bereizketa hirukoa da. Alegia, *hiru aldi* izango genituzke.

Bestalde, forma beretan oinarri hartuta bestelako adizki batzuk ‘atera’ ditzakegu. Berezik, hor goiko hirukote horri -(TE)KE morfema erantsirik ahalezkoak edo geroaldiko formak sortzen zaizkigu:

- (10) a) da --> dateke/ dator --> datorke/ daki --> dakike...
 b) zen --> zatekeen// zetorren --> zatorkeen/ zekien ---> zekikeen
 c) litz --> litzateke// letor --> letorke// litz ---> litzateke

Ezarri ditugun taula hauek laburbiltzen ahal dute adizki sintetikoen jokoa. Morfologiari begira, bi multzo nagusi ditugu: DA/DU saila lehenbizi, orainaldiari dagokiona. Honen ezaugarri morfologikoak begien bistan ditugu: irregulartasun batzuk gorabehera, denetan elementu berak aurkitzen ditugu. IZAN-en irregulartasuna nahikoa nabarmena da, nolana ere. Normala, horixe baita maiztasun handiena duen aditza, corpusen azterketa arrunt batek bistaratzen ahal duen bezala. Erabiliena bera da eta erabiliena izaki, ez da harritzekoa historian zehar aldaketa zenbait jasan izana. Eredu gisa jarri dugun etorri aditzak ongi erakusten dunola osatua den: IZAN eta *EDUN-en kasuan erroa, esate baterako, askoz itxuraldatuago dagoela ematen du. Dena dela, hauek ere ‘gordetzen’ dute nolabait adizki guztien taxu bera.

Eskema berean sartzen ahal ditugu adizki jokatu guztiak, sintetikoak nahiz analitikoak.

Adizki hauei erreparatzen badiegu, irregulartasun batzuk gorabehera, erraz ikusten ahal da nola osatuak dauden. Adibidez, ETORRI edo EGON aditzaren kasuan, orainaldian, -TOR- / -GO- erroa ageri-agerikoa da: dator < da-TOR; nator < na-TOR; hator < ha-TOR; gator < ga-TO-Z; dago < da-GO; nago < na-GO. IZAN aditzaren kasuan gauza ez dago hain agerian: da < da-ø; naiz < na-IZ; haiz < ha-IZ, haiz < ha-IZ, baina gara < ga-RA, zara < za...

*EDIN/ *EZAN laguntzaileek ere kontraste bertsuak erakusten dituzte egiturari dagokionez. Hirugarren pertsona singularra hartzen badugu kontuan:²

hori gerta da-di-n	hori gerta ze-di-n	hori gerta ba-le-di-ai-ledi gerta ²
hori egin de-za-n	hori egin Ze-za-n	hori egin ba-Le-za-ai-LEZA egin
hori gerta da-i-teke	hori gerta ZI-teke-en	hori gerta LI-teke
hori esan de-za-ke	hori esan ze-za-ke-en	hori esan le-za-ke

² *ailedi/ aileza* forma zaharkituak dira baina idazleen eskura daudeke betiere.

Hortaz, euskal aditz laguntzaileak ere betetzen du aditz sintetikoen funtsezko egitura, sintetikoak baitira laguntzaile direnean ere.

Jakina, eskema honek badu bere konplexutasuna, baina adizki guztiek eskema berari erantzuten diote. Horretan jarri behar lukete arreta euskara ikasleek eta irakasleek. Eta mamuak uxatu.

Morfologiari dagokionez, aditz jokoaren konplexutasuna gaztelaniaz eta euskaraz aski parekoa dela esango nuke. Hala ere, euskarak badu alde on bat: sintetiko gisa jokatzen duten aditzak gutxi batzuk besterik ez dira, laguntzaileak barne. Aldiz, nolabait gure adizki sintetiko urrien aldean (*nator*, *daki*, *dabiltza*, *daramagu*...) askoz gehiago dira gaztelaniazkoak, hemen aditz guztiek joka baitezakete nolabait ‘sintetiko’ gisa (*vengo*, *sabe*, *andan*, *anduvimos*, *llevamos*, *lleve*, *crecierón*, *dijera*, *comprariámos*...) euskaraz ez bezala (*hazi ziren*, *esango banu*, *erosiko genuke*...), non aditz gehienek ez duten joko sintetikorik.

Gaztelaniazko gramatiken arabera, aditz morfologian erroa eta desinentzia bereizten dira. Euskaraz ez dugu desinentziaz hitz egiten, hitzaren ‘amaiera’ baita ‘desinentzia’ hori, adizkiaren eskuin muturrean doana. Euskaraz adizkion atalkatzea egiten dugunean, erroa ez doa hasieran, erdian doa nolabait. Hortaz, desinentziaz ez hitz egiteak badu arrazoia. *Caminábamos* aditzaren erroa *camin-* bada, honi eransten zaiona da desinentzia: *caminábamos* < *camin-avt-batm-mospn*. Desinentzia honetan biltzen da aditz jokatuaren flexioari buruzko informazioa: a) *vt* ‘bokalk tematikoa’: *caminar* aditzaren kasuan, -a, *temer*-n kasuan, -e, *servir*-enean, -i; b) *tm* ‘denbora eta modua’; *pn* ‘pertsona eta numeroa’. (RAE 4. atala)

Konplexutasunaren aldetik, *caminábamos* aditzaren analisiak ez dio enbidia handirik euskal aditz jokatuari eskaini ohi diegun analisiari. Ingelesarenak ere ez, bere sotiltasunean. Euskaraz bokalk tematikorik ez denez, edozein aditzek berdin balio digu, baldin eta jokabide sintetikoa onartzen badu. Har dezagun, esaterako, *ibili* aditzaren forma bat: *genbiltzan* < *g-en-bil-tza-n*. Adizki honen atalkatzea eginez gero, euskal aditzak gaztelaniazkoak baino atal bat gehiago duela ohar-tuko gara: *g1-en2-bil3-tza4-n5*. Izan ere, gaztelaniazko *caminábamos* aditzean bertan ere, morfema bera bikoitza da: *-mos*-ek pertsona eta numeroa ‘biltzen’ ditu. *-ba-* morfemak ere denbora eta modua (eta aspektua) ‘biltzen’ dituen bezala. Elkarren parekotzat har litezke, dena dela, nahiz eta hemen ere euskarak ‘alderantzizko’ hurrenkerara behartzen gaituen. ‘Desinentzia’ delako hori, gure kasuan, erroaren ezker-eskuin kokatua ageri da: *g-en-bil-tza-n*. Hori bai diferentzia, baina konplexutasuna, bertsua. Morfemen hurrenkerak eta kokaguneak goian eman dugun sintagmaren kasuan bezala, kontraste bera erakusten dute. Hemen ere, morfemak nonbait kokatu behar direnez eta nahas-mahas ezin direnez ezarri, beren lekua izan behar dute.

Euskaraz aditz jokatuak bi multzotan banatzen direla ikusi dugu: a) *adizki sintetikoak* (*naiz*, *nator*, *geundeke*..., *dugu*, *dakigu*, *lekarke*...), eta b) *adizki analitikoak*, adizki jokatu batez eta aditz laguntzaile jokatu batez osatuak. Adizki sintetiko tankeran erabiltzen direnak guxi batzuk baino ez dira: hauetako batzuk (*EZAN eta *EDIN) laguntzaile gisa baino ez dira erabiltzen. Gaztelaniaz, aldiz, denak erabiltzen ahal dira aditz ‘soil’ gisa, sintetiko gisa alegia: *veníamos*, *temisteis*, *cantaríamos*... nahiz analitiko edo konposatu gisa: *habíamos venido*, *habéis temido*, *hubierais cantado*, *habríamos protestado*... Gainera, aditz hauetako asko irregularrak dira: *caminar* -> *caminasteis* vs *andar* -> *anduvisteis*; *temer* -> *temió* vs *placer* -> *plugo* eta abar. Euskarak holakorik ez.

EGLU II-k proposatzen duen aditzaren egitura hirukoitz horren jatorri zuzena Allieres-en (1983) proposamena da eta honen azpian Lafitte-rena (1944) dago. Morfologian oinarritua da,

ez semantikan³. Arazoa, batez ere hirugarren zutabeko adizkietan dago. Zer da subjuntiboa, adibidez? Lafittek behin baino gehiagotan aipatzen du bere gramatika ospetsuan halako formak frantsesez subjuntiboa deitzen zaion horren itzulpen direla, baina berariaz ez du subjuntiborik aipatzen euskararentzat. Gogoan dut ikasle aleman batzuk nola etorri zitzaizkidan behin gaztelaniazko ikastaro bati buruz galdezka. Haien kezka nagusia, aitortu zidatenez, gaztelaniazko subjuntiboa ikastea zen. Buruko min handia ematen zien nonbait. Subjuntiboa zer datekeen ez da ulertzen erraza, ez eta zein balio semantiko duen ere. Ez dago garbi euskaraz subjuntiborik baden edo behar ote den ere. EGLU-n ez da jasotzen, EBE-n bai. Alboan gaztelania bezalako hizkuntza bat dugularik, ihes ezin egin diogun kontzeptu bat da subjuntiboa. Euskaraz subjuntibotzat hartzen dira *etor dadin*, *etor dadila*, *egin dezan*, *egin dezala* moduko aditzak. Forma hauek, ordea, [aditza + beste zerbait] ere badira: hor dute eskuinean -en/-ela konplementatzailea *edin/*ezan aditz laguntzaileari itsatsirik. Baina berez, *etor dadila/ egin dezan* eta *etorri dela/ egin duela* bezalakoxeak dira. Diferentzia bakarra aditz laguntzailean eta adizki ‘nagusian’ daukagu: {*etorri/ etorriko/etortzen*} *da/*dadin* vs *etor* {**da/dadin*}. Alegia, *edin/ezan* laguntzaileek aspektu markarik gabeko aditz bati itsasten zaizkio. Aldiz, *ukan/izan*, aspektu marka duenari. Hori morfologian argi dago. Baina honek zer esan nahi du? Aspektu markarik ez duen oro subjuntiboa dela? Ezin hori esan, zeren ahalezko adizkiek ere ezaugarri horixe dute: *etor daiteke* baina {**etorriko/*etortzen/*etorri*} *daiteke*. *Datorrela hurrengo*, *goazen kalera* diogunean eta *etor dadila hurrengo*, *joan gaitezen kalera* diogunean ere berdin. Morfologiaren ikuspegitik ez dago alde handirik *joan gaitezen* eta *goazen*-en artean. Aditz bera dugu. Diferentzia atzizkian dago. Atzizki hau, bistan da perpaus osagarrietan erabiltzen dugun bera da. Hots: Allieres-en ikuspegia zuzena da. Arazoa da ez dakigula zer den juxtu subjuntiboa. Dakigun bakarra da beste hizkuntzetako subjuntibo deituei modu batera edo bestera ordainak aurkitzen ahal dizkiegula, baina ordain horiek ez dira beti bat eta bera izaten.

Horregatik, EGLU II-n paradigmetan ez da jasotzen *subjuntibo* hitza. Dena dela, euskara gaztelaniarekin eta beste hizkuntza askorekin alderatuta, nekez esango dugu besteak baino konplexuago denik alde honetatik. Baina bai, nire gazte alemaniar haien kezka ikusita, subjuntiboaren erabilera ere arazotsua da.

Terminologia ere aski nahasia gertatzen da, euskaraz bakarrik ez, hala ere. -edin- eta -ezan- aditz laguntzaileen gainean eraikitako formei ‘subjektibo’ deitzen die Allières-ek eta *edun/ izan-en gainean eraikiei berriz, ‘objektibo’. Hau ere, beharbada, hankamotz gelditzen da baina bide zuzenagotik doala esango nuke, indikatibo/subjuntibo bereizketa baino. Nolanahi ere, EGLU-ko taulan formen banaketa argia dela uste dut. Baina taula horretako zein ‘laukitan’ kokatu beharko genuke subjuntiboa? Garbi dago lau laguntzaileak adizki jokatu gabe desberdinekin elkartzen direla: erro soila dugunean (*etor/lagun...*) *EDIN eta *EZAN laguntzailea hautatzen dute. Beste hiruak DA/DU. «Mais le problème qui se pose est celui de leur dénomination, car là presque chaque descripteur a créé son propre vocabulaire fondé sur la sémantique (...) la terminologie reste souvent hésitante et mal adaptée parce que toujours plus ou moins rattachée aux usages romans» (Allières 1983: 41). Honen ondorioz, morfologian oinarritutako sailkapena proposatzen du. Eus-

³ Hain zuzen, Allières-ek lanaren hasieran honela dio: «On peut distinguer trois principaux types d’erreurs parmi celles où ces auteurs sont tombés. (...) Seconde fausse manoeuvre: légiférer sur le verbe basque en general en prenant pour base exclusive de référence un système local. Et l’inconvénient d’une telle pratique s’aggrave si vient s’y joindre un autre choix néfaste: celui de définir les divers tiroirs verbaux en termes sémantiques plutôt que morphologiques. En effet, comme le même type de formation peut revêtir selon les régions des valeurs et des fonctions différentes, on perd de vue le lien qui unit morphologiquement les divers systèmes locaux en un ensemble à peu près homogène partout, compte tenu des particularités d’usage ou des phénomènes de supplétion propres à chaque localité ou ensemble dialectal.

kaltzaindiaren paradigmek ere bide horri segitu beharko liokete. Gaztelaniazko sistemak ez digu asko laguntzen. ‘Subjuntibo’ deitura horrekin ezagutzen ditugun aditzak ez dira mota berekoak euskaraz eta gaztelaniaz. Hona hemen adibide batzuk, non euskarak eta gaztelaniak gehienetan bat egiten ez duten:

- (6) a. No conozco a nadie que *haya leído* ese libro.
a’. Ez dut liburu hori *irakurri duen* inor ezagutzen
- b. Pocos alumnos habrá *que hayan aprobado*.
b’. Ikasle gutxi izango dira *gainditu dutenak*.
- c. Se necesita secretario *que sepa* inglés.
c’. Inglesez *dakien* idazkari bat behar da.
- d. La próxima vez *que venga*, te haré una visita.
d’. Hurrena *natorrenean*, bisita bat egingo dizut.
- e. Si *hubiéramos* tenido tiempo, *hubiéramos* ido a visitarte.
e’. Astia *izan bagenu*, zu ikustera joango ginen.

(ikus Porto Dapena, J. (1991), 240 eta hurr.)

Gauzak horrela, tentu guztiarekin jokatzearabaki zuten EGLU II-ren egileek. Han egiten zen proposamen bat, badaezpada beharbada, baina bidezkoa, nire ustez. Testu liburuetan ere jasoa izan da (ik. J. Butron & J. De Pedro (1988): 127 eta hurr.), Gramatika Batzordeari baimena eskatuta, gainera.⁴

6. Pertsona/numero komunztadura eta abar

Hona hemen euskal aditzaren egiturari dagokion beste mamua: subjektuaren pertsona-numero komunztadura nahikoa ez, eta objektu zuzenarekin eta zehar-objektuarekin ere komunztadura egin behar. Hizkuntza ikasteko ahaleaginean ari denarentzat zailtasun bat gehiago. Baina ondo pentsatzen jarrita, gure inguruan eta gure artean nagusi eta harro dabilen beste hizkuntzetan, gaztelaniaz esaterako, antzeko zerbait gertatzen dela esango nuke izenordain klitikoekin. Eta inor ez da asaldatzen horregatik. R.A.E-ko gramatikak orrialde mordoa eskaintzen dio kontu honi. Bestalde, klitiko horiek maiz beharrezkoak dira, R.A.E-k dioen bezala. Hain zuzen, euskaldun askori, ustez dotore-antzean erdaraz mintzo direlarik, ateratzen zaizkie holako batzuk, behar den klitikoa ‘irentsiz’. Adibidez: «*Cómprame el periódico*» eskari/agindu edo dena delakoari erantzunez: «*No te he comprado; luego te compraré*» moduko erantzuna aski ohikoa da. Hala ez balitz, RAE-ko gramatikaren egileek ez zuketen honen aipurik egingo.⁵ Dena dela, espainol estandarrean, honela behar luke erantzun horrek: «*No te lo he comprado; luego te lo compraré*». Gaztelaniazko *te-lo-he comprado* hitz-segida eta gure *erosi dizut* aditza (<eros *-i+d-i-zu-t*) gauza

⁴ Harrigarriago egiten da, ordea, aurten bertan, Euskaltzaindiaren «ehuntzea» ospatzeko kaleratu den *Euskara Batuaren Eskuliburuan (EBE)* noizbaiteko aurkezpen zaharrarekin jarraitzea (514-522 orr.) oharririk txikiena ere egin gabe gauzen zergatia azaltzeko. Eta hori, bibliografian, EGLU II oinarritzko erreferentzia gisa aipatzen dutelarik.

⁵ «En el español hablado en ciertas zonas de las áreas rioplatenses y andina, así como del País Vasco (España), se elide a menudo el objeto directo de los verbos transitivos sin que ocupe su lugar un pronombre átono. El proceso se registra incluso con grupos nominales de interpretación específica: *Cuando termine la carta llevaré al correo (...)*. La ausencia de pronombre átono en estos contextos parece deberse a la influencia del quechua (...) en el primer caso y a la del euskera, en el segundo» (RAE 2010 657-658). Ez dakit hori datekeen euskaldunen kasuan gertatzen denaren arrazoia, euskaraz klitiko horiek hor segitzen baitute aditzean itsatsirik, Ø gisa bada ere: [eramango Ø-d-u-t] .

bera ez badira, gutxi falta zaie. Gaztelaniaz ari garenean *te-lo-h-e comprado* gertatzen zaigun bezain arrunta, hain zuzen.

Jakina, erantzuna berehala etorriko zaigu esanaz hor pertsona eta numero informazioa hitz bakarrean ematen zaigula euskaraz eta gaztelaniaz hirutan. Ahoskerak ordea beste zerbait adierazten digu: gaztelaniaz *nóteloh-e /comprado*, modu arruntean emanda. Euskaraz berdin: *eztítut /erosi*. Idazkera alde batetik doa eta ahoskera bestetik. Hortaz, hemen gaztelaniaz eta euskaraz elkarren antz pittin bat badutela esango nuke. Gaztelaniazko izenordain horiek *klitikoak* dira. Eta klitiko horiek, gainera, euskaraz ez bezala, gaztelaniaz erredundanteak gertatzen dira batzuetan: *yo me tomo todos los días un vaso de leche fría* perpausean hiru aldiz jasotzen da lehen pertsonari buruzko informazioa. Objektuari buruzko informaziorik ez da jasotzen, baina batzuetan jaso beharra dago (cf. *‘esos vasos de leche yo me los tomo de un trago, sin respirar’*). Euskaraz hainbestearinoko erredudantziarik ez da gertatzen subjektuari dagokionez behintzat: *egunero baso bat esne hotz hartzen dut* (<*ø^{abs3sg}-du-t^{erg1sg}*). Beraz, hizkuntzen konplexutasunaz ari baldin bagara, ez dago, funtsean hizpide ditugun bi hizkuntzen artean hainbesteko alderik: biak eder, biak on, biak antzeko. Hau oso modu argian erakusten du B. Fernandez-ek (2016: 121-142 orr.). Bide batez, gure hizkuntzalari zuhurrak Wikipediako eta beste hainbat lekutako euskal aditz jokoaren *‘gehiegizko konplexutasunaz’* baino nahiago du *‘euskal aditz inflexioaren izaera sofistikatuaz’* hitz egin. Dotore; bejondeiola!⁶

7. Aditzekin jolasean: *Darabagu, datsegit, bazeniezazkigukete, nirudike, naiozu...*

Lantxo hau amaitzeko ahaleginean, ezin diot uko egin egunotan zabaldu berri den datorren urteko *Araba Euskaraz* festari erantsi dioten goiburua: *Darabagu*. Atsegin dut txantxa, baita hizkuntza bezalako gai serioetan ere. Hizkuntzak eta gramatikak ere txantxarako aukera eman dezakete eta eman behar ligukete. Erratza sudurraren puntan jarrita, tente, metro batzuetan eramatea bezala da: erratzaren funtzioa ez da hori, jakina, baina lurreko zaborrak batu ostean metro batzuk egitea erratzak sudur puntatik ihes egin gabe ez dago gaizki: horregatik ez zaio errespetorik galtzen erratzari. Tresna sofistikatu horrek berez dituen aplikazioak *‘ukitzeke’* gaindi daiteke ariketa. Gure kasuan adizki edo adizki usteko horiek ere ez dute adierazten euskal aditz beneragarriarekiko errespetu falta. Txantxa horrek hizkuntzaren balizko konplexutasuna gaintzeko edo garrantzia kentzeko (!) balio dezakeen neurrian ongi etorria izan bedi. Aditz jokoak gure kirioak lasaitzeko balio badu, ongi etorria, askotan kontra aritzen baikara.

Jar gaitezen serio. *Darabagu* aditz bat da ala aditz tankerako asmazio bat, asmatutako aditz *‘berri’* bat, alegia? Eta zer esan nahi du? Nahi duzuna, edo ezer ez. Fantasia hutsa izan daiteke, euskal adizki peto-petoaren tankera duena.

Gorago aipatu dudan Altunaren liburutxoan, hasiera-hasieratik esaten zitzaigun euskal aditza hiru zutaberen gainean eraikitzen dela: NOR, NORI eta NORK zutabeen gainean. Saia gaitezen aditz hauen arkitektura osatzen hiru zutabe horien gainean. Badakigu, NOR dela ezinbesteko zuta-

⁶ *Bejondeiola* honek ere adizki sintetiko baten itxura guztia du. Euskaltzaindiaren *Hiztegiak* *‘norbaiti zerbait onuragarri desiratzeko esaten zaion esapidea’* dela dio. Euskara-irakasle hasi berri nintzen garaian, egun batez, ikasgelan sartu eta arbelean norbaitek idatzia zuen honako hau: «Patxi bejondeizula bat da». Ez nuen ulertu. Ikasleak, eserita, irribarrez, begirik kendu gabe, arbelari so, etsita. Azkenean, batek erakutsi zidan ikasliburuko hiztegitxo, non *bejondeizula*-ren ordain gisa *‘bravo’* ageri zen. Jakina, ez nien azaldu esapide horren egitura konplexua zein zen. Aski izan zen interjekzio bat zela esatea. Denek barre.

bea eta beste biak komunztadura premien arabera erabiltzen direla. Forma alokutiboak albo batera utziko ditugu.

NOR	NORI	NORK
N-	-IDA- /-IT	-T
H-	-I(K)A-/-IK(-INA-/-IN	-K
d-	-IO-	ø
G-	-IGU-	-GU
Z-	-IZU-	-ZU
Z-	-IZUE-	-ZUE
d-	-IE-	-TE

Zutabe hauen gainean aditz erroa ezarri behar da. Gure *darabagu* horretan D_{NOR} -ARABA-GU_{NORK} horren erroa RABA edo ARABA bada... hor dugu aditza dagoeneko eratua: D-ARABA-GU. Eta hau daukagun neurrian, balizko D-ARABA-KE-GU, edo D-ARABA-TE, N-ARABA-ZU... Datiboa ere behar bada, D-ARABA-KIO-GU eta antzekoak osa daitezke erraz-erraz. Baina txantxetako adizki horri erroa alda diezaiokegu -RAMA- esaterako eta horrela, benetako aditzetara ere irits gaitezke: *daramat*, *daramate*... *Datsegit* ere txantxetako aditza litzateke. Baina euska aditzaren zirrikietan barrena trebatzeko egokia izan liteke hau ere. Horra beste adizki ‘fantastiko’ bat.

Eta zer esaten didazue *zenieazkigukete* adizkiak? «Lekua hartu, mingaina trabatu, burmuina berotu eta komunikazio mailan deus ere gaineratzen ez duena» (Salaburu 2002: 177). Baina, jakina, *darabagu* ez bezala, *zenieazkigukete* egiazko aditz bezala sortua da, *zenizkigukete* izan litekeen bezala. Eta hau bezalakoa litzateke *zenigukete* edo *zeniguten*. Hortaz, non dago *zenieazkigukete* horren akatsa? Salaburuk aipatzen dituen trabetan, hain zuzen. Zer egin orduan? Euskara batuko adizki izaeraz jantzia ageri da hor zehar dabiltzan paradigma guztietan, nahiz eta inork ez duen erabiltzen, corpusek erakusten duten bezala. Esango nuke, hala ere, aditz batuaren zerrendetan ageri bada, badagoela nonbait eta badela. Baina inor ez dago behartua erabiltzera. Eta pentsatzekoa da, halaber, inork ez duela erabili ez erabiliko ere. Idatziz betiere errazago litzateke, baina testu idatzi hori norbaitek irakurri behar badu, jendaurrean adibidez, hobe, *izki*, *gu* eta *ke* artean ito nahi ez badugu, beste bideren bat hartzea: hasteko datiboa bera ordezkia daiteke modu askotara. *Eraman* aditzarekin, adibidez errazago gertzatzen da *X-entzat* edo *X-ra* erabiltzea. Horrek adizkiari pisu dezente kenduko lioke. Ahalezko forma denez, gainera aditz analitiko bat osatu beharrean perifrasi modura ere saia gaitezke *ahal*-ez baliatuz: (*zuek*) *eraman ahal(ko) zenioke*/ *Andonirengana eraman ahalko zenukete*.

Atal honen izenburuan beste aditz bat ere ezarri dugu: *nirudike* nahiz *nirudien*. Hau ez *dadarabagu*-ren antzeko aditz fantastikoa, izatekotan ere aditz ‘fantasma’ litzateke, aurrekoa bezala. Honek, noski, egitura eta itxura soilagoa edo sotilagoa du, baina lehenbiziko aldiz aditz zerrenda batean lekua bazuela ikusi nuenean zer demontre izan zitekeen ezin asmatu ibili nintzen. Gero *zirudien*, *dirudi* eta holakoen ‘arrimuan’ zentzua hartzen joan zen. Hala ere, kontrabandoko adizkia dela esango nuke. Nire euskaraz behintzat hirugarren pertsonako subjektua eskatzen du eta ergatiboa, *horrek* adibidez: *horrek inozoa dirudi* ‘inozoa ematen du’. Gaurko gramatiketan, aditz inergatiboa dela esango genuke, beharbada. Alegia, subjektu ergatiboa eta objekturik gabea, *ira-*

kin-en antzekoa. Dena dela, *iruditu* aditzaz ari al gara? Ezetz uste dut. *Iruditu*, *Euskaltzaindiaren Hiztegi*an, ‘zaio’ motako aditza da. *Irudi* aditza, ordea, ez da ageri hiztegi horretan, nahiz eta aditz batuaren paradigmetan ageri-agerian dagoen. Beraz, aditz defektibotzat hartua izatea merezi luke, beharbada. Baina ez hiztegi ‘falta’ delako, baizik teoriarik izan beharko litzatekeen forma batzuk ez dituelako: *dirudi*, *zirudien* bai, baina agian ?*dirudit* edo ?*nirudien*/*nirudike* modukoak handiegi gelditzen zaizkio hain aditz kokorrari. Zerbaitengatik du aldameneari gure zerbitzurako eman aditza bera: *apaiza dirudit/nirudien-en ordeztu, apaiza ematen dut*. Gauzak nik uste bezala izatekotan, zerrenda zuzentzea eskatuko luke. Euskara ikasleek merezi dute horrenbeste.

Gauza bera gertatzen da *iraun* aditzarekin ere. Hau inergatiboa dela argiago dago, *irakin* bezalaxe. Ez bata ez besteak ez du objekturik onartzen: **zer iraun duzu?* / **bost urte hauek iraun ditut* (baina bai *bost urte hauetan iraun dut*). Beharbada, *irudi*-rekin gauza ez dago hain garbi, zeren *mutil honek {zuhurra/ gure aitona...} dirudi* zuzenak dira, baina *zuhurra* eta *gure aitona* sintagmak hemen ez dira objektu, izen predikatu baizik. Hortaz, *irudi* eta *iraun-en* paradigmak berriusi beharko lirateke eta dagokien lekura bidali.⁷

Beste kontu bat: EKARRI, ERAMAN aditzen berririk ez da ematen ‘alegiakoei’ formetarako EGLU II-n. Pentsatu nahi dut Euskaltzaindiaren arauan ere agertzen ez delako dela. Baina, neure harriturarako, EBE-n nola edo hala jasoak ageri dira. Hori bai, inolako azalpenik eman gabe. Bide horretatik segituz gero sor dezakegun nahaste-borrastea handia izan daiteke. Eskerrak adizki konplexu horiei inork ez diela begiratu ere egiten, nonbait. Konplexutasunaren konplexuak ez du, antza, hainbesteko indarrik. Baina balizko irakurleari esan behar diot, *nirudien* edo *nirudike* ez bezala, *balekar*, *lekarke* eta antzekoak norbaitek erabili dituela idatziz. Beraz, EBE-koei zorionak, gehiegi luzatu ez badira ere azalpenak emateko orduan.⁸

Azken hitza: errefraura berriro

Gaur hasierako errefrau hori honela parafrazea genezake beharbada: «Ai ederragoa balitz ... eta ez ona» edo «Ai ederragoa balitz ona ez izan arren», «ederra nahiago, ona baino», alegia. Hau euskararekin zenbaitek duen maitasun harremanari aplikatu ahal zaiola iruditzen zait: euskara maite bide dute, baina ez ona delako edo delakoan, duen edo dukeen edertasun miresgarriagatik, nahiz eta praktikan, oharkabea bada ere, ez duten ezertarako nahi, ez behar. Euskaraz bizi nahi dugunok ‘eder eta on’ nahi genuke gure hizkuntza, bi kualitateak bat eginik. Zenbat bider kontatu ote zigun Patxi Altunak Azpeitian ezagutu zuela mutiko bat, zeharoko ergela nonbait, eta buruari bi eskuez heldurik honela esaten omen zuena betiere: «Ai tontoagoa banintz!» Itxura horretako esaldia dugu izenburu gisa hautatu dugun errefrau hori. Baina euskararen kasuan, eta munduko hizkuntza guztien kasuan ere, honela bihur genezake errefrau hori: *euskara ailitz ederra, ona eilitz*. Edo gaurko euskaraz, *ai ederragoa balitz, eta bertute handiagokoa!* Eta *ailitz* moduko adizki zaharkitu batek hizpide eman badigu, gure taulan eskuineko zutabea kokatu beharko genuke, alegia: «IZAN-en adizki sintetikoa, balizkoa, aspektu-markarik gabea, 3. NOR singularra...». Luzetxo, agian. Ea noiz hemen hizpide izan ditugun adizki motentzat ere izen egokiak aurkitzen ditugun.

⁷ EBE-n hemen zailtasun jarri ditudan hiru aditzon paradigmatetik bat bakarra jasotzen da: *iruditu*-rena. Baina gorago esandakoaren arabera, *irudi*-rena beharko zukeen. Edo, beharbada, *irudi izan* moduko sarrera bat beharko luke, *nahi izan*-en parekoa. Ez dakit. Nolanahi ere, defektiboa litzateke.

⁸ Baina noraezean bezala ibili naiz EBE-ren informazioarekin. Izan ere, LEKARKE moduko ‘alegiakoei’ leku egin ostean, IHARDUN-en kasuan ez dute ezer jartzen. Justu EGLU II-ko jokabidearen kontra, hemen LIHARDUKE modukoak jasotzen baitira baina ez *lekarke* modukoak (238-240 orr.).

Eta azken burutazioa: *ailitz*, *ailu* modukoak zaharkituak bide dira eta bazter utziak. Ez ahal zaigu geuri ere gertatuko aditz ‘botibo’ hauei gertatua! Hots, zaharraren zaharraz baztertu eta ahaztu. Adizki zaharrak dira, baina gaurko idazle batzuek erabili izan dituzte (*Ereduzko prosa gaur-en* egin daiteke bilaketa). Ez zait gelditzen azkenean honako hau esatea baino: *Aikina gazte, euskara bezala: urtetsu baina bizi*.

Bibliografia

- Allieres, J. (1983): «De la formalisation du système verbal basque», *Piarres Lafitte’ri omenaldia* IKER 2, Bilbao, 37-91.
- Altuna, P. (1971): *Euskal Aditza. Batua? ... Baturakoa*. Etor, Bilbao.
- Amuriza, X. (2010): *Euskara batuaren bigarren jaiotza*. Lanku, Oiartzun.
- Butron, J & J. De Pedro (1988): *Euskal Gramatika Praktikoa*. Erein, Donostia.
- Euskaltzaindia (2012): *Euskaltzaindiaren Hiztegia. Adierak ea adibideak*. Euskaltzaindia, Bilbao.
- Euskaltzaindia (1987): *Euskal Gramatika. Lehen urratsak II*. Euskaltzaindia, Bilbao.
- Euskaltzaindia (2018) :*Euskara Batuaren Eskuliburua (EBE)*. Euskaltzaindia, Bibao.
- Fernandez, B. (2017): *Euskara eta hizkuntzari buruz. Hizketan semearekin*. Erein Donostia.
- Fernandez, B. (2016): *Euskera para castellanohablantes*. Erein Donostia.
- Lafitte, P. (1944): *Grammaire Basque (Navarro-Labourdin Littéraire)*. Librairie «Le Livre», Baionne.
- Lafon, R. (1943): *Le système du verbe basque au xv^e siècle*. Berrinprimatua Elkar, (1980). Donostia.
- Moreno Cabrera, J. C. (2000): *La dignidad e igualdad de las lenguas. Crítica de la discriminación lingüística*. Alianza Editorial, Madrid.
- Moreno Cabrera, J. C. (2008): *El nacionalismo lingüístico. Una ideología destructiva*. Panínsula, Barcelona.
- Porto Dapena, J. A. (1991): *Del indicativo al subjuntivo*. Arco/Libros S.A. Madrid.
- Real Academia de la Lengua Española. (2010): *Nueva gramática de la lengua española. Manual*. Espasa Libros S.L.U. Madrid.
- Rebuschi, G. (1984): *Structure de l'énoncé en basque*. SELAF, Paris.
- Salaburu, P. (2002): *Euskararen etxea*. Alberdania, Irun.
- Zubiaur, J. R. & J. Arzamendi (1976): «El léxico vasco de los refranes de Garibay», *ASJU* X, 47-143. Donostia.

Baliabide lexikoen sarea: Baldintza filologiko eta tekniko zenbait

DAVID LINDEMANN

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

IÑAKI SAN VICENTE

Elhuyar

1. Sarrera

Humanitate Digitalen (HD) arloan, berrikuntza garrantzitsuak izan dira azken urteotan, eta informatikarekiko lankidetzaz dela medio, metodologia arlo konputazionalerara zabaldu egin du filologiak. Testu klasikoaren edizio digitalak burutu eta prozesamendu konputazionalerako prestatzea da Humanitate Digitalen ardatz nagusi bat. Argitalpen historikoen edukiak testu digital bihurtzeaz gain, edizio digitala metadatuak osatzera igaro dira HD arloko ikerlariak, hala nola, anotazio filologikoak (esaterako, forma historikoa egungo formarekin lotzen dituztenak), edo anotazio linguistikoak, adibidez, hitz bati dagokion lema, analisi morfosintaktikoa edota semantikaren inguruko anotazioak. Bestetik, errekurtsoen artean loturak daramatzaten poligrafiak sortzeko ahalmena ematen du edizio digitalak (Smith, 2004).¹

Hiztegigintza historikoan ere metodologia konputazionalen etorreraren lekukoak gara, Europako hainbat hizkuntza-komunitatetan burutu diren ikerlanei erreparatzen badiegu. Hiztegi historikoa banakako baliabide itxi bezala ulertu beharrean, haren edukiak beste hainbat baliabidetako edukiekin lotzeak abantaila nabarmenak ekartzen dizkio edizioaren erabiltzaileari. Hala egitea ez da batere ideia berria, baina, Orotariko Euskal Hiztegiaren (OEH, Mitxelena & Sarasola, 1988) ekoizpen prozesuan egin zen moduan, hiztegi historikoen edukiak kartoizko txarteletan idatzi eta indexatu beharrean, hiztegi sarrerak eta haien metadatuak errepresentatzeko eredu digital aurreratuak ditugu egun. Eta haiek erabilita, edizioaren egileak eta erabiltzaileak biek dituzte onurak. Edizioen elkarreragingarritasuna bermatzeko asmotan, estandarrak garatzen ari dira. Artikulu honetan, joera berri hauen oinarri batzuk ikusi eta metodologia konputazionalaren zenbait aplikazio ikusiko ditugu, baita nazioarteko arloaren egoera euskararen eremura ekartzeko zenbait baldintza identifikatu ere. Oso oinarritzko esperimendu bi aurkeztuko dugu, baliabide lexikal ezberdinen edukiak elkartzeak zer aukera zabaltzen duen irudikatzeko.

¹ Cologne Centre for eHumanities erakundeak burututako *Vedaweb* proiektua aipa daiteke edizio digital aurreratuaren adibide gisa, Rig-Veda-ren edizio ezberdinak lerrokatu eta metadatuarekin batera interfaze publiko baten bitartez eskaintzen dituenak, ikus <https://vedaweb.uni-koeln.de/>. Kopenhagenerko ONP *Old Norse Prose*, berriz, baliabide ezberdinak interfaze bakarretik atzitzeko eskaintza da, ikus <http://onp.ku.dk/onp/onp.php>

2. Hiztegien sareak

Euskarazko hiztegi guztiak, klasikoak nahiz egungoak, banakako *stand-alone* baliabide gisa ditugu eskura, elkarrekiko lotura espliziturik gabe. Hiztegien ohiko egituran eta ekoizpenerako ohiko lan-fluxuan datza atomizazioaren zergatia: lematagia zehaztu ondoren, lema bakoitza argibide lexikografikoez hornitzera igaro ohi da hiztegitilea. Lemaren azpian, desanbiguazio sintaktiko edota semantikoa txertatzen du, hiztegi-sarrerari zer egitura eman nahi dionaren arabera. Aldi berean, edukien aurkibide bakarra lematagia izan ohi da, kasu batzuetan, sarrera batetik beste sarrera batera bidaltzen duten zeharkako erreferentziez osaturik. Hiztegi batean hiztegi beretik kanporako erreferentziak jasotzen badira, OEH-k lema-ikuraren edo haren aldaeraren agerpena hiztegi aurrekarietan aipatzen duen bezala, kanpo-erreferentzia horiek ez dira, orokorrean, edizio digitaletan landu, hau da, ez dago aplikazio informatiko batek interpreta lezakeen lotura espliziturik aipuaen jatorrizko agertokira edo sarrera horrek beste kanpo baliabideetan izan dezakeen informazio osagarrietara.

Hori dela eta, lema-ikurra (hots, lema ordezkatzeko duen hizki segida) baino ezin izan da usiatu baliabideetan zehar loturak eraikitzeke. *Euskalbar* bilaketa-tresna² eta *Elhuyar Hiztegi App*³ horrelako ekimenak dira, bilaketa bakarra hainbat hiztegitan aldi berean egitea ahalbidetzen dutenak. Bilatzen den hizki-segida bat euskarazko hainbat online-hiztegitako sarbidetara bidaltzen dute tresna horiek, emaitzarik izango ote den aurretik argitu edo bilaketa lema-ikurraz bestelako edukietara murriztu edo hedatu ahal izan gabe. Bilatzen den hizki-segida bera hainbat hiztegitako bilaketa-eremuetan eskuz sartzeko (eta emaitzak ikustarazteko) prozesua bizkortzea baino ezin diote erabiltzaileari eskaini, beraz, aipatutako tresnek. Murriztasun horren zergatia hiztegi elektronikoen gehienek datu-base egituran datza: lema-ikurrak baino ez ohi dira bilaketak ahalbidetzeko in-dexatu, ezta datu-baseon edukiak API⁴ bidez zuzenean atzitzeko prestatu ere.

Gauzak horrela, *hiztegi sare* motako baliabidea ezin izan da ekoitzi euskararako, hau da, hainbat hiztegitako edukiak ezin izan dira batu eta lema-ikurra loturak antolatzeke pibot elementutzat harturik elkarrekin eskura jarri, nazioartean hainbat kasutan egin den legez. Alemanezko *Wörterbuchnetz* hiztegien bilduma,⁵ galegozko *Diccionario de Dicionarios*,⁶ portugesezko *Corpus Lexicográfico do Português*⁷ eta nederlandararen *Geïntegreerde Taalbank*⁸ gertuko adibideak dira.⁹

Nederlandararen *Geïntegreerde Taalbank* baliabidea Europa mailan aurrekaria dela esan daiteke. Aipatutako beste ekimenetan bezalaxe, hiztegi historikoen edukiak bateratu egin dituzte, baina besteetan ez bezala, lema-ikur historikoak egungo nederlandarazko formekiko lotura esplizituek hornitu dituzte; horrela, lema-ikur historikoak nahiz egungoak erabil daitezke bilaketak egin eta beti forma historiko zein egungoari lotutako hiztegi-edukiak ikustarazteko. Jokabide hori OEHk euskararako erakusten duen irizpide filologikoarekin bat dator, hau da, forma historikoak

² Ikus <https://github.com/euskalbar/euskalbar>

³ Ikus <https://www.elhuyar.eus/eu/site/prentsa-aretoa/147/hiztegiapp-mugikorretarako-aplikazio-berria>

⁴ Ingelesez, *Application programming Interface*, aplikazio edo plataforma batek bere edukiak hirugarrenen baten eskura jartzeko erabiltzen den zerbitzua. Eduki horiek programa bezero baten bidez atzitzen dira.

⁵ Ikus <http://www.woerterbuchnetz.de>

⁶ Ikus <http://sli.uvigo.es/DdD/>

⁷ Ikus <http://clp.dlc.ua.pt/DICIweb/>

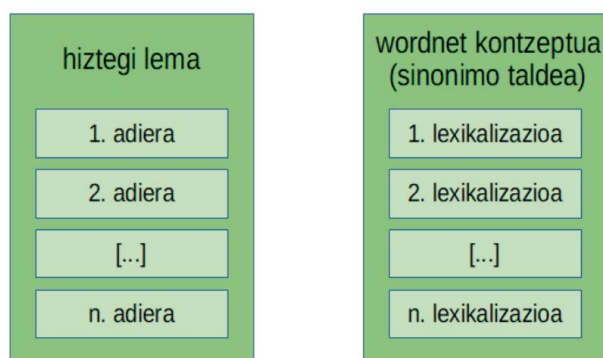
⁸ Ikus <http://gtb.inl.nl/search/>

⁹ Artikulu honetan arazo filologiko eta teknikoak ditugu hizpide. Alabaina, aipatu beharra dago hiztegien egile-eskubideak ere oztopoa izaten dira hiztegien edukiak berrerabiltzeko. Egile-eskubideak argitalpena baino 70 urte beranduago desagertzeak ahalbidetzen du hiztegi historikoen edukiak berrerabiltzeko; hiztegi berriagoen kasuan, berrerabiltzeko hitzarmenak behar dira, lizentzia irekiko baliabideak ez badira.

egungo formarekin lotzea, OEhk hiztegi historikoetako lema-ikurrak egungo lema-ikurren azpian erreferentzia gisa zerrendatzen dituen gisan. Baina nederlanderazko baliabidean, egungo lema ez ezik, forma historikoak ere bilaketarako indexatu dituzte, horrela aurkibidea corpus lexikografi-koan agerpena duten lemen aldaera guztietara hedatuz.

Esan bezala, lema-ikurraz bestelako edukiak ezin izan dira elkarrekiko loturaz hornitu, eta horrela gertatzen da aipatu ditugun nazioarteko hiztegi sareetan. Egoera hori gaindi dezakeen Hiztegi Matrix izenda litekeen baliabideak¹⁰ lema-ikurrez gain adierak ere hartuko lituzke aintzat hiztegietan zehar loturak ezartzeko. Adierak elkarrekin lotzea (*sense linking* deritzona) egun pil-pilean dagoen erronka dela esan daiteke. Adierak (*word senses*) lema homografo eta polisemikoen kasuan desanbiguatzea Lengoaia Naturalaren Prozesamenduaren (LNP) erronka nagusietako bat da. Azken urteetan hitzen errepresentazio bektorial berriei esker (*word embeddings*) aurrerapauso garrantzitsuak eman dira (Raganato, Camacho-Collados & Navigli, 2017). Hala ere, emaitzarik onenak lortzen dituzten sistemak gainbegiratuak dira, eta adierak eskuz etiketatuta dituzten datu bildumak behar dituzte desanbiguazioa egiten ikasteko. Tamalez, baliabide horiek oso urriak dira eta gainera kostu handia du horiek sortzea (Camacho-Collados & Pilehvar, 2018). Jakintza lexikografiko preziatua paperezko argitalpenetatik atera eta ustiapen konputazionalerako prestatzea, beraz, hutsune haiek betetzeko jokabide baliotsua izan daiteke.

Polisemiaren arazoaren aurrean, eta hiztegi gintzari dagokionez, lema-ikurretatik abiatu beharrean kontzeptuetan oinarritzen diren baliabideak sortzea izan da erantzun bat. *Wordnet* bezalako baliabide lexikoek, esaterako, sinonimo-talde banaren bitartez lexikalizatzen diren kontzeptuak biltzen dituzte, hiztegietan adierak lemaen azpian biltzen diren bitartean (ikus 1 irudia). Kontzeptua sinonimo-taldeko kideen (hots, kontzeptuaren lexikalizazioen) adieretako bat da, definizio bategi deskriba daitekeena eta beste kontzeptuekiko erlazio semantikoak dituena. Sinonimiak, eredu honen arabera, kontzeptu bera denotatzea esan nahi du, dagokion unitate lexikoaren erregistroari, dialektoari edo datazioari, adibidez, erreparatu gabe. Xehetasun horiek unitate lexikoari lotuta daude ikuspegi horretatik, eta ez kontzeptuari. Hiztegiak adiera mailan elkarrekin lotu ahal izateko jokabidea ematen digu Wordnet ereduak, beste hizkuntzetako Wordnet-etara nahiz bestelako baliabide lexiko-semantikotara,¹¹ kontzeptutik abiatzen diren loturek adierak zuzenean lotu, eta adieren desanbiguazioaren arazoa gainditzen baitute.

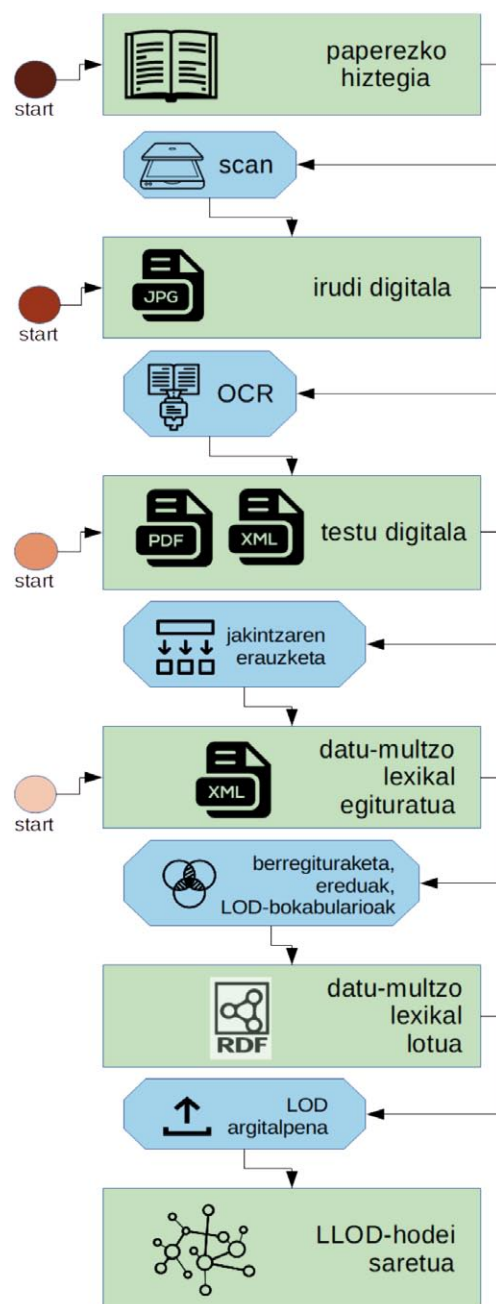


1. IRUDIA

Hiztegi lema eta WordNet kontzeptua

¹⁰ *Elexis Lexicographic Infrastructure* ikerketa-azpiegitura proiektu europarrean (2019ean abiatuta, ikus <http://elex.is>) lema eta adiera mailan lotutako hiztegien bilduma eleanitza sortu nahi dute, *Dictionary Matrix* izenpean.

¹¹ Adibidez, danierazko hainbat baliabide lexiko adiera mailan lotu dituzte, danierazko Wordnet kontzeptuak pibot elementutzat harturik (Pedersen, Nimb, Olsen & Sørensen, 2018).



2. IRUDIA

Hiztegiak digitalizatzeko lan-fluxu orokorra

3. Hiztegiak sartzeko lan-fluxua

Hiztegien edizio digitalak burutzeko bide berriak jorratu dira azken hamarkadatan. Bibliotekonomin orok har gertatzen den bezalaxe, hiztegiak historikoan ere egungo irizpideen arabera edizio digitalak, jatorrizko edizioaren edukiaren babesa bermatzeaz gain, edukiak makinak uler dezakeen formatuan erreformatzea du helburu, hau da, testu-geruza duen faksimile digitala eskaintzeaz gain, hiztegiaren makroegiturari (lemategiari) eta mikroegiturari (sarreraren barruko

antolamenduari) dagozkion anotazioak gehitzea, makinak testu-zatia bakoitza item lexikografiko zehatz gisa interpretatu ahal izateko. Bestela esanda, testu-geruza hutsari anotazio semantikoak dituen metadatu-geruza bat gehitzen zaio. Errekurtsoaren barruko bilaketa aurreratuak ahalbidetzeko eta baliabide barruko nahiz kanporako lotura esplizituak esleitzeko baldintza teknikoak betetzen dira horrela. Nazioartean, hainbat hiztegi klasikotako edukiak horrela prozesatu izan dira, eta egun *machine readable* formatuan ditugu eskura.¹²

Hiztegi historikoen kasu askotan, Larramendi (1745) euskarazko klasiko ezagunaren kasuan, esaterako, digitalizazioa faksimile-irudiak ekoiztean gelditu egin da, eta, ondorioz, pixel-irudi haiei begira dagoen erabiltzailek ikus ditzakeen testu-edukiak ezin dira ordenagailuz lagunduta aurkitu ezta bestela prozesatu ere. OCR (karaktereen ezagutza optikoa) izango litzateke testu-edukiak azaleratzeko beharrezkoa den lan-urratsa. Hiztegi berriago gehien kasuan, testu-geruza duen PDF formatuko argitalpena edota argitalpenaren oinarria izan den testu-fitxategia dugu, eta zenbait kasutan, are gehiago, mikroegitura lexikografikoa islatzen duen markaketa semantikoz osatua da hiztegiaren testua; hala nola, Euskaltzaindiaren 2010ko Hiztegi Batuaren XML bertsioa.¹³ Horregatik, eta 2 irudian azaltzen dugunez, hiztegi bat datu-multzo lexikal egituratu bihurtzeko lan-fluxuak abiapuntu ezberdina izan ahal du («start»), jatorrizko formatuaren arabera.

Hiztegiaren testu-geruza metadatu lexikografikoez osatzeko, ordenagailuz lagundutako metodologiak ere urrats garrantzitsuak eman dituela esan daiteke. «Lema-ikurra», «adiera», «definizioa» edo «itzulpel-ordaina» bezalako mikroegitura-elementuak atzeman eta markatzeko, ikasketa automatikoko algoritmoak darabiltzan tresna batekin esperimentuak egin ditugu, lizentzia libreko OEH-ren PDF bertsioa adibide hartuta.¹⁴ Zenbait orrialdetan markaketa eskuz egin eta tresnak entrenatzeko erabili ondoren, hiztegi-sarrera bakoitzaren hasiera eta amaiera markatzeko gai izan gara hiztegi osoan zehar, sarrera oso laburra izan edo sarrera batek orrialde bat baino gehiago hartzen badu ere. Ikasketa Automatikoko algoritmoak malgutasuna erakutsi du eginkizun horretan, erregeletan oinarritutako sistema batek nekez izango lukeena. Sarreraren hasiera eta amaiera zehazturik, sarrera bakoitzaren barrukoak zatitu eta etiketatzea izango litzateke hurrengo lan-urratsa, handitik txikira aurreratuz («cascaded approach», ik. (Khemakhem, Foppiano & Romary, 2017)). Tresna entrenatzeko datuak eskuz prestatu behar dira, zatiketa automatikoaren emaitzen (eskuzko) ebaluazioak doitasun maila egokia lortu arte.

Edizio digitalen elkarreragingarritasunaren bila, formatu-estandarra sortzeko ahaleginak ere bultzatu dira. Hiztegien barruko antolamendua errepresentatu eta, loturak ezartzeko asmotan, estandar baterantz moldatzeko ekimenen artean, TEI-XML ingurukoak aipatu behar ditugu,¹⁵ hiztegi digitalen egiturak eta markaketa-lengoiak berdintzea helburu dutenak. Hiztegiak eta beste-lako baliabide lexiko-semantikoak elkartzeko, berriz, azken urteotan maiz aipaturiko estandarrak sortzen ari dira *Linguistic Linked Open Data* eremua lantzen duen komunitatean, haren baitan dagoen *Ontolex* W3C lan-taldean,¹⁶ bereziki. Web semantikoaren RDF estandarra (Berners-Lee, 2006) erabili eta hiztegien edukiak eredu formal batean jasotzeko ahaleginak dira *Ontolex-core* («lemon») eta *Ontolex-lexicog* ereduak. Ereduok eskaintzen dituzten klaseak elkarrekin lotuta

¹² Digitalizazio lan-fluxuen inguruan, ikus, adibidez, <https://digilex.hypotheses.org/> blogean baturiko txostenak. Oraingoz lema mailan baino elkarrekin loturarik ez duten hiztegi historikoen bilduma bat ikus <http://woerterbuchnetz.de> helbidean.

¹³ <http://www.euskaltzaindia.eus/dok/eaeb/hiztegiabatua/> helbidean eskuragarri.

¹⁴ Ikus <https://digilex.hypotheses.org/250>

¹⁵ Ikus <https://github.com/DARIAH-ERIC/lexicalresources>

¹⁶ Ikus <https://www.w3.org/community/ontolex/>

daude, zehazki definituriko propietateen bitartez, eta haien arabera sailkaturiko baliabide lexikalen edukiak aurkigarriak izango dira mikroegitura-elementuak islatzen dituzten klaseak eta haiekin lotutako propietateak zehazturik. Horretaz gain, *Ontolex-core* ereduko klaseek *Wikidata* baliabidearen zenbait klaserekiko loturak eskaintzen dituzte.¹⁷ *Wikidata* baliabidearen barruan, berriz, milioika kontzeptu daude bilduta, hiztegietakoa adierekin lotu daitezkeenak. Kontzeptuak biltzen dituen *Wikipedia* ez bezala, *Wikidata* datu lexikalak ere batzen hasi da. Osagai horiekin, *Wordnet* bezalako egitura bat eraiki daiteke *Wikidata*-n bertan, eta *Wikidata* bera izan liteke, beraz, Hiztegi Matrixaren ardatz eta kokapena. Nabarmentzekoa da Euskararen pozisionamendua *Wikidata*-ko datu lexikalei dagokienean, Euskal Wikilarien Elkartearen eta Elhuyar fundazioaren elkarlanari esker.¹⁸

Adiera mailan lotutako Hiztegi Matrix delakoa etorkizuneko erronka bada ere, laburbiltzeko, bi baldintza nagusi bete behar dira Hiztegi Matrix-aren parte izan daitezkeen datu-multzoak sortzeko; lehenik, datu lexikalak egituratuak izan behar dute, XML edo RDF bezalako ereduak erabiltza. Hiztegi-sarrera baten barruan, gutxienez honako mikroegitura-elementu hauek markaketa eraman behar dute, hiztegia sareratu ahal izateko: lema-ikurra, kategoria eta adiera.¹⁹ Bigarren baldintza litzateke loturak ezartzea: Alde batetik, lema-ikur hutsaz haratago, kategoria duen lema bakoitzari (*lempos entity* bakoitza) eta, bestetik, adiera bakoitzari identifikatzailea esleitzea, identifikatzaile hori propietate sorta batez kanpoko baliabideetako beste identifikatzaile batzuekin lotu ahal izateko. Lema mailako loturak ezartzeko bide nahiko zuzena dagoen bitartean, lema ikurra bera horretarako baliabideak daitezkelako, sense linking deritzona askoz ere eginkizun konplikatuagoa da. Joera berriek hemen ere ikasketa automatikoa oinarritutako teknikak erabiltzen dituzte, eskuzko lana errazteko asmotan (ikus eztabaida Sauri, Mahon, Russo & Bitinis, 2019 lanean). Erabateko doitasunari hurbiltzeko, alabaina, hiztegi-gilearen eskuzko lanak ezinbestekoa izaten jarraitzen du, eta algoritmoen garapenean ez ezik, eskuzko lanean ere inbertitzen jarraitzen dute era masiboan arloko egungo proiektuek. Polonierazko *Wordnet* baliabidea hedatu eta haren kontzeptuak ingelesezko *Wordnet*-ekoekin lotzeko, esaterako, dozenaka hiztegi-gilek dihardute polonieraren inoizko baliabide lexiko-semantiko handiena sortu eta mantentzeko ahaleginean (Maziarz, Piasecki, Rudnicka, Szpakowicz & Kedzia, 2016).

4. Bestelako baliabide lexikalak

Hiztegietan zehar ez ezik, lema eta adierak dituen bestelako baliabide lexikaletan zehar ere loturak ezartzea ezinbesteko baldintza da hainbat aplikazio burutzeko. LNP alorrean kokatu ohi diren baliabide lexiko-semantikoetatik bat aipatu dugu goian: *Wordnet* kontzeptu-sarea (Miller, Beckwith, Fellbaum, Gross & Miller, 1990). Horretaz gain, testu-corpusetatik erauzitako maiztasun lemategiak eta hizkuntzalaritza konputazionalako lexikoiak ere lotu ditugu lema-ikur mailan elkarrekin eta hiztegietakoa lemategiekin. Hiztegi elebidun berri bat sortzeko, euskarazko oinarritzko lemategia sortzea baitzen gure asmoa (EusLemStd, Lindemann & San Vicente, 2015).

¹⁷ Ikus https://www.wikidata.org/wiki/Wikidata:Lexicographical_data eta https://www.mediawiki.org/wiki/Extension:WikibaseLexeme/RDF_mapping

¹⁸ Ikus euskarazko lexemen kopurua *Wikidatan* <https://tools.wmflabs.org/ordia/language/> helbidean, eta adibide bat hemen: <https://www.wikidata.org/wiki/Lexeme:L48738>. *Wikidata*-ko euskarazko lema guztiak honako SPARQL-query honen bitartez ikus daitezke: <https://w.wiki/Byt>

¹⁹ Mikroegitura-elementu gehiago markatuz gero, haien edukia ere aurkigarri eta loturak ezartzeko gai bihurtzen dira, hala nola, aipua, *Geintegreerde Taalbank* gertatzen den moduan.

Horretarako, gaurko corpus handienetan, hau da, ETC Egungo Testuen Corpus (Sarasola, Landa & Salaburu, 2013) eta Elhuyar 200M-Webcorpusean (Elh200, Leturia, 2014), eta, aldi berean, hiztegi nagusietariko batean²⁰ agerpena duten lema-ikurrak zehaztu ditugu. Horrela, hiru baldintza betetzen zituzten lema-ikurren zerrenda lortu dugu. Lehenik, lema-ikur historikoak baztertu ditugu, hau da, gaurko erabilera islatzen duten corpusetako batean agerpen-atalasea²¹ gainditzen ez dutenak. Bigarrenik, hiztegi nagusietan inon jaso ez direnak baztertu ditugu, corpusetan agertzen den guztiaren artean jasotzeko mereziko luketeen neologismoak iragaztea ez baita hiztegiaren elebidunaren zeregina, Lexikoaren Behatokia (Artola, Ezeiza, Gurrutxaga, Sagarna & Urkia, 2017) bezalako ekimen berezituaren eginkizun erraldoia baizik. Eta hirugarrenik, hiztegien eta corpusen intersektzioa islatzen duen lematagia corpusetako maiztasun-datuekin hornitu dugu, horrela EusLemStd lematagia abiapuntua duen hiztegi berrian lehen maiztasunari buruz argibideak eman ahal izateko.

```
<homograph homograph="aditu" corpus_counts="42042">
  <ADI lemma="aditu" pos="ADI SIN" corpus_counts="18989">
    <sense synset="30-00588888-v" equivs="understand"/>
    <sense synset="30-02169702-v" equivs="hear"/>
    <sense synset="30-02571901-v" equivs="heed mind listen"/>
  </ADI>
  <IZE lemma="aditu" pos="IZE ARR" corpus_counts="13945">
    <sense synset="30-09617867-n" equivs="expert"/>
    <sense synset="30-10557854-n" equivs="scholar scholarly_person bookman"/>
  </IZE>
  <ADJ lemma="aditu" pos="ADJ ARR" corpus_counts="5486">
    <sense synset="30-02226162-a" equivs="adept expert skillful"/>
  </ADJ>
</homograph>
```

3. IRUDIA

‘aditu’ lema EDBL eta EusWN datuak eta corpus maiztasunak

Aipatutako lanean euskal Wordnet (EusWN, Pociello, 2007) eta EDBL (Aldezabal *et al.*, 2001)²² ere aintzat hartu genituen, corpus eta hiztegiekiko intersektzioak definitzeko. Hutsetik sortzekoa litzatekeen hiztegi elebidun berria eduki gehiagoz osatzeko ere erabili genituen bi baliabide horiek (Lindemann & San Vicente, 2016). EusTagger tresnak (Aduriz *et al.*, 1996) corpuseko formei lema eta kategoria esleitzen dizkie, EDBL-eko datuetan oinarritua. EusLemStd zerrendako lema, hau da, Elh200 corpusean gutxienez 20 aldiz eta, aldi berean, hiztegi nagusietako batean agertzen direnak, EusTagger-rek asoziatutako hitz kategoriez anotatu genituen, bigarren pauso batean EusWN-eko adierak ere gehitzeko. EusLemStd sarreraren heren bat estali ahal izan dugu horrela gutxienez kategoria eta adiera banarekin. 3 irudian erakusten dugu emaitzen adibide bat, aditu lema, lehenengo maila batean hitz kategoriaren arabera antolatuta, haren azpian *Open Multilingual Wordnet* (Bond & Foster, 2013) bitartez lortutako ingelesezko itzulpen-ordainak eta Elh200 corpuseko agerpen-kopuruak adierazten ditugula.

²⁰ Orotariko Euskal Hiztegia (Mitzelena & Sarasola, 1988), Hiztegi Batua (Euskaltzaindia, 2010), Euskal Hiztegia (Sarasola 1996) eta Elhuyar Euskara-Gaztelania hiztegiaren euskarazko lematagia (Azkarate, Kintana & Mendiguren, 2006) dira eskuragarri izan ditugun lematagiak.

²¹ Hogeia agerpenetan finkatu dugu atalasea, corpus hizkuntzalaritzako usadio bati jarraituz.

²² Euskararen datu base lexikala (EDBL) corpusetan metadatu morfosintaktikoak gehitzeko eta zuzentzaile ortografikoetan erabiltzen da, besteak beste.

Horrela lortutako hiztegi-zirriborroa hiztegi elebidunaren abiapuntua balitz, eta eduki guztiak irudian dugun bezalako markaketaz hornituko balira, zirriborroa erauzteko baliatu zituen iturriak aberasteko lanean jardungo luke hiztegi-gileak aldi berean, eskuzko lanaren balorea bikoiztuz, el-karreragingarritasun iraunkorreko *bootstrappingloop* deitu daitekeen lan-fluxu batean. Corpus maiztasun atalasea gainditu eta EusTagger-rek kategoria esleitu arren EusWN adierarik ez duten EusLemStd zerrendako lemak izango lirateke landu beharrekoak. Eta lema hauek hiztegi-gintza elebidunean lantzeak haien adierak zehaztea esan nahi duenez, EusWN hedatzeko ekarpena d-akar. Bestetik, lema batek dituen adiera guztiak EusWN baliabidean islatuta dituen edo ez galdetu beharko da; horretarako, hiztegi-gietan duen adiera-kopuruari begiratzea izan daiteke arazoari hurbil-tzeko bide bat (Lindemann & San Vicente, 2016).

Euskarazko hiztegi-ek ez dutela baliabide bakoitzetik kanpora bidaltzen duen lotura expli-ziturik esan dugu. Corpusetako adibideak hiztegi sarrerei edo adieretako bakoitzari lotuta ager-tzen badira, hiztegi-giaren parte gisa jaso ditu hiztegi-gileak, Egungo Euskararen Hiztegi-gian (EEH, Sarasola, 2008), esaterako, gertatzen den moduan. Hiztegi-gitik kanpora lotura espliziturik ezean, erabiltzaileak baliabide ezberdinen artean salto egiteko moduak badaude, hala ere. Elhuyar hiz-tegi elebiduneko erabiltzaile-interfazeak²³ eskaintzen dituen aukerak adibide moduan azalduko ditugu.

Batetik, hiztegi horretako sarrera elebidunetan, corpus paraleloan bilaketa egiteko aukera ematen da. Bilaketa egiteko bi hizkuntzetako lemak erabiltzen dira, eta horrela bilaketa bi ordai-nak agertzen diren esaldi-bikoteetara mugatu. Adieren desanbigua-zioa, beraz, ordainaren bidez egiten da, bilaketa egiten den momentuan bertan. Jokabide horren abantaila begi-bistakoa da: cor-pusetako adibideak ez dira euskarazko lema-aren beste adierei dagozkienekin nahasten. Aldi berean, horrek adibideak hiztegi-gian agertzen diren ordainetara mugatzen ditu, bestelako ordainak bilatzeko corpusera zuzenean jo beharko genukeelarik.

Bestetik, Elhuyarrek bere hiztegi tekniko eta terminologikoak elkar-lotuta ditu. Horri esker, erabiltzaileak hiztegi-sarrera bat kontsultatzen duenean aukera du sarrera hori aipatutako kanpo baliabideetan kontsultatzeko. Horren muga, berriz ere, erakunde ezberdinek garatutako baliabi-deak elkarlotzeko aukeran datza. Hori bilaketa bidez ahalbidetu liteke (*Euskalbar* tresnak egiten duena), baina ez hiztegi-sarrera desanbiguatuen mailan.

5. Sarearen aplikazioak

3. atalean aurkeztutako lan-fluxuaren emaitza datu-multzo lexikalen sorta da, eta haien elkar-keta lema-ikurren edota adieren bitartez baliabide lexikalen sarea, edo, neerlanderazko izendapen-ereduari jarraituz, hizkuntza-banku integratua. Horrelako sarearen aplikazioen artean, erabiltzai-le-interfaze baten bitartez bilaketa aurreratuak eskaintzea aipa daiteke lehendabizi, *Geïntegreerde Taalbank*-ek duen interfaze publikoak ahalbidetzen duen bezala. Galizierazko baliabide lexikalen sareak *Recursos Integrados da Lingua Galega* du izena,²⁴ idatzizko lekukotza eta lema-ikurraren araberrako bilaketak ahalbidetzen ditu hainbat baliabideetan aldi berean, hiztegi-gietan, GalNet *Word-net*-ean nahiz testu-corpusetan, lema-aren kategoria gramatikala zehazturik.

²³ <http://hiztegiak.elhuyar.eus> helbidean eskuragarri.

²⁴ Ikus <http://sli.uvigo.es/RILG>

Bigarrenik, sare bihurtutako datu lexikalen bilduma horrelako beste sare batzuekin elkar daiteke, egiturak TEI-XML edo *Ontolex*-RDF bezalako nazioarteko estandar bati jarraitzen badio. Lema mailan elkarturik dauden bilduma batzuk goian aipatu ditugu. Bestetik, *Wordnet* sare semantikoaren ereduari jarraitzen dioten bildumak aipagarriak dira, azken urteotan maiz aipaturiko *BabelNet* ekimena tartean (Navigli & Ponzetto, 2010),²⁵ *Open Multilingual Wordnet* eta *Wikipedia* baliabideen edukiak, besteak beste, kontzeptu mailan batzen dituen, doitasun handiko algoritmo baten bitartez.²⁶ *BabelNet*ek bere edukia aberasten jarraituko du, eta hiztegi historikoen edukia *BabelNet*-eko kontzeptuekin lotzea da, hain zuzen ere, *Elexis* ekimenean jarri duten erronketako bat (Declerck, McCrae, Navigli, Zaytseva & Wissik, 2018).

Hirugarrenik, baliabide lexikalen sarea filologia arloko ikerketa kuantitatiboen iturri gisa erabil daiteke, edota hurbiletik aztertzeke laginak sortzeko. Euskarazko zenbait datu lexikal elkartu ditugu, baliabide ezberdinetatik erauzketak eginda. Jarraian, elkarketa horren aplikazioetarako baliabide emango dugu.

5.1. *Larramendiren lema berrienerabilera gaur*²⁷

Goian aipatu dugun euskarazko hiztegi klasikoa, Larramendirena, alegia, ez dugu oraindik edizio digitalean. Alabaina, digitalizazioa burutu ezean, bidezidorra dugu beste hiztegi baten edukia erabiliz. Izan ere, Ibon Sarasolaren hiztegi elebazarretan ‘*1745’ marka darama lema, Larramendiren hiztegia lehenengo agerpen dokumentatua baldin bada. Bestetik, egungo lema-ikur baten azpian jasotzen du Sarasolak Larramendik erabilitako forma (OEH-n gertatzen den bezalaxe). Larramendiren ‘beguioordeac’ (s.v. ‘anteojos’), adibidez, islatzeko, grafiari eta, pluraleko forma denez, morfologiari dagokionez gaurko lematizazio estandarrari erantzuten ez dion forma beharrean, ‘begi-orde’ bezala zerrendatzen dute Sarasolak eta OEH-k, eta begioordeak da Larramendiren hiztegiko aipuan OEH-n ematen duen forma. Egungo euskararen lematizaziorako egokitzapena ere badugu, beraz, Sarasola/OEH hiztegieta. OEH-tik lemaz bestelako eremuetara murriztutako bilaketak egun ezin ditugu ordenagailuz lagunduta egin. Alabaina, OEH-ko begi-orde adibideak prototipikoki erakusten du lematizazio historikoa interpretatzeko jokabidea (‘beguioordeac’>‘begioordeak’>‘begi-orde’), eta hain zuzen Sarasolak eta OEH lantzen duen taldeak egindako egokitzapen horregatik izango gara gai Larramendiren lema gaurko euskararen beste baliabide batzuetan bilatzeko.

Euskal Hiztegiaren 1996ko edizioa XML bihurtu dutenez (Arriola *et al.*, 2003), hiztegi osoan zehar aurki eta erauz ditzakegu edukia, mikroegitura-elementua zehazturik, datazioarena, esaterako. Larramendiren lanean lehendabiziko agerpena duten lema 3.420 ‘*1745’ marka daramaten horiek direla onartuz gero, eta lema haien zerrenda egungo corpusetatik erauzitako maiztasun-zerrendekin erkatzen badugu, Larramendiren lema berrien arrakastaz zerbait esan dezakegu, eskuzko inolako bilaketarik egin gabe.

²⁵ Ikus <http://babelnet.org>

²⁶ *BabelNet* baliabidetik euskaraz-ingelesazko hiztegi elebiduna erauzteko gure esperimentu eta ebaluaketaz ikus (Lindemann & Kliche, 2017). *EusLemStd* zerrendaren estaldura nabarmen handiagoa eskaini du *BabelNet*-ek (%40), *WordNet* soilik erabiltzearen aldean (%31), doitasuna maila berean mantentzen duen bitartean (%85).

²⁷ Azkarate & Lindemann (2018) lanean antzeko ariketa bat aurkeztu dugu, Sabindarren neologismoen gaurko erabileraren inguruan.

1 eranskinean ‘*1745’ lemen maiztasun-zerrenda ematen dugu, ETC eta Elh200 corpusetik erauzitako maiztasun-zerrendan dituen tokiak adierazita. Gaurko euskararen 750 gehien erabilitako lemetatik 40 Larramendik lehen aldiz jasotakoak direla izango litzateke ondorioa, ereduazko testuetan nahiz internetean topaturiko euskarazko edukietan. Oro har, 3.420 lema-ikur haien erdiak (1.742) Elh200 corpusean hogeia agerpen edo gehiago dutenez, gaurko erabilera frogatutzat eman daiteke, beste erdiaren %44 (1678tik 733) hogeia aldiz baino gutxiago agertzen den bitartean; gainontzeko 945 lema ez dira behin ere agertzen.

5.2. *Twitter corpusaren azterketa*

Azkenik, sare sozialetara begira jarri nahi izan dugu. Horretarako Twitterretik erauzitako 40 Miloi hitzeko corpusa erabili dugu. Corpus orokorra da, 2017 urtetik 2019ko iraila bitartean bildua, eta hizkuntza automatikoki detektatuta. Twitterren erabiltzen den hizkuntza aztertzeak berez erronkak ditu. Izan ere, Twitterren erabiltzen den hizkuntza askotan kaleko ahozko hizketatik hurbilago dago idatzizko hizkuntzatik baino. Ohiko testuetan topatzen ez ditugu zenbait fenomeno tratatu behar dira. Adibidez, gaztelania eta euskara nahasten dituzten esaldiak (Code Switching izenez ezagutzen dena, adib. ‘Ta atzo kede fatal!!!’, ‘Ingleseko azterketa de puto kulo itezunen...’), enfasia adierazteko hitzak maiuskulaz idaztea (‘A ze pelikula TXARRA ikusi dugun gaur’) ala bokalak errepikatzea (adib., ‘ooona!’), edo lekua aurrezteko hitzak moztea (adib., ‘msdz’ → ‘mesedez’, ‘mlskr’ → ‘milesker’, ‘pxkt’ → ‘pixka bat’). Fenomeno horiek denek noski prozesamendu automatikoa zailtzen dute, harik eta berariazko aurreprozesatze bat eskatzen dutelarik (Alegria *et al.*, 2015). Aurreprozesatze hori bereziki da beharrezkoa hiztegi gintza arloko aplikazioetarako, edo lemak beste baliabideetako loturez hornitzeko, forma ez estandarren lemak identifikatu nahi baditugu.

Sare sozialetako datuek erronkak planteatzen badituzte ere, abantailak ere badituzte. Hizkuntzaren aldaketak sare sozialetan idatzizko lekukotza izaten du, beste inon baino bizkorrago. Eta lekukotza horiek erauz ditzakegu, era masiboan. Termino berrien agerpena, edo gazte hizkera bezalako erregistroen azterketa egiteko informazio iturri egokia izan daiteke, beraz (Nguyen, Gravel, Trieschnigg & Meder, 2013).

Iturri honetatik eratorri dezakegun informazioa baliatzeko adibide gisa, Twitter corpusean agertzen diren lemen azterketa egin dugu. Corpusaren lematizazioa beste corpusekin egin bezala Eustagger tresnarekin burutu dugu. Goian aipatutako aurreprozesatzerik ez dugu egin kasu honetan, gure helburua beste corpusekin konparatzea izan delako. Hortaz interes berezia genuen Sare sozialetan ematen diren fenomeno horiek azaleratzen. 2 eranskinean, hiru corpusetan agerpena duten lema-ikurren lagina eskaintzen dugu. Lema-ikurrek hiru corpusetan dituzten maiztasun erlatiboak (rfreq, ikus Lindemann & San Vicente (2015)) konparatzen ditugu, terminologia-erauzketan erabiltzen den weirdness ratio neurria erabiliz (Ahmad, Davies, Fulford & Rogers, 1994; ik. Schäfer, Rösiger, Heid & Dorna, 2015). Weirdness baloreak adierazten du ikur baten maiztasuna corpus batean, erreferentzia-corpus batean duenarekin alderaturik. Taulan zerrendatutako lema-ikurrek corpus zehatz batean askoz ere agerpen gehiago dute dagokion erreferentzia-corpusean baino, eta, hortaz, corpus zehatz horren osaeraren ikuspegi berezia eskaintzen dute. Ez dira corpus batean maizen agertzen direnak, erreferentzia-corpuseko maiztasunetik urrunenak gelditzen direnak baizik. Zerrendako lehenengo posizioei erreparatuta, Internetetik erauzitako edukiak biltzen dituzten Elh200 corpusak eta Twitter corpusak antzekotasuna erakusten dute, es-

kuz hautaturiko ereduzko testuez osatutako ETC corpusaren aldean. Bestalde, Twitter corpusak, Elh200 erreferentzia harturik, bere berezitasunak erakusten ditu, hemengo lehenengo tokietan ere erakuskari esanguratsuak agertzen baitira. Hemen aurkezten duguna oso oinarritzkoa bada ere, maiztasunak erkatzeko ariketa hau ezingo genuke burutu iturri ditugun baliabideak elkartu ahal izan ezean.

6. Ondorioak

Hiztegi historikoak eta bestelako baliabide lexikalak saretzea izan dugu hizpide artikulu honetan. Lan-urratsak ikusi ditugu, karaktereen ezagutza eta testu-edukien markaketa semantikotik datu-multzo lexikal egituratuz osatutako bilduma sortzeraino. Bestetik, horrelako bilduma sare semantikoekin erlazionatzeko baldintzak ere aztertu ditugu, horien artean, datu-ereduen estandarrei erreparatzeko premia.

Euskarazko hiztegi historikoei dagokienez, hastapen batzuk baino ezin izan ditugu gauzatu egun arte. Lemen aldaera historikoak egungo estandarrentz moldatzea beharrezkoa dela ikusi dugu lema-ikurretan oinarritutako baliabide-sareak eraikitzeke. Beste puntu batean ere agertu zaigu premia bera: Sare sozialetan erabiltzen diren formak ere estandar diren edo lirakekeen formekin lotzea. Lan-urrats horrek, aurrean ditugun aldaerak historikoak ala berri-berriak izanda, irizpide filologikoak eskatzen ditu. Euskarazko hiztegitantza historikoan lematizazioaz sortutako ereduak aztertu eta erabilgarri bihurtzea, corpus hizkuntzalaritzako tresnek aplikatu dituzten, gerorako eginkizun garrantzitsua izan daitekeela aipatu nahi dugu, eta OEH-n lema historikoen kodifikatze lexikografikoaz egindako eskuzko lan erraldoia banan-banan erauztea horretarako bide egokia izan litekeela.

Datozen urteetan ildo horretan lanean jarraitzeko beharra eta aukera ikusten dugu. Elhuyar Fundazioa hasia da dagoeneko bide horretan. Batetik, fundazioak berak garatutako baliabide lexikalen inguruko loturak eskaintzen dizkie erabiltzaileei, corpusetatik eratorritako adibideak, edo hiztegi terminologikoetan egon daitezkeen sarreretara loturak. Horiek «barne loturak» direla esan dezakegu, erakunde berak sortutako baliabideen artekoak. Baina, Euskal Wikilarien Elkartarekin elkarlanean, Elhuyar hiztegi-tako datu lexikalak Wikidata ekimenean txertatzen hasiak dira. Euskarazko bestelako baliabideekin zein beste hizkuntzetako lexikalizazioekin loturak sortzeko pauso garrantzitsu bat eman da honezkero, aurrera begira, euskarazko hiztegien erabiltzaileari ere onurak ekarriko dizkiona, edozein bilaketa egitean, euskaraz ez ezik beste hainbat hizkuntzako informazioa ere eskura izango duelako.

H2020 *Elexis* ikerketa-azpiegitura lagungarria izango zaigu, lan-fluxuak, datu-ereduak eta teknologiak Europako hizkuntza-komunitateetan zehar eztabaidatu, garatu eta zabaltzeko asmoa duena. Europako beste hainbat erakundek bezala, EHUK ere *Elexis* proiektuan hartzen du parte observer gisa, konsortzioari euskarazko datu lexikalak eskuragarri jarritz, azaldutako lan-fluxuaren urrats bakoitzean esperimenduak egin daitezela, eta esperimendu haien emaitzak ebaluatuz, tresna automatikoen garatzaileek ondorioak ateratu dituzten. Trukean, garatutako teknologien early adopters izango gara. Tresnen garatzaileen eta datuen emaitzen, informatikoen eta filologoaren artean dugun nazioarteko lankidetzak hori Humanitate Digitala arloa gorpuzten duten haietarikoa da, mundu biek irabazteko moduko lankidetzak, zalantzarik gabe.²⁸

²⁸ David Lindemanek Eusko Jaurlaritzaren laguntza eskertzen du (IT1169-19).

Bibliografia

- Aduriz, I., Aldezabal, I., Alegria, I., Artola, X., Ezeiza, N., & Urizar, R. (1996). EUSLEM: A lemmatiser/tagger for Basque. *Proceedings of EURALEX 1996*, 17-26. Sarean: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.44.9004&rep=rep1&type=pdf>
- Ahmad, K., Davies, A., Fulford, H. & Rogers, M. (1994). What is a term?: The semi-automatic extraction of terms from text. In M. Snell-Hornby, F. Pöschhacker & K. Kaindl (Arg.), *Benjamins Translation Library* (Libk. 2, or. 267). <https://doi.org/10.1075/btl.2.33ahm>
- Aldezabal, I., Ansa, O., Arrieta, B., Artola, X., Ezeiza, A., Hernandez, G. & Lersundi, M. (2001). EDBL: A General Lexical Basis for the Automatic Processing of Basque. *Proceedings of the IRCS Workshop on linguistic databases*, Philadelphia. Sarean: <http://artxiker.ccsd.cnrs.fr/docs/00/08/11/54/PDF/2001-IRCS.pdf>
- Alegria, I., Aranberri, N., Comas, P. R., Fresno, V., Gamallo, P., Padró, L., San Vicente, I., Turmo, J., Zubiaga, A. (2015). TweetNorm: A Benchmark for Lexical Normalization of Spanish Tweets. *Language Resources and Evaluation*, 49(4), 883-905. Sarean: <https://doi.org/10.1007/s10579-015-9315-6>
- Artola, X., Ezeiza, N., Gurrutxaga, A., Sagarna, A. & Urkia, M. (2017). Lexikoaren Behatokia: Leiho bat XXI. mendeko hedabideetako euskarari. *Senez: itzulpen aldizkaria*, (48), 16. Sarean: <https://dialnet.unirioja.es/servlet/articulo?codigo=6206911>
- Azkarate, M., Kintana, X. & Mendiguren, X. (Arg.). (2006). Elhuyar hiztegia: Euskara-gaztelania, castellano-vasco (3. arg.). Usurbil Gipuzkoa: Elhuyar Edizioak. Sareko interfazea: <http://hiztegiak.elhuyar.org>
- Azkarate, M., & Lindemann, D. (2018). Basque Lexicography and Purism. *International Journal of Lexicography*, 31(2), 132-150. Sarean: <https://doi.org/10.1093/ijl/icy003>
- Bond, F. & Foster, R. (2013). Linking and Extending an Open Multilingual Wordnet. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 1352-1362. Sarean: <http://anthology.aclweb.org/P/P13/P13-1133.pdf>
- Camacho-Collados, J. & Pilehvar, M. T. (2018). From word to sense embeddings: A survey on vector representations of meaning. *Journal of Artificial Intelligence Research*, 63, 743-788. Sarean: <https://jaair.org/index.php/jair/article/view/11259>
- Declerck, T., McCrae, J., Navigli, R., Zaytseva, K. & Wissik, T. (2018). ELEXIS - European Lexicographic Infrastructure: Contributions to and from the Linguistic Linked Open Data. In I. Kernerman & S. Krek (Arg.), *Proceedings of the LREC 2018 Workshop «Globalex 2018 – Lexicography & WordNets»* (or. 17-22). Sarean: <https://globalex.link/globalex2018/proceedings/>
- Euskaltzaindia. (2010). Hiztegi batua (3.). Sarean: <http://www.euskaltzaindia.net/hiztegiatua/>
- Khemakhem, M., Foppiano, L. & Romary, L. (2017). Automatic Extraction of TEI Structures in Digitized Lexical Resources using Conditional Random Fields. In I. Kosem, C. Tiberius, M. Jakubíček, J. Kallas, S. Krek & V. Baisa (Arg.), *Electronic lexicography in the 21st century: Lexicography from scratch. Proceedings of eLex 2017*. Sarean: <https://elex.link/elex2017/proceedings-download/>
- Larramendi, M. (1745). *Diccionario trilingüe castellano, bascuence y latin dedicado a la M.N. y M.L. provincia de Guipuzcoa*. Sarean: <http://www.kmliburutegia.eus/Record/203133> eta <http://www.kmliburutegia.eus/Record/26577>
- Leturia, I. (2014). *The Web as a Corpus of Basque* (PhD Thesis). UPV-EHU Lengoia eta Sistema Informatikoak Saila, Donostia. Sarean: <https://www.elhuyar.eus/media/content/files/ILeturia-The%20Web%20as%20a%20Corpus%20of%20Basque.pdf>
- Lindemann, D. & Kliche, F. (2017). Bilingual Dictionary Drafting: Bootstrapping WordNet and BabelNet. In I. Kosem, C. Tiberius, M. Jakubíček, J. Kallas, S. Krek & V. Baisa (Arg.), *Electronic lexicography in the 21st century: Lexicography from scratch. Proceedings of eLex 2017* (or. 23-42). Sarean: <https://elex.link/elex2017/wp-content/uploads/2017/09/paper02.pdf>

- Lindemann, D. & San Vicente, I. (2015). Euskarazko maiztasun lematagia gaurko teknologien ikuspuntutik. In B. Fernández Fernández & P. Salaburu Etxeberria (Arg.), *Ibon Sarasola, gorazarre. Homenatge, homenaje* (or. 441-456). Sarean: <http://www.ehu.es/ehg/sarasola/liburua/SarasolaGorazarre36.pdf>
- Lindemann, D. & San Vicente, I. (2016). Bilingual Dictionary Drafting: Connecting Basque word senses to multilingual equivalents. In T. Margalitadze & G. Meladze (Arg.), Proceedings of the 17th EURALEX International Congress: Lexicography and Linguistic Diversity (or. 898-905). Sarean: <http://euralex.org/publications/bilingual-dictionary-drafting-connecting-basque-word-senses-to-multilingual-equivalents/>
- Maziarz, M., Piasecki, M., Rudnicka, E., Szpakowicz, S. & Kedzia, P. (2016). plWordNet 3.0—A Comprehensive Lexical-Semantic Resource. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics, 2259-2268. Sarean: <https://www.aclweb.org/anthology/C/C16/C16-1213.pdf>
- Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D. & Miller, K. J. (1990). Introduction to WordNet: An On-line Lexical Database. *International Journal of Lexicography*, 3(4), 235-244. <https://doi.org/10.1093/ijl/3.4.235>
- Mitxelena, K. & Sarasola, I. (1988). Diccionario general vasco—Orotariko euskal hiztegia. Euskaltzaindia; Editorial Desclee de Brouwer. Sarean: <https://www.euskaltzaindia.eus/hizkuntza-baliabideak/online-hiztegiak>
- Navigli, R. & Ponzetto, S. P. (2010). BabelNet: Building a Very Large Multilingual Semantic Network. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, 216-225. Sarean: <http://dl.acm.org/citation.cfm?id=1858681.1858704>
- Nguyen, D., Gravel, R., Trieschnigg, D., & Meder, T. (2013). «How old do you think I am?» A study of language and age in Twitter. Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media. -(E)n aurkeztua Cambridge MA. Sarean: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewPaper/5984>
- Pedersen, B. S., Nimb, S., Olsen, S. & Sørensen, N. H. (2018). Combining Dictionaries, Wordnets and other Lexical Resources – Advantages and Challenges. In I. Kernerman & S. Krek (Arg.), *Proceedings of the LREC 2018 Workshop «Globalex 2018 – Lexicography & WordNets»* (or. 101-104). Sarean: <https://globalex.link/globalex2018/proceedings/>
- Pociello, E. (2007). *Euskararen ezagutza-base lexikala: Euskal WordNet* (PhD Thesis). UPV/EHU Euskal Filologia Saila, Donostia. Sarean: <http://ixa.si.ehu.es/node/4117>
- Raganato, A., Camacho-Collados, J. & Navigli, R. (2017). Word sense disambiguation: A unified evaluation framework and empirical comparison. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, 99-110. Sarean: <https://www.aclweb.org/anthology/E17-1010/>
- Sarasola, I. (1996). Euskal Hiztegia. Donostia: Kutxa Gizarte-eta Kultur Fundazioa.
- Sarasola, I. (2008). Egungo Euskararen Hiztegia (EEH). Berreskuratua 2019(e)ko azaroakaren 12a, -(e)tik UPV/EHU website: <https://www.ehu.es/eeh/>
- Sarasola, I., Landa, J. & Salaburu, P. (2013). Egungo Testuen Corpora. Sarean: <http://www.ehu.es/etc/>
- Saurí, R., Mahon, L., Russo, I. & Bitinis, M. (2019). Cross-Dictionary Linking at Sense Level with a Double-Layer Classifier. In M. Eskevich, G. de Melo, C. Fäth, J. P. McCrae, P. Buitelaar, C. Chiarcos, M. Dojchinovski (Arg.), *2nd Conference on Language, Data and Knowledge (LDK 2019)* (or. 20:1-20:16). <https://doi.org/10.4230/OASIS.LDK.2019.20>
- Schäfer, J., Rösiger, I., Heid, U. & Dorna, M. (2015). Evaluating noise reduction strategies for terminology extraction. In T. Poibeau & P. Faber (Arg.), *Proceedings of the 11th Int. Conference on Terminology and Artificial Intelligence (TIA'15)* (or. 10). Granada: Universidad de Granada. Sarean: http://ceur-ws.org/Vol-1495/paper_7.pdf
- Smith, M. N. (2004). Electronic Scholarly Editing. In S. Schreibman, R. Siemens & J. Unsworth (Arg.), *A Companion to Digital Humanities*. Sarean: <http://www.digitalhumanities.org/companion/>

1. eranskina

Sarasolaren (Larramendiren) «*1745» lemen maiztasun-zerrenda, ETC eta Elh200 corpusetik erauzitako maiztasun-zerrendan dituen tokiak adierazita (40 lehenengoak).

Posizioa	Lema	ETC rank	Lema	Elh200 rank
1	talde	39	talde	42
2	erabaki	100	aurkeztu	93
3	egoera	121	ondoren	95
4	azaldu	128	elkarte	108
5	maila	136	erabaki	120
6	ondoren	143	maila	130
7	biztanle	174	egoera	134
8	arazo	183	azaldu	173
9	aurkeztu	207	era	208
10	elkarte	233	arazo	215
11	asmo	241	atal	268
12	era	281	irail	271
13	arrazoi	301	asmo	311
14	saiatu	355	honako	319
15	zuzendari	373	zehir	348
16	sarrera	381	hona	359
17	bukatu	394	sarrera	364
18	ordaindu	423	bultzatu	367
19	bultzatu	455	albiste	379
20	ondorengo	459	ordaindu	395
21	adina	510	martxa	411
22	doan	543	arrazoi	412
23	zehir	545	zuzendari	416
24	atal	549	bukatu	444
25	giro	550	lasai	512
26	egitura	559	saiatu	526
27	baztertu	560	hitzaldi	538
28	lasai	565	hondakin	559
29	antzeko	569	ale	581
30	desagertu	615	egitura	587
31	irail	630	bideratu	617
32	iparralde	641	giro	620
33	garaipen	651	arautu	640
34	zalantza	664	gehitu	649
35	honako	693	ipuin	652
36	aski	700	zalantza	675
37	benetan	716	bizitu	685
38	hainbeste	719	egiazatu	710
39	sasoi	721	neu	712
40	proba	747	antzeko	721

2. eranskina

Twitter corpusean gutxienez hogeitaz aldiz nahiz ETC eta Elhuyar corpus bietan agertzen diren lema-ikurrak, weirdness-ratio neurriaren arabera ordenaturik, erreferentzia-corpus ezberdinak oinarritut hartuta (100 lehenengoak). Hirugarren eta laugarren zutabeetan, Twitter corpuseko maiztasun erlatiboak ETC eta Elh200 corpusetako maiztasunekin, eta bostgarren zutabeetan, berriz, bien batezbestekoarekin duen proportzioaren arabera hurrenkera dugu, hau da hirugarren zutaberako, Twitteren soilik internet osoan baino maiztasun askozaz ere handiagoaz erabiltzen direnen erakuskaria.

Posizioa	Elh200_rfrec / ETC_rfrec	Twitter_rfrec / ETC_rfrec	Twitter_rfrec / Elh200_rfrec	Twitter_rfrec / ETC_Elh200_rfrec
1	batera	batera	ezker	etxeondo
2	arteko	noranzko	mikroskopiko	zapra
3	aurrez	arteko	aspirazio	patatx
4	bertatik	zuzeneko	astintzaile	antifaxismo
5	arriskutsu	arriskutsu	berritsukeria	manugaitz
6	baserritar	bertatik	botagura	torraka
7	zuzeneko	baserritar	burugorri	egibide
8	arrakastatsu	aurrez	indargetu	txantxibiri
9	bigarren	arrakastatsu	saharau	infrasoinu
10	inon	bigarren	amatista	txerren
11	zehazki	inon	tematu	elikagune
12	berrikusi	eguberri	interrogatorio	mozal
13	garapa	urrezko	gedar	salabardo
14	zelan	aspaldiko	librepentsalari	noranzko
15	jatorriz	etxeondo	patroi	mertzero
16	zeharka	zelan	tiratu	eztabai
17	aspaldiko	ospa	artalde	enkantatu
18	bada	orduko	tesla	prospero
19	urrezko	indartsu	frustazio	nabaro
20	bertako	berrikusi	tuneatu	spoiler
21	zai	zai	txeketegi	bermiotar
22	indartsu	fuerte	egozentriko	zerraldatu
23	hirugarren	txiolari	kontari	golatu
24	bakarka	zehazki	fiskalizatu	brutal
25	orduko	koloretsu	pito	azarri
26	eskuko	larriki	bikini	berjabetu
27	berreskuratu	hamargarren	hispanoamerikar	euskarafobo

Posizioa	Elh200_rfrec / ETC_rfrec	Twitter_rfrec / ETC_rfrec	Twitter_rfrec / Elh200_rfrec	Twitter_rfrec / ETC_Elh200_rfrec
28	berreraiki	klaru	pertso	behizain
29	berriki	berreskuratu	hirixka	zibereraso
30	nora	zabalik	tupla	proximo
31	gogoko	eskuko	bengaliar	armagabetu
32	ageriko	bizirik	graduatu	fronton
33	bizirik	nora	marigorringo	basomutil
34	ahozko	gogoko	normalitate	txiotxio
35	nahikoa	bertako	enbaxada	horrible
36	hamargarren	garapa	kultista	bonito
37	hare	hirugarren	arrastiri	alegrías
38	neu	ikusarazi	israeldar	sorki
39	klaru	zeharka	dibisa	jornada
40	gehiégizko	oriotar	domeinatu	txiolari
41	amen	ageriko	arilo	lagunkoitasun
42	gaineko	bada	zorun	playoff
43	onuradun	mago	zabalketa	erretura
44	interesdun	keatu	zakuto	restaurante
45	koloretsu	aitzina	konopial	bidelari
46	misteriotsu	basura	dolore	desakatu
47	zatika	hagoatu	irudikeria	maravilloso
48	delako	berreraiki	atralaka	salio
49	zentsu	gogotsu	hamargarren	estatalizazio
50	berregín	ikustarazi	espainiarzale	bypass
51	zabalik	hiruko	gexo	habuin
52	balioitsu	neu	zentralizataile	larretxe
53	berraztertu	biziarazi	armagabe	taup
54	eguberri	ustezko	momotxorro	inguralde
55	zeia	bakarka	indigaztainondo	migrante
56	dirudienez	barrura	txortalo	oholtzape
57	zegi	jatorriz	ezpatadantza	fototeka
58	guztizko	irundar	glass	movil
59	kartzeatu	adimentsu	tentaldi	zalagarda
60	ustezko	polideportibo	mahi	ziztor
61	legezko	nahikoa	ur	mineri

Posizioa	Elh200_rfrec / ETC_rfrec	Twitter_rfrec / ETC_rfrec	Twitter_rfrec / Elh200_rfrec	Twitter_rfrec / ETC_Elh200_rfrec
62	baketsu	baliotsu	trikatu	txinta
63	maket	berregin	ingurukatu	baluarte
64	saltoka	erdutu	izain	txamarreta
65	garrantzitsu	delako	tease	cadiztar
66	lapurtar	kartzeatu	arazle	oxala
67	mutxi	dirudienez	hargindegi	ugaitz
68	esplikazio	berriki	beroketa	traol
69	belasko	aberia	beno	ikasmin
70	ezagutarazi	belasko	hondeagailu	ahi
71	ikusarazi	gehiegizko	begigorri	kurkubi
72	gutxiegi	saltoka	tridente	mordaza
73	historikoki	lapurtar	sprinter	euskaitz
74	oriotar	haizegune	ondoezik	txotor
75	formalki	misteriotsu	hondarpe	kupe
76	sutsu	baldintzape	ikusketak	aurrestreinaldi
77	odoltsu	zeu	goizale	buruto
78	kementsu	baketsu	andetar	urgatzi
79	geroko	zentsu	jiratu	ihitza
80	moduzko	gutxiegi	txamarra	artefaktu
81	eregi	laguntxo	lorezale	twit
82	gehi	amen	relazio	hankazabal
83	aitzina	pentsiodun	apiril	turf
84	apurka	dozenaka	elkarreragile	haizeberritu
85	baiezko	maket	petraldu	animos
86	zeu	xitu	itsusikeria	gizaeskubide
87	zazpigarren	pago	mandatu	polideportibo
88	halakoxe	garrantzitsu	alergeno	moreno
89	berridatzi	ahozko	berantetsi	tanda
90	bertsu	zapra	sintetiko	puto
91	borondatezko	zazpigarren	errekurrente	galgorri
92	larriki	joko	zarpaildu	saratu
93	ereduzko	trabela	apurkor	argazkigile
94	berrantolatatu	gaineko	ezkutaleku	semifinal
95	giza	kostata	klasifikazio	estreno

Posizioa	Elh200_rfreq / ETC_rfreq	Twitter_rfreq / ETC_rfreq	Twitter_rfreq / Elh200_rfreq	Twitter_rfreq / ETC_Elh200_rfreq
96	pasonibel	pina	barrendero	ahaldungai
97	berregituratu	giza	zarabanda	rikar
98	orokorki	sutsu	itsasgora	gipuzkoano
99	adimentsu	pozoitsu	kontrapuntu	jenial
100	latindar	baiezko	esperoan	sorteo

Tesiko bosgarren kapituluaren historia txikia («Zer arraio idatzi dok zorioneko bosgarren kapitulu horretan?!»)

JESUS MARI MAKAZAGA

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

Pello Salaburu irakasle, lankide eta adiskidearen erretiroa dela eta, omenaldi modura, erreko-nozimenduzko hitz batzuk juntatu nahi nituzke. Urte mordo bat bada elkar ezagutzen dugula, eta, halako batean, erretiroa hartzeko ordua etorri zaio, ondo irabazitako erretiroa hartzeko ordua.

Pello irakasle

1986. urtea izango zen Pello ezagutu nuenean: irakasle izan nuen, Euskal Filologia karrerako azken urteetan. Sintaxia irakasten zigun, mugimendu generatibistaren ikuspegitik, eta, gaia hain abstraktua eta berezia izanik, beti iruditu izan zait azalpen ezin argiagoak ematen zizkigula; gaiak eta edukiak oso txukun ordenatuta eskaintzen zizkigun, eta atsegina eta erakargarria egiten zitzaidan zenbaitentzat hain gai lehorra izan zitekeenaren irakaskuntza. Beti, gainera, pazientzia handiz eta patxadaz ematen zizkigun eskolak.

Pello Euskaltzaindiko Gramatika Batzordeko kide

Handik urte gutxira, Euskaltzaindiko Gramatika Batzordean bekadun-lanetan sartu nintze-nean egin nuen topo berriz Pellorekin. Artean ez zen batzordeburu (Patxi Altuna zen garai artan, oker ez banago), baina batzordearen sortzaileetariko bat zen, eta ni, batzordean sartu berri zen bekadun lotsati bat. Aukera polita izan zen niretzat Pelloren ez ezik batzordekide guztien jakintza-ren berri izatea, lehen eskutik gainera: oso aberasgarriak izan dira urte hauek guztiak niretzat, bertatik bertara izan baitut aukera batzordekide jakintsuen ezagutza handi hartatik edateko. Batzorde horretan hainbeste urte egin baldin baditut, han ikusi dudana adiskidantza-giro bikainagatik ere badela aitortu behar dut. Ni bezalako gazte sartu berri hark beti izan du lankidetzaren eta konfiantza-giro ezin hobea batzordean, eta ez du inongo kexurik. Alderantziz, den-dena aldekoa iruditu izan zait beti: giro akademiko eta adiskidantzazko ezin egokiagoa.

Pello nire doktoretza-tesiaren zuzendari

Doktoretza-tesia Elgoibarko ahozko euskararen gainean egiteko asmoa nuen, eta horretarako, Pellori berari eskatu nion zuzendari izateko, harekin neukan konfiantzagatik, haren ezagutzagatik eta Baztango euskalkiari buruz egin zuen tesi bikainagatik. Baiezkoa eman zidan, eta hor ere beste lotura bat sortu zen gure artean. Lehenbiziko urteetan, irakaskuntzako kontuak zirela eta, batzordeko lanak zirela eta, ez zuen aurrera egin tesiak; baina behin, halako batean, gaiari serio heldu nionetik, Pelloren erantzuna bikaina izan zen beti: galdera bat egiten nionean, materiala pasatzen nionean berak gainbegiratzeko eta iritzia emateko, berehala iristen zen erantzuna. Berak seguru ezin duela gauza bera esan negatik, hark eskatutakoei ematen nion erantzuna askotan luzatu egiten baitzen. Baina inoiz ez zuen pazientzia galtzen: «egingo du halako batean» pentsatuko zuen seguruenik, eta itxoiten geratzen zen, eserita badaezpada. Askatasun handia ematen zidan, eta, bere iritzia ematearekin batera, ondo onartzen zuen nirea ere, baita biak bat ez zetozen kasuetan ere.

Pello lankide

Sailkide izan gara Euskal Herriko Unibertsitateko Euskal Filologia sailean aurrena, eta Euskal Hizkuntza eta Komunikazioa sailean ondoren. Ni neu zerbitzu-betekizunetan nabil azken urte luzeetan UPV/EHUko Euskara Zerbitzuan, eta han ere, 2011tik aurrera (ni zerbitzuburu naizenetik) paraleloki zuzendu ditugu hark Euskara Institutua eta nik Euskara Zerbitzua, konfiantza-giroan, lankidetza onean eta era ahalik eta koordinatuenean betiere.

Pello adiskide

Urte hauetan guztietan, errespetuzkoa eta konfiantzazkoa izan da beti elkarrekiko harremana; hainbeste gauza partekatu ondoren, adiskidantza-harreman on bat izatera iritsi gara, eta hala jarraituko dugu etorkizunean ere.

Nire doktoretza-tesia eta bosgarren kapitulua

Herri jakin bateko hizkerak aztertzeke dialektologiako ohiko lan monografikoetan, hizkuntzaren arlo hauek izan ohi dira aztergai normalean: fonologia, izenaren morfologia, aditzaren morfologia, sintaxia eta hiztegia.

Horietatik aparte, berri-emaileengandik jasotako testuak ere transkribatu izan dira, hizkeraren erakusgarri orokorrago bat emateko, eta, berri-emaileen munduaren eta mundu-ikuskeraren testigantza bat finkatuta, utziz, lanari balio etnologiko erantsi bat emateko.

Horiek ere egin nituen nire tesian: testu sorta oparo bat transkribatu nuen berri-emaileen bizitzaren alderdien arabera sailkatuta: eskola-kontuak, umetako jolasak, neskametzak, soldadutza, Gerra Zibila, ogibideak, baserriko lanak, jaiak, erromeriak eta bestelako ospakizunak, hizkerari buruzko testuak (oharmenezko dialektologiaren arloan), eta abar. Zentzuzko aholku batzuk salbuetsita, arazorik ez halakoetan ere tesi-zuzendariaren partetik horiek ere sartzeko.

Dialektologiaren arloan ere, ekarpen bat egin nahi izan nuen tesian: euskalkien eta azpieuskalkien azterketan pauso bat eman nahi izan nuen, Mendebaleko eta Erdialdeko hizkeren arteko muga horretan, argi pixka bat egin inguruko hizkerekiko konparazio kualitatiboa eta kuantitatiboa eginez lanaren amaieran.

Ondo joan ziren kontu horiek, oro har, tesiaren defentsaren egunean. Izan ziren ongi arrazoitutako oharrek ere, noski, asko estimatu eta aintzat hartu ditudanak ondorengo bertsioetan; ohar xebreak ere izan ziren. Bat aipatuko nuke, halako tentsio-unean ere barrea eragin zidalako: «Abarka asko, apreta¹ gutxi» esamoldearen bidez laburbildu zuen Patxi Goenaga epaimahaikideak bere kritika, gaztelaniazko «El que mucho abarca poco aprieta» esamoldearen egokitzapen zoragarriaren bidez. Tesian gauza asko sartu nahi izan nuelako, gai batzuetan gehiegi sakontzera ez nintzela iritsi leporatu zidan; eta ados neu ere. Gero etorri zen Orreaga Ibarra epaimahaikidea nik aurkeztu nuenaz gainera ia-ia beste tesi bat egiteko eskatzen zidana, alegia belaunaldi gazteagoen euskara ere ikertzea eta bi belaunaldien arteko euskarak konparatzea: epaimahaikideak ez zetozen bat, eta, jakina, ni erdibidean, poztxo horrekin: ezin, ba, biei kasu egin, esaten zituztenak nire ustez kontraesanean zeudenez gero.

Idatziz eman zizkidan Miren Azkaratek bere jarduneko hitzak, eta biziki eskertzen diot. Ez dut esango epaimahaikide guztien iritzien adierazle denik Mirenena, baina guztienak ez ziren urrun izan, eta ekar ditzagun hona Mirenen hitz batzuk, haren baimenarekin:

«Oso argi adierazten duzu atal bakoitzean, Fonologian, Izenaren morfologian, aditzean... zer egin duzun eta zer hartu duzun oinarritzat. Alde honetatik, sintaxiaren atalean esaten duzun bezala, dialektologian ohikoak izan diren aztergaietatik «haruntzago» joan zara. Egia da oso aberasgarria dela bai Esaera, Esapide eta Interjekzioen bilduma, eta zer esanik ez Hiztegiarena (kontuan izanik, gainera, zeuk diozun bezala, zenbat diren inguruko hizkerez ondutako hiztegiak).»

Atal horietan, zehaztasun gehiagorekin edo gutxiagorekin, era honetako lan monografikoetan egin ohi diren kontuak aztertu nituen.

Berri-emaielen partetik jasotako ahozko testuen azterketa egiten ari nintzela, haien ahozko jarduneko honako ezaugarri hau fonologiaren atalean sartzen nuen, horrako beste hori sintaxiaren *zakuan*, harako hura aditzaren morfologiarenean, eta abar eta abar.

Hala ere, baziren tarteka-tarteka sailkapen zurrin horretan (fonologia, izenaren morfologia, aditzaren morfologia, sintaxia eta hiztegia) lekuri ez zuten ezaugarri batzuk, kokapen egokirik ez zuten hizkuntza-item batzuk. Betidanik izan dut Diogenes sindrome ikutu bat, ustez ez-morbidoa, eta badaezpada ere halako ezaugarriak gorde egiten nituen, «Bestelakoak» edo «Saski-naski» (ez naiz ondo gogoratzen) izeneko *zaku* sortu berri batean. Beste ataletan leku garbirik ikusten ez nuen aldietan, zaku berri horretan sartzen nituen nire arreta erakartzen zuten hizkuntza-ezaugarri berezi horiek.

Zaku berri hori gero eta handiagoa eginez joan zen; ahal nuenean, garbi samar ikusten nuenean kontzeptua azpialat berriak sortu, eta haiek elikatzen nituen poliki-poliki. Guztiz gozagarria egiten zitzaidan zaku berri hori elikatzeke lana, eta horrek ere laguntzen zuen, ondo lagundu ere, doktoretza-tesia bezalako arnas luzeko lan batean interesa eta arreta ez galtzen.

¹ Zenbait hizkeratan, *abarketa* eta *espartina* hitzen ordaina da.

Tesiaren lehenbiziko helburuei uko egin gabe (hizkera jakin baten ohiko ezaugarritzea eta haren kokapen dialektala egitea), helburu berri baten ateak ireki zitzaizkidan aurrez aurre: neurri batean hizkera horren ezaugarritzea egiten ari nintzen atal berriarekin ere, baina ikusten nuen orokorragoa zen zerbait jasotzen ari nintzela, eta argi neukan ez niola ezikusiarrena egin nahi.

Iruditzen zitzaidan jasotzen ari nintzen ezaugarri horietan zegoela gure guraso eta aitona-amonen euskararen sena, haien euskararen adierazgarritasuna, komunikagarritasuna eta poesia, nolabait esateko. Ezaugarri horiei esker da haien euskara hain komunikagarria, hain adierazgarria, hain polita eta, azken finean, hain eraginkorra.

Zer esanik ez, Pellori aipatu zaku berri honen kontua, eta, «aurrera, bai», «interesgarria oso»² bat haren erantzuna. Ados hura ere gai honen garrantziarekin. Pozarren ni: lan gozagarria egiten ari nintzen, eta banuen *nagusiaren* oniritzia, gainera. Gure artean esaten den bezala, *Gustoko mendixan, aldarrik ez*.

Tesiko beste atal guztiak adinakoa edo, kasu askotan, handiagoa izatera iritsi zen, bolumenari dagokionez, atal berri hau. Bazegoen, bai, zer hustua gure berri-emaileen ahozko testu ezin interesgarriago eta aberatsagoetatik. Eta gero esango dute zenbaitek, edo berri-emaileek berek, edo haietako batzuek behintzat, gure euskara ez dela ona, Azkoiti aldekoa edo, areago, Tolosa aldekoa dela ona, garbia, eta abar. Bai zera! Edo bai, haiek onak dira, baina, Erdialdeko euskara horiei ezer kendu gabe, eman diezaigun geureari dagokion balioa, handia da eta.³

Gai horrekin nuen lehenbiziko helburua beteta zegoen: jasoak nituen nire ustez pertinentek ziren hainbat ezaugarri, eta hor geratuko ziren jasota behintzat, hobeto edo okerrago sailkatuta eta azalduta, baina jasota behintzat. Eguneroko jarduneko elkarrizketetan, ahozko jardunean, berri-emaileek darabiltzaten esaerak, esapide eginak, interjekzioak eta hitz adierazgarriak, harridurazkoak; guztiak hizkuntzaren ahozko erabileraren bereizgarri, hizkuntzaren bizigarri, freskagarri eta, zergatik ez esan, edergarri (metafora eta baliabide literarioz josiak haietako asko).

Izan ere, garai bateko euskaldunen bizimodua, pentsaera, sentitzeko eta hitz egiteko era badoakigula ikusita, horien guztien ispilu diren hizkuntza-baliabideak bildu nahi nituen, ahalik eta era sistematikoenean, ordenatuenean. Horietako asko hiztegiaren atalean sar zitezkeen, baina hiztegiaren *zaku* handiaren hurrenkera alfabetiko eta «arbitrario» horretan, iruditzen zitzaidan «galdu»-edo egiten zirela altxor horiek, hainbeste sarrera eta azpisarreraren artean. Horregatik ekarri nituen zaku berrira, irizpide tematiko-kontzeptualen bidez ikusgarriago egiten nituelakoan.

Bihar-etzikoko euskaldun gazteei, haien euskarari ekarpen bat egingo niela uste nuen ahozko jardunaren baliabide haiek bilduta: ulerterrazagoak, sinpleagoak, naturalagoak eta adierazgarriagoak izango dira haien testuak, idatzizkoak nahiz ahozkoak, gure aitona-amonak erabiltzen zituzten baliabideen laguntzarekin.

² Beharbada, nire atzerapenetara hain ohitua, era honetako jarraipenen batekin: «... baina ez gehiegi luzatu; laster bukatzea komeni zaik».

³ Bada gure berri-emaileen artean konplexu bat, darabilten euskarari dagokiona: diote ez dela garbia, nahaste handia daukala («Bizkai aldeko berbak eta Gipuzkoa aldekuak, nahasketa haundixa dagok»; berbak ... eta ezaugarriak, gaineratuko nuke nik), eta garbiagoa dela Erdialdera egiten duena, Azkoiti edo Tolosa aldera. Bestalde, bada gure inguru hurbilean, Eibarren, Soraluzen, Elgoibarren..., hiztegi kontuetan mailegutzarako, erdarakadetarako joera markatu samar bat (beti aipatu izan diren *kaballua* eta *bentania*, horren adierazgarri), gure euskararen pertzepzioa iluntzen eta estigmatizatzen duena. Ez dira ohartzen, ordea, darabilten euskara zoragarria dela, haien helburu komunikatiboetarako tresna ondo zorroztua dutela, eta eraginkorra, dinamikoa eta ederra dela; beste leku hizkeraz garbiago batean erabiltzen dena bezain ederra, ederragoa ez bada.

Izan ere, lagunarteko erregistroan, batez ere belaunaldi gazteagoen ahozko jardunean, hutsuneak nabaritzen nituen. Hizkera mota horrek behar-beharrezko duen adierazgarritasuna eta indarra bilatzeko, gehiegixotan doaz gure gazteak erdaren iturrietara, euskaraz aurkitzen ez duten ur freskoaren bila. Izan ere, ahozko euskararen transmisioan, hutsuneak izan dira azken urteotan, eta gazteek ez dute, oro har, Elgoibar bezalako herri askotan behintzat, horren berri zehatzegirik izan, gero beren eguneroko lagunarteko giroan erabiltzeko. Erdararen kasuan guztiz bestelakoa da egoera gure herrian, oro har, osasuntsuagoa; gure herriko euskararen kasuan, herriko gazteen arteko elkarriketei zaharren elkarriketen naturaltasuna, adierazgarritasuna falta zaie.

Gaur egun, Euskaltzaindiaren arauen arabera euskara nahiko ondo idazten eta erabiltzen dute gazteek, oro har. Euskararen gramatika ere ondo ezagutzen dute jeneralean; baina euskara egokiaren eta eraginkorraren kontuan zenbaitetan herren dabilta, nire ustez. Kalean adin batetik beherakoek darabilten euskara ere askotan neutroa da, aukeran; zenbaitetan ez du indarrik, gatz eta piperrik, eraginkortasunik mintzakidearengan bilatu nahi den erreakzioa lortzeko. Argi dago azken urteetako transmisioak eta irakaskuntzak zerbaitetan huts egin duela: auzi honi dagokionez behintzat, nabarmena da ahozkoasuna ez dela behar bezala landu. Hori garatzeak asko lagunduko liguke hizkuntza era naturalagoan eta adierazgarriagoan erabiltzen.

Hor ikusten dut nik gaur egungo euskararen hutsune larrienetako bat: ahozko lagun arteko jardunean erabiltzeko tresna egokia falta da askotan. Kezka hori izanik, gure herriko (eta inguruetako) zaharrek (errespeturik handienarekin erabiltzen dut berba hori) darabilten euskararen sakondu nahi izan dut, euskara horretatik atera daitekeen etekinik oparoena atera ahal izateko: gazteen ahozko jarduna naturaltasunez eta adierazgarritasunez hornitu, idatzizko testuak ere eraginkorrako egiteko hobekuntza-bide batzuk aztertu, eta abar.

Hor kokatzen dut tesiko bosgarren kapitulu honekin egin nahi izan dudana ekarpena. Elgoibarko ahozko jarduneko adibide batzuk bildu nahi izan ditut, adineko hiztunek ahozko jardunean darabiltzaten baliabide estilistiko edo poetiko zenbaiten erakusgarri gisa⁴. Egutegiko-egutegiko darabilgun euskararen adierazgarritasunaren eta eraginkortasunaren mesedetan, betiere.

Gure ahozko tradizioetik elikatu nahi nuen; handik bildu nahi nituen hainbat adibide. Ikusten bainuen euskararen hobekuntza-bideetako bat tradizio hori lantzea dela; ez bakarrik, sortu ere egin behar baitute hiztunek. Edonola ere, sormena eta berrikuntzak tradizioa ezagututa egitea litzateke egokiena: tradizioan dugunetik hartu (erabil dezakegun neurrian), eta sortu (tradizioan ez duguna edo zaharkitua geratu dena), baina ez sortzen jardun tradizioan baduguna eta erabil dezakeguna.

Adierazgarritasunaren bila aritu nintzen, ahozko hizkerak eskaintzen dizkigun baliabide horien erakusgarrien bila: alde horretatik ikusten nion nik interesa *zaku* berri honi, ahozko hizkeraren baliabide garrantzitsu batzuen bilduma den aldetik; eta ez hainbeste dialektologiaren aldetik.

Nik, aditu izan gabe, eman nitzakeen argibideei baino garrantzi handiagoa eman nahi nien adibideei berei; adibide kopuru zabala eskaini nuen, eta haietan sumatu uste nituen joerak azaldu nituen. Pozik nengoke helburu horretan behintzat asmatu badut.

⁴ *Poetiko* hitzaren zentzurik zabalenean, Joxerra Garziak (1998) dioen bezala, «gauzak modu itxuraz berri eta berezian» adieraztearen zentzuan; hain bereak dituen hitzak hona ekarriaz, «esana era bakan-betean» esatea da harentzat poesia, bere zentzurik zabalenean. Esanaren bete-bakantasan irudipen hori gudan (mintzakidearengan, mezu-hartzailearengan) sortzeko amarruak dira, haren esanetan, literatura-baliabideak, baliabide poetikoak.

Arloko adituek hizkera kolokialaren edo lagunartekoaren definizio eta ezaugarritzat aipatzen zituztenak ere bildu nituen kapituluaren sarreran.

Arlo honetan aitzindarietako bat dugun W. Beinhauer-ek gaztelania kolokialari buruzko definizio hau ematen du:

«El habla tal como brota, natural y espontáneamente en la conversación diaria, a diferencia de las manifestaciones lingüísticas conscientemente formuladas, y por tanto más cerebrales, de oradores, predicadores, abogados, conferenciantes, etc., o las artísticamente moldeadas y engalanadas de escritores, periodistas o poetas.»

Berezkotasuna, naturaltasuna eta planifikazio falta dira definizio horretatik atera daitezkeen hizkera kolokialaren ezaugarriak.

Hona M.C. Lassalettak dioena:

«El lenguaje coloquial es una fracción o nivel de lenguaje total que se destaca por su carácter pintoresco reflejado en multitud de expresiones y vocablos intraducibles a otros idiomas, fundados muchas veces en alusiones metafóricas y que posee una gracia, viveza, gran espontaneidad, concreción y expresividad que lo distinguen de otros niveles.»

Hor aipatu direnetako asko aurkitzen dira nik zaku berrian jasotako ezaugarrietan: esapide berezi itzulezinak eta esapide metaforikoak; esaeren eta esapideen atalean ditugu horiek, eguneroko bizitzako metafora eta konparazio bidez osatuak.

Lagunarteko hizketa deskribatzen duten adituek honako ezaugarri hauek aipatzen dituzte (ikus Briz): orainaldian gertatua, ahozkoa, interaktiboa, berehalakoa, informala, berezkoa, gai ez-berezituak tratatzen dituen, gutxienez bi hiztunen artean gertatua, prestatu gabea, berdinaren artean gertatua, eta abar.

Jardun lagunartekoaren konstanteak ere aipatzen ditu Brizek:

- Sintaxi-egitura parataktikoak.
- Egitura parafrastikoak erabiltzea.
- Geldialdien ugaritasuna.
- Erredundantzien erabilera joria.
- Enfasiak leku nabarmena.
- Elipsia nagusi.
- Anakoluto ugari agertzea.
- Estilo zuzenaren erabilera.
- Hitz-hurrenkera pragmatikaren araberakoa.
- Mintzakideak osatu behar dituen enuntziatu bertan behera utziak.
- Entonazio goranzkoak, adierazgarriak.
- Egitura indargarriak (hiperbolea).
- Aukera lexiko urria.
- Parahizkuntzak duen garrantzia.
- ...

Behin ezaugarri horien bilketa eta nolabaiteko sailkapen-lana eginda, emaitzak ezagutzea gertatzen zen, epaimahaikideen ahotik.

Pozgarria izan zen, adibidez, Miren Azkaratek esandakoa:

«Dialektologian ohikoak izan diren aztergaietatik «haruntzago» joan zara (...) Haruntzago joan izan horrek berarekin dakar emaitzen erabilgarritasuna, bai gramatikaren ezagutza, bai dialektologian, baina, batez ere, ahozko hizkuntzaren eta herriko hizkuntzaren ezagutza. Ahozkotasunak eta gaurko gazteen komunikaziorako ahalmen kamutsaz kezkatuak dabilzanak (gabiltzanok) benetan eskertuko dute esapideetan-eta egindako bilduma. Gai labainkorra da, berak bakarrik tesi bat osatzeko beste emango lukeena (sailkapen irizpideetatik hasita), baina bilduma hor dago. Hortik aurrera, gehiago sakondu nahi duenak, hortxe du lehengaia.»

Eta Miren bezala, gainerako epaimahaikide guzti-guztiek (Orreaga Ibarrek, Xarles Bidegainek, Patxi Goenagak eta Beñat Oihartzabalek) aipatu zuten bosgarren kapitulua, era nabarmenean aipatu ere. Ohikoa den bezala, bakoitzak bere arloan egin zituen ekarpen jorienak eta lanaren eragozpen handienak (Goenagak eta Oihartzabalek sintaxian, Azkaratek morfologian, Bidegainek hiztegian eta Ibarrek dialektologian eta hizkera lagunartekoan), baina denak gelditu ziren, denbora luzea eskainiz gainera, tesiaren bosgarren kapitulu hori iruzkintzen. Nobedadea izan zen haientzat halako lan batean gai hori sartzea: hala aitortu zuten; ikusten zuten hor erabilgarria izan zitekeen material bat, nahiz beharbada beste era batera sailkatuko zuten...; denek izan zuten hizpide bosgarren kapitulu hori⁵.

Eta ni pozarren, nire ekarpenak berritasuna ekarri zuelako, eta horrekin epaimahaikideen arreta eskuratu nuelako. Ordura arte ez nengoen lasai, baina hortik aurrera zertxobait lasaitu nintzen.

Defentsa-lana amaitu eta gero ere, aretotik kanpo, jarraitzen genuen bosgarren kapituluaren kontuarekin. Dirudenez, ez zen nire inpresioa bakarrik epaimahaikideen arreta eskuratu izana: Koldo Zuazo adiskidearen lehenbiziko hitzak hauek izan ziren: «Baina zer arraio idatzi dok zorioneko bosgarren kapitulu horretan?!». Jose Luis Ormaetxea *Txipik* ere berdintsu: hark ezagutzen zuen, Koldok bezala, nire tesia, eta aldezturik zoriondu ninduen kapitulu berri hori sartzeagatik. Koldok ere bai bere garaian, baina, jakina, tesi-defentsaren egunean ez zen gogoratzen zer zenbaki zeukan kapituluak.

Ez zen nire inpresioa bakarrik, beraz: entzuleena ere bazen.

Zer idatzi nuen? Koldoren galderari erantzunez, hona hemen ezaugarri horien erakusgarri labur-txiki bat beren adibide eta guzti (halaxe ulertzen dira ondoen eta).

1. *Konparazio bidezko esapideak*

Solaseko norbaiten ezaugarri bat edo egoera bat era enfatikoan adierazteko, jakina da konparazioak erabiltzen direla sarritan. Besterik gabe, era neutroan, izenondo arrunt baten bidez, pertsona horren ezaugarri hori eman beharrean, konparazio enfatiko bat egiten da, eta hala, berebi-

⁵ Tesi-defentsarako prestatu nuen tesiaren kopian, «Esaerak, esapideak eta interjekzioak» izenburupean jaso nituen; geroago etorri zen argitalpenean, berriz (ikus Makazaga 2010), «Ahozko hizkeraren ezaugarri batzuk» izenburua eman nion.

ziko indarra hartzen du deskribapenak, bai eta, asko erabiliak ez diren kasuetan, asmatuak diren konparazioetan, grazia ere.

Hona halakoen erakusgarri batzuk⁶:

- *Zorria baino alferragoa da Joxe.*
- *Ezpala baino argalago geratu da aita, gaixo egon eta gero.*
- *Txoria baino arinagoa da ume hori.*
- *Ardi galdua bezala dabil, beti bakar-bakarrik.*
- *Uztarrian joateko modukoak dira lagun bi horiek, beti elkarrekin.*
- *Infernuko deabrua baino gaiztoagoa da mutil hori.*
- *Urrea baino finagoa dut alaba.*
- *Txita batek bezainbeste jaten du ume horrek.*
- *Txakurra eta katua bezala dabilza beti seme-alabak.*
- *Ume batek baino pena handiagoa hartu nuen Errealak galdu zuenean.*
- *Garizuma baino luzeagoa zen emakume hura.*
- *Beldurraren beldurrez, tximista baino azkarrago etorri zitzaigun umea.*
- *Telarik gabeko aterkia baino argalagoa da mutil hori.*⁷

2. Galdera erretorikoak

Gure diskurtsoan erabiltzen dugun beste baliabide bat da. Galdera erretorikoetan, dakigun bezala, ez dugu mintzakidearen erantzunik espero; beharbada berrespen txiki bat bai, baina, jaso ditudan kasuetan, batez ere diskurtsoa antolatzeke baliabide bat da, diskurtsoaren erritmoa antolatzeke, diskurtsoaren harian, dena asertzio bidez eman beharrean, bariazio estilistikoa egiteko.

Hona adibide batzuk:

- *Eta bazen, ba, aldeko mutila? Gero, joan zen Santaeske egunean gerrara.*
- *Zapatak apurtu samartuta zeuzkan. Eta zer egin zuen? Brigadari eskatu berriak.*
- *Eskolan, baserritarrek ere, zer? Neguan sei bat hilabete egin, eta gero lanera.*
- *Zer da, eta azkenean zigorra maisuak!*
- *Zein etorriko, eta etsairik gaiztoena!*

3. Onomatopeiak

Ahozko jardunean, hitzean eta hortzean erabiltzen dira gure inguruan entzuten ditugun hotsak imitatzeke erabiltzen diren berba hauek, bai adinekoen hizkuntzan, bai umeenean ere. Ekintza komunikatiboan, esaten duguna era zuzenagoan, argigarriagoan, ilustratiboagoan adierazteko balio dute berba hauek.

⁶ Lan honetan, euskara batuko forman eman ditut Elgoibarren batutako ezaugarri hauek.

⁷ Nire berri-emaileetako batek asmatutako konparazioa.

Hona Elgoibarren entzun ditudan onomatopeia batzuk:

- *Bolo-bolo* *dabil* herrian *txutxumutxu* hori.
- *Kalamua* [*mendia*]n geratu ziren, eta sei hilabetea hortxe, *dinbi* eta *danba*, bazter guztiak garbitzen [*bonbak*].
- *Hori* zen ardoa erraz edatea, *dzanga-dzanga*!
- *Barra-barra* erdaraz kaletarrak.
- Amarekin joaten ginen bost lagun edo. Koxkorrak denak, *ttiki-ttaka*, *ttiki-ttaka*, txarriku-meak amaren ondoan bezalaxe joaten ginen.
- Haien artean beti erdaraz. Euskara ahaztu! Eskoletan laga ez, eta gero kalera joan, eta *tto-tto-tto*, dena erdaraz.
- Belarretan, *ras-ras-ras* garbitzen genituen zapatak [erromeriara iristean].
- *Ri-ra* egin dugu dena, hasi eta amaitu egun batean.

4. Elipsia

Elipsiarena da ahozko jardunaren ezaugarri garrantzitsuenetako bat; komunikazio dinamikoa goa, biziagoa eta naturalagoa lortzen da, zalantzarik gabe, jakintzat ematen den zenbait informazio isilduz; kontakizunak egitean, elipsiaren bidez asko irabazten du gure jardunak erritmoan. Aditzaren elipsia da ohikoena: aditz laguntzailearena, edo, are, aditz osoarena ere bai.

Hona adibide batzuk (kakotx artean jarri dut eliditutako parteak):

- Amak arrazoia: jakina, *liburu asko behar* [*ziren*].
- Zintzoak izango zineten? - Horixe! *Duda* [*egiten duzu*].
- Bueno, martxan zagoz grabagailua, ala ez? - Bai, aspalditik. - *Adarrik ez* [*iezadak jo*].
- Hogeita hamabostean *ezkondu* [ginen], eta hogeita hamaseian *gerra* [piztu zen].
- Eta gero, txarren-txarrena zen, zera, *biderik ez* [*zegoela*] baserrietara.
- *Laga ez* [*zien*] hark neskak eramaten; bestela, eramango zituzten neskak ere frentera.
- Umeak *laga* [*zituzten*] gurasoei, eta pixka bat haiek deskantsatu egin behar zutela, eta *etorri* [*ziren*].
- Euskaraz *ezin* [*genuen*] egin, eta erdaraz *jakin ez* [*genekien*]: haiek ziren komeriak!
- Dirurik ez, eta *zegoenarekin konformatu behar* [*izaten genuen*] ...

5. Errepikapenak

Ahozko jardunean erruz erabiltzen dira errepikapenak ere, zerbaiti enfasi edo indar berezia eman nahian edo: errepikapen indargarriak dira, osagai baten gainean enfasi berezia egin nahi dugun kasuetan erabiltzen direnak. Errepikatzen den osagaia edozer kategoria gramatikaletakoa izan daiteke: batzuetan, erakusleren bat izan daiteke; beste batzuetan, izenondoa, aditza edo perpausa bera ere bai.

Hona Elgoibarko jardunetik bildutako adibide batzuk:

- *Hori duk marka, hori*!
- *Hura zen ona, hura*!

- *Hi bizi haiz, hi!*
- *Andres ume txiki-txikia zen.*
- *Egunero bi buelta ematen ditu gizon horrek. Ordu-orduan, gainera.*
- *Bai, askotan etortzen dira. Beti-beti ez dira etortzen, baina askotan bai.*
- *Lotsagabeagoa zen, lotsagabeagoa...*

Errepikapenak hirukoitzak ere izan daitezke, indar handiagoa emateko:

- *Ur zuri-zuri-zuria ikusten da, gainera, garbi-garbi-garbia barruan.*
- *Isil-isil-isilik etxera orduantxe.*

Perpausak eta sintagmak ere errepika daitezke; halakoetan, bariazioagatik, oso ohikoa da hurrenkera-aldaketa egitea perpaus edo sintagma batetik bestera:

- *Gerra da azkena! Azkena da gerra, e!*
- *Aste Santuan eta, [jende] gaztea dena, dena gaztea.*

6. *Kontakizunetako betegarriak eta makuluak*

Ahozko jardunean arruntak dira isiluneak: hitz bat ez zaigunean gogora etortzen, esaterako. Isilune horiek betetzeko, makuluak edo betegarriak ditu eskura ahoz ari den hiztunak, harik eta etortzen ez zaigun hitz hori ahora etorri arte. Txantxa moduko esaldi hau da horren adierazgarri muturrekoa:

- *Zereko zera zertu da zerarekin zerean.*

Ekar ditzagun hori baino adibide errealago batzuk:

- *San Migel, hemen inguruetan dagoen zerik ederrena dago [auzorik].*
- *Ezin, ba, joan zerera... oportetara.*
- *Andre bat dago hemen, Zestoakoa da izatez, baina hor zerean bizi da, Markinan.*

7. *Ezezko formen ugaritasuna*

Nabarmena da ezeztapenaren erabilera joria ahozko jardunean. Gertaera eta deskribapen asko eta asko adierazten dira ezezko forman, era markatuan adierazi ere; horretarako, esan nahi denaren kontrakoa ezezko forman jarri besterik ez da egin behar:

- *Ez duk berehalakoan etorri!*
- *Gerraren berri baldin badakik, ez haiz gaur goizekoa.*
- *Asko maite du anaia. Ez dauka, ez, saltzeko.*
- *Eskapada ederra egin dugu autoarekin: ibaira joan ez garenean!*
- *Triste egongo da, ba, lana galduta? — Ez dakizu asko!*
- *Pakita hil da. — Ez da egia izango!*

Ezeztapenen barruan, bada beste mota bat, ugaria hau ere; eta bitxia, zeren eta horrelakoetan gauza bat adierazteko literalki haren kontrakoa esaten baitugu: ukapenak dira (ezezko formak), baina baiezko adiera dute; ez dira, hortaz, literalki hartu behar, zentzu ironikoan baizik, hitzez hitz esaten denaren kontrako zentzuan.

- *Esan diot gehiago ez molestatzeko.* — *Ez alferrik!*
- *Txerrikeria eta ustelkeria nagusi da gaur egun.* *Ez da giro bazterretan!*
- *Telebistako euskara ulertzen duzue?* — *Ulertuko ez dugu, ba!*
- *Ikasi, zeozter ikasten zenuten eskolan?* — *Baaai! Ikasiko ez genuen, ba!*
- *Eta nongoa nintzen ez dit, ba, galdetu artaburu honek!? Jakin behar zuen!*
- *Donostiako Semana Grande-ri ez zaio esaten* Aste Nagusia?

8. Hitz-ordena berezia

Atal honetan hizpide dudan gaiak ere azterketa sakonak ekarri ditu eta ekarriko ditu oraindik ere. Ur sakonegietan sartu gabe, hitz-ordena markatuaren adierazgarritasun-ahalmena aipatu besterik ez orain. Hanpatu edo markatu nahi den osagaia era jakin batean markatu nahi izaten da ahozko mintzamolde adierazgarrietan. Horretarako, aditzaren aurreko kokalekua edo, mintzagai hanpatuaren kasuan, perpaus hasiera izaten dira indar berezi hori markatzeko gunek.

Kasu hauetan, lehenbizi garrantzitsuagoa den informazioa ematen da, eta gero bigarren mailakoa. Adibideen eskuin aldean, kakotx artean, ordena neutroa zein litzatekeen adierazi dut.

- *Orain badaukazue ganadurik?* — *Ardiak orain.* [Orain ardiak (dauzkagu)]
- *Eta gero beste gauza bat zen: garoa mendian.* [Mendian garoa (egoten zen)]
- *Zer ganadu daukazu?* *Ahuntzak, hiru. Astoa, bat. Behorak, bi.* [Hiru ahuntz; Asto bat. Bi behor]
- *Joango gaituk kalera?* *Ez oraindik!* [Oraindik ez (gaituk joango kalera)]

Oro har, esan daiteke hitz-ordena malgu samarra aurkitzen dugula ahozko jardunean, mezua jasotzeko moduak (ahoz eman eta belariz jaso) baldintzatuta, neurri handi batean. Kontua da oso ugaria dela esaldiaren pisua aurrera ekartzea, adibidez, aditza (esaldiaren antolatzailea) aurrean ematea. Nahi adina adibide aurkituko ditugu edozein testu ahozkotan. Hona hemen batzuk:

- *Sukaldaria ere zen Hondarribikoa.*
- *Orduan hori zen sakratua.*
- *Orain batzuek egiten dute ezkondu hemendik kanpora.*

Galdegaia aurreratzeko estrategien artean, egitura kataforikoena dugu ahozko jardunean nahi-kotxo erabiltzen dena:

- *Futbolari gauza bat ikusten diot nik txarra: kanpoko jendea ekartzen dute.*
- *Hor ere (gurasoak kanpoan lanean) problema txarra dago: familia apurtuta geratzen da.*
- *Beste batean pasatu zen polita: (...)*
- *Eta, akordatzan naiz detailea: zeuden denak militarrek, batzuek ez zeukaten ihesbiderik, haiek fusilatu egingo zituzten seguru.*

Azkenik, deiktikoen erabilera bat aipatu nahi nuke: lehendabizi deiktikoa erabiltzea elkarriketakideak kokatzeko, eta horren ostean kokapen hori mugatzea, zehaztea. Hona, adibidez, hauek:

- *Santa Luzietan egiten dira horiek. Eta gero, hemen Markinan ere bai, eta hor Elorrion ere bai.*
- *Hemen irratian, ba, Arrate [irratia] ipintzen dugu.*
- *(...) beharbada bidea ebakiko dute hor Urkaregin, eta ezin pasatuko naiz Markinara; hor beharbada Alzola aldetik igo eta irits daitezke hona Urkaregira.*

9. *Hasiera-markatzailea duten mendeko perpausak*

Ahozko jardunean erabiltzen diren mendeko perpausen artean, deskodetzen errazenak dira nagusi. Arreta berezia merezi dute mendekoaren aurreko buruan markatzaile sintaktikoren bat dutenek: perpaus mendekoaren hasiera-hasieratik hura zer eratakoa den iragartzen digutenak dira.

Jaso ditzagun hemen adibide batzuk:

- *Esaten jardun zuten eze hura egiteko berrogeita bi milioi arbola kendu zituztela.*
- *Harik eta etxeko lan denak amaitu arte, ez ginen oheratzen.*
- *Lo egiten dut, baina esna ere asko. Nik, behin goizeko hirurak ezero, gau oso bat ez dut egin lo.*
- *Ez zen komeni han gelditzerik, zergatik eze berri ere beste tiro bat edo igual ailegatuko zen.*
- *Ba, zure ama eta aita nola San Lorentzokoak diren biak, ba, guk erabili dugun euskara berdina da.*

10. *Estilo zuzena*

Etengabea da estilo zuzenaren presentzia ahozko testu lagunartekoetan, hauetan gertaerak eta elkarriketak kontatzea oso arrunta baita. Estilo zuzena erabiltzen da bai beste pertsona baten hitzak jasotzeko, bai berri-emailearenak berarenak (bere buruari esaten dizkionak) jasotzeko ere. Hona adibide batzuk:

- *Nik aitari galdera: «Aita, han hola eta hola badago, ba, txabola bat? (...)».*
- *Hala esan nion neure buruari: «Ez hadi negozio horretan sartu!».*
- *Eta neure artean: «Hau duk pagotxa, hau!».*
- *Esan nuen: «Hau nola ordaintzen dute?».*

11. *Mailegutza berezia*

Hitz mailegatu batzuk erabilera adierazgarrian bakarrik entzuten dira; erabilera neutroan ez dute lekurik, Elgoibarko euskararen bederen (adibideen eskuin aldean, barra etzanaren eskuin aldean, adierazi nahi izan dut ezinezkoa dela erabilera neutro hori)⁸:

⁸ Kasu berezi honetan, Elgoibarko eran jaso ditut adibideak, euskara batuan oso bitxiak geratuko lirakeelako hitz hauen erabilera bereziak.

- Lehen grandia zuan motorra edukitzia, gero, e! / *Etxe grandia erosi dot.
- Hik dakak, hik, sanidadia, Patxi! Ordu bixetan Bilbon egon bihar, eta ordu bat t'erdietan Elgoibarren hago! / *Sanidade [osasun] ona dauka Inaziok.
- Han espezial ibili nitzuan! / ?Gizon espeziala da hori.
- Terriblia haiz, Inazio! / ?Gertaera terriblia (beldurgarria).
- Hori da printzipalena. / ?Kamiño printzipala hartu bihar dozu.
- Kantidadia etortzen zittuan Eibartik futbolera! / ?Jende kantidade egokixa etorri da.

12. Esaerak

Idea bat garatu ondoren, kontakizun bat burutu ondoren, amaieran laburpenak egitea, kontatu dena irudi batekin deskribatzea modu egokia da entzulearekiko komunikazioa bermatzeko, esan nahi duguna entzuleari hobeto ulertarazteko eta gogorarazteko, esan nahi dena gogoangarri egiteko: ondorioak ateratzea, esaldi edo esapide egin batekin nahiz esaera batekin amaitzea. Gauza bera gertatzen da gertakizun bat kontatu edo zerbait deskribatu nahi denean ere: orduan ere oso ondo etor dakioke hiztunari esaera edo esapide egoki-egoki bat, mintzakideari begi aurrean argazki bat jartzeak sortzen dion eragina sor dezakeena. Eguneroko bizitzako elementuei egiten zaie erreferentzia, eta oso ugariak dira baliabide metaforikoak eta konparazio bidezkoak.

Ahozko jardunean askotan erabiltzen den estrategia da.⁹

- Urte batzuetan ibili nintzen lanean neure kontura; gero, anaiaren bitartez hasi nintzen haren tailerrean: kanpoan hotz egiten zuen.
- Batera eta bestera ibili zen neskalgun bila, baina ezin egokia bilatu; ezin inorekin konformatu. Azkenean, on bila, on bila, bonbilla!
- Mendaroko gazte batek bota zuen bertso «berde» hori. - Bai ba, sasoian gereziak!
- Ezin izan dut asteburuan lana amaitu: azkar eta ondo, usoak hegan.
- Beti dabilta elkarrekin: hara zer pareta, karakola eta bareta!
- Kanpokoekin ondo geratzea gustatzen zaio: kanpolarrosa samarra da hori.
- Kalean beste aurpegi bat darabil emakume horrek: kalean uso, etxean otso.

13. Interjekzioak

Batzuetan, aurretik esan dena berresteko balio dute interjekzioak:

- Pagatu behar zen dantzaz egiteko?! — Erregular!
- Ondo bizi zinen zu, e! — Bai, horixe!
- Lagunduko didak etxea garbitzen? — Akabo ba!
- Aurten ere negua hotza dator. — Larri ibili!
- Gustatzen zaizu baserriko lana? — Beharko! Nora joango naiz, bestela?
- Aurten ere Kanarietara joango haiz oporretara? — Bienke!

⁹ Hona Ong-en hitzak: «Las fórmulas ayudan a aplicar el discurso rítmico y también sirven de recurso mnemotécnico, por derecho propio, con expresiones fijas que circulan de boca en boca y de oído en oído: *Divide y vencerás*, *Fuerte como un roble*. Las expresiones fijas (...) en las culturas orales no son ocasionales. Son incansables. El pensamiento, en cualquier manifestación extensa, es imposible sin ellas, pues en ellas consiste.».

Beste batzuetan, berriz, aurretik esan dena ukatzeko erabiltzen dira:

- *Zahartu egin zara. — Ezta gutxiagorik ere! Ni txabalitoa sentitzen naiz.*
- *Gorospa izan da bizkaitar txirrindularirik onena. — Adarrik ez!*
- *Laster ezkonduko haiz, ezta? — Bai, zera!*

Norbait agurtzerakoan, norberaren egoeraren berri emateko balio dute beste batzuek:

- *Zer moduz zaude? — Oso ondo: ni baino hobeto dagoena ez dago gaizki!*
- *Zer moduz zaude? — Oso ondo: hobeto, nahi ere ez!*
- *Zer moduz bizi haiz? — Espezial!*
- *Zer moduz zaude? — Ondo edo, ondo ustean.*
- *Zer, ondo? — Ondo edo, ondo nahian.*
- *Aspaldiko! Zer moduz zabilta? — Hainbestean!*
- *Zer moduz zabilta? — Ez hain ondo, denaren beharrean.*
- *Zer moduz zara? — Nahi eta ezinean, emakumea!*

Lokailu bidezko egitura adierazgarriak ere erabiltzen dira. Presuposizio-kontuetan, berezi-berezia da *ere*, eta indar berezia ematen dio gure jardunari: berehalako erantzunen, interjekzioen arlo honetan, beraz, aukera egokia ematen du gure erantzunei eta esamoldeei adierazgarritasuna emateko. Hona adibide batzuk:

- *Laster ezkonduko haiz, ezta? — Baita zera ere!*
- *Hartu genuen erabakia. — Bai azkar hartu ere!*
- *Egun on! — Bai zeuri ere!*
- *Egun on! — Baita zeuri ere!*
- *Ez zarete alferrik galduko zuek! — Ezta komeni ere!*

14. Barren

Azkenik, Elgoibarren berezkoa den *barren* lokailua izango dugu hizpide. Berezia delako gure herritarrontzat egin diot atal berezi bat. Egitura adierazgarrietan izaten du berezko lekua, eta primeran datorkigu Elgoibarkoa eta inguru hurbilekoa bakarrik den hitz honen erabileraren berri emateko. Hona adibide batzuk:

- *Eta Patxi? — Zein Patxi?! Adarrik ez duk jo behar: Patxi ni lanean nenbilenean hil zuan, barren!¹⁰*
- *Umeek ez daukate, barren, inon jolasean egiteko lekurik!*
- *Eta gero, plazara joaten direnean, «garestia» esaten dute [barazkia]. Ez, jauna! Haiek ez dakite, barren, zer den baserriko lana!*

¹⁰ *Barren* lokailua Elgoibarren eta inguruko herri gutxi batzuetan bakarrik erabiltzen da; *Euskal Gramatika. Lehen Urratsak III (Lokailuak)* liburukian (ikus 131 or.), Gramatika batzordeak aurkaritzakoen artean sailkatu zuen. Berez markatuak diren perpausetan erabiltzen da batez ere, harridura-edo adierazteko. Haren esanahia ulertzeko, *baina* edo *eta* erabil daitezke haren ordez, kasuan-kasuan. Elgoibarko Izarra euskara elkarteak astero argitaratzen duen euskarazko herri-astekariak ere *Barren* du izena.

Bibliografia

- Alcoba, Santiago (coord.) (1999): *La oralización*. Barcelona: Ariel.
- Amonarriz, Kike (1998): «Euskarazko umorea eta hizkera erakargarria», in Rodríguez, Fito & Esnaola, Imanol (ed.), *Euskara mintzatuaren erronkak*. Leioa: EHU.
- Beinhauer, Werner (1929): *El español coloquial*. Madrid. Gredos.
- Briz, Antonio (1996): *El español coloquial: situación y uso*. Madrid. Arco.
- Elxepuru, Juan Martin (2005): *Bergara aldeko hiztegia. Leintz-Aramaio, Oñati eta Eibar inguruko berbe-kin osotuta*, Bergarako Udala.
- Esnaola, P. (1987-1989): «Ahozko euskarari buruz», Euskararen Kutxa atalean argitaratutako artikulua-sorta, *HABE*.
- Euskaltzaindia (1990): *Euskal Gramatika. Lehen Urratsak III (Lokailuak)*. Bilbo.
- Garate, Gotzon (1988): *Erdarakadak*. Mensajero.
- (1998): *27.173 Atsotitzak. Refranes. Proverbs. Proverbia*. Bilbo. BBK.
- Garzia, Joxerra (1998): «Hedabideetako euskara mintzatua», in F. Rodríguez & I. Esnaola (arg.), *Euskara mintzatuaren erronkak*. Bilbo. UPV/EHU.
- (1999): *Gaur egungo bertsolarien baliabide poetiko-erretorikoak. Marko teorikoa eta aplikazio didaktikoa*. Leioa: EHU/doktore tesiak.
- (2008): *Jendaurrean hizlari. (Ahozko) Komunikazio gaitasuna lantzeko eskuliburua*. Alberdania.
- Ibarretxe, Iraide (2006): *Euskal onomatopeien hiztegia. Euskara-Ingelesera-Gaztelania* [Diccionario de onomatopeyas vascas. Euskara-Inglés-Español]. Donostia. Gaiak.
- Izagirre, Koldo (1981): *Euskal lokuzioak*. Hordago.
- Laspiur, Imanol (1999): *Eibarrera jatorraren bidetik*. Eibar. *Eta kitto!* Euskara elkarteak.
- Lassalletta, M.C. (1974): *Aportaciones al estudio del lenguaje coloquial galdosiano, etc.* Insula Madrid.
- León, Jabi (2004): *Elgoibartarren esanetan*. Elgoibar. Elgoibarko Udala.
- Makazaga, Jesus Mari (2006): «Non dago ahazko jardunaren adierazgarritasuna? Erakusgarri batzuk Elgoibarko euskaratik», in B. Fernández & I. Laka (arg.), *Andolin gogoan. Essays in honour of Professor Eguzkitza*, 663-678. UPV/EHU.
- (2008): «Ahozko jarduna komunikazioaren lagungarri: ekarpen bat ahazkoaren estrategia komunikatiboez», in (X. Artiagoitia & J.A. Lakarra arg.) *Gramatika jaietan: Patxi Goenagaren omenez*, 545-560. UPV/EHU.
- (2010): *Elgoibarko ahazko euskara*. IKER-25. Euskaltzaindia-UPV/EHU.
- Mendebalde Euskera Alkartia (2003): *Ahazkotasuna aztergai*. Bilbo. BBK.
- Mokoroa, Justo M. (1990): *Ortik eta emendik. Repertorio de locuciones del habla popular vasca*. Bilbao: Labayru / Eusko Jaurlaritza. Etor.
- Ong, W. J. (1987): *Oralidad y escritura*. Mexico. Edit. Fondo de Cultura Económica.
- Osa, Eusebio (1990): *Euskararen hitzordena, komunikazio zereginen arauera*. Leioa. EHU-ren Argitarapen Zerbitzua.
- Plazaola, Estepan, Agirregabiria, Karmele & Bikuña, Karmen (2005): *Ahazko jarduna hobetzeko proposamenak*. Badihardugu Euskara Elkarteak.
- Salaburu, Pello (1984): *Hizkuntza teoria eta Baztango euskalkia* (2 liburuki). Leioa. EHU.
- (2005): *Baztango mintzoa: gramatika eta hiztegia*. Nafarroako Gobernua-Euskaltzaindia.
- San Martin, Juan (1960): *Zirikadak*. Berrargitaratua: Elkar (1984).
- (1967): *Ibilarixanak*. Ego ibarra (1993). Kutxa.

Sobre el orden básico de palabras en euskera: un nuevo estudio de corpus¹

LUIS PASTOR SANTAMARÍA

Universidad del País Vasco / Euskal Herriko Unibertsitatea (UPV/EHU)

Abstract

Basque is classified as an SOV language, like Japanese or Korean. However, unlike these OV languages, which have rigid word order, Basque has free word order and it allows postverbal arguments. Some corpus studies have determined the basic word order is SOV (De Rijk, 1969; Aldezabal *et al.*, 2003) whereas others have claimed it to be SVO (Hidalgo, 1995a; Aske, 1997). I conducted a new corpus study of Basque to disentangle whether the most frequent word order (i.e., the basic word order) is SOV or SVO. Unlike previous Basque corpus studies, this new corpus study analyzes and compares different sources (press, magazine, books and TV scripts) in order to have a heterogeneous corpus. The results indicate that SOV is the most frequent word order in Basque. Furthermore, I compared the data of all corpus studies in Basque (including the data of this chapter) and, once again, SOV emerged as the most frequent word order. All together, these results show that SOV is the basic word order in Basque, contrary to claims by Hidalgo (1995a) and Aske (1997), who defend that SVO is the basic word order. I present a critical discussion of these two studies.

El orden básico de palabras es el orden que aparece en una oración transitiva declarativa con todos los argumentos expresados fonológicamente (Greenberg, 1963). Greenberg (1963) fue el pionero en utilizar el orden básico de palabras como criterio principal para clasificar tipológicamente las lenguas naturales. En su trabajo, titulado *Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements*, observó que de las posibilidades lógicas de la combinación del sujeto (S), el objeto (O) y el verbo (V), que son seis, los órdenes que se encuentran predominantemente en las lenguas del mundo son SOV, SVO y VSO. Posteriormente, Dryer (2013b) ha confirmado, mediante un corpus de 1377 lenguas, la prevalencia de estos tres órdenes principales observados por Greenberg (1963): los órdenes SOV (41%) y SVO (35%) son los más frecuentes, seguidos con a mucha distancia por el orden VSO (7%) (TABLA 1).

Tabla 1

Porcentajes de los órdenes básicos de palabras en el WALS (Dryer, 2013b).

Entre paréntesis aparecen las frecuencias absolutas

SOV		SVO		VSO		VOS		OVS		OSV		Sin orden		TOTAL	
41%	(565)	35,4%	(488)	6,9%	(95)	1,8%	(25)	0,8%	(11)	0,3%	(4)	13,7%	(189)	100%	(1377)

¹ Este trabajo ha recibido financiación del Gobierno Vasco (IT1169-19), MINECO (FFI2015-64183-P) y MICINN (FPI-2012 EEBB-I-13-06977).

El orden básico de palabras se asocia con dos propiedades: es el orden más frecuente y el menos marcado morfológica y pragmáticamente (Hawkins, 1983; Comrie, 1989; Dryer, 1995; Whaley, 1997; Dryer, 2013a, entre otros), aunque como afirma Greenberg (1966) la propiedad de ser el orden menos marcado está correlacionado con ser el orden más frecuente. Por tanto, podemos decir que ser el orden más frecuente es la propiedad principal del orden básico de palabras en la oración. Así, de todos los órdenes posibles que puede tener una oración en una lengua, el más frecuente será su orden básico de palabras. Por ejemplo, el castellano puede hacer uso de diferentes órdenes de palabras dado que tiene cierta libertad en la distribución de los argumentos; sin embargo, de entre todos ellos el orden básico SVO es significativamente el más frecuente (92%) (López, 1997).

En este artículo trato sobre la frecuencia del orden de palabras de las oraciones en euskera. En la siguiente sección 1.1, presento y discuto los diferentes estudios de corpus que he llevado a cabo en euskera sobre las frecuencias de los órdenes de palabras y trato de explicar por qué dichos estudios observan diferencias en los resultados que obtienen para la distribución de los órdenes de palabras. En la sección 1.2, presento un nuevo estudio de corpus en euskera con la intención de disipar la divergencia de los resultados de los estudios de corpus anteriores, y analizo estadísticamente las frecuencias de los órdenes de palabras. El artículo termina con la discusión (sección 1.3) de los resultados y las conclusiones (sección 1.4).

1. Revisión de los estudios de corpus sobre el orden de palabras en euskera

Como he comentado en la introducción, el orden básico o canónico se caracteriza por ser el orden de palabras que se usa con mayor frecuencia. Si el euskera es una lengua OV, el orden de palabras SOV debe de ser el más frecuente. Sin embargo, los estudios de corpus existentes reflejan frecuencias opuestas sobre el orden de constituyentes en las oraciones: unos muestran que SOV es el orden más frecuente (De Rijk, 1969; Aldezabal *et al.*, 2003); y otros que no, porque sería aventajado por el orden SVO (Hidalgo, 1995a, 1995b; Aske, 1997).

De Rijk (1969) es el primero argüir que el orden básico de palabras en euskera es SOV, y entre los criterios que aporta, uno es ser el orden más frecuentemente usado. Para examinar la frecuencia de los órdenes de palabras posibles en euskera De Rijk (1969) lleva a cabo un estudio pionero de corpus escrito en el que solo tiene en cuenta las oraciones declarativas transitivas en las que aparecen expresados el sujeto (S), el objeto (O) y el verbo (V). Del conjunto total de oraciones transitivas, De Rijk (1969) deja a un lado las transitivas interrogativas y negativas, las transitivas que no tienen el verbo conjugado, y transitivas en las que el sujeto o el objeto son oraciones subordinadas (salvo cuando la subordinada es una oración relativa que modifica el sujeto o el objeto), y las transitivas directas como «*Bihar jatera joango naiz*» *esan zuen mutilak*» [*Mañana iré a comer*] *dijo el chico* e indirectas como «*Mutilak esan zuen bihar joango zela jatera*» [*El chico dijo que mañana iría a comer*]. El corpus está compuesto por tres muestras: La muestra I consiste en textos folklóricos recogidos por José Miguel Barandiaran y contados por hablantes de Gipuzkoa y Bizkaia entre los años 1920-1936; la muestra II está compuesta por pequeñas obras teatrales escritas por Nemesio Echániz; y la muestra III la componen las obras *Mateo Falcone* y *Oillasko Iturri* de Mérimée (traducidas del francés por Nemesio Echániz). Los resultados de su cómputo muestran que de los seis posibles órdenes de palabras que se pueden utilizar en euskera SOV es el orden predominante en las tres muestras (57% en total. Desglosado por muestra: I: 66%; II: 44%; y III: 61%), seguido por SVO (Total: 30% = I: 23%; II: 37%; y III: 31%) (TABLA 2, GRÁFICO 1).

Tabla 2

Frecuencias absolutas de los órdenes de palabras en euskera
en el corpus de De Rijk (1969)

	I	II	III	TOTAL
SOV	138	80	41	259
SVO	48	67	21	136
OVS	11	17	3	31
OSV	5	13	1	19
VSO	6	4	1	11
VOS	1	2	0	3
TOTAL	209	183	67	459

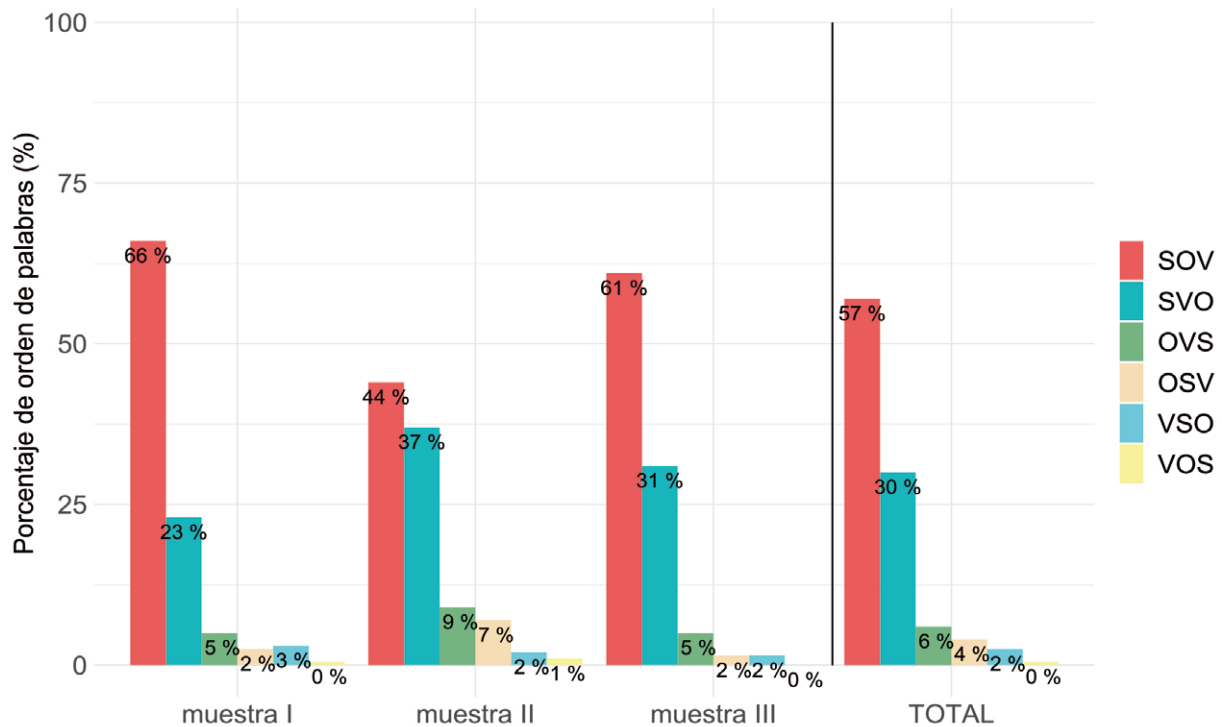


GRÁFICO 1

**Porcentajes de los órdenes de palabras en euskera en el corpus de De Rijk (1969),
por muestras (I, II y III) y en total**

De estos datos, De Rijk (1969) concluye que, pese a la libertad de orden de palabras en la oración, el euskera sí tiene un orden básico o canónico: el orden SOV. Aldezabal *et al.* (2003) encuentran la misma distribución encontrada por De Rijk (1969) mediante un estudio de corpus de mayor tamaño. Este corpus está compuesto por 5.639 oraciones de artículos periodísticos del periódico en euskera *Euskaldunon Egunkaria*, entre los años 1999-2000. Las oraciones han sido eti-

quetadas de manera automática mediante un analizador sintáctico computacional según el orden del sujeto (S), el objeto (O) y el verbo (V) en las oraciones. De total de oraciones etiquetadas, solo en 512 oraciones aparecen los tres argumentos expresados. Los resultados muestran que SOV es el orden de palabras que aparece con mayor frecuencia (56,8%), seguido de SVO (14,8%) y OVS (13,8%) (TABLA 3, GRÁFICO 2).

Tabla 3

Frecuencias absolutas de los órdenes de palabras en euskera
en el corpus de Aldezabal *et al.* (2003)

	TOTAL
SOV	291
SVO	76
OVS	71
OSV	51
VSO	6
VOS	17
TOTAL	512

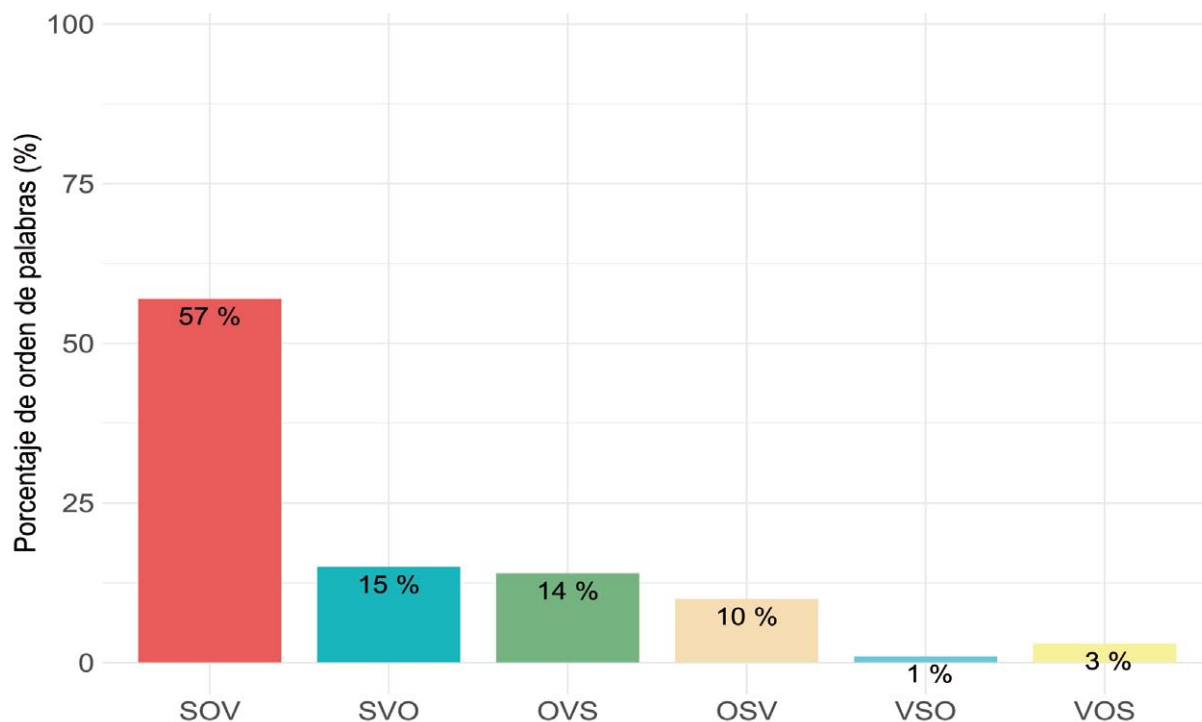


GRÁFICO 2

Porcentajes de los órdenes de palabras en euskera
en el corpus de Aldezabal *et al.* (2003)

Como vemos, la distribución obtenida por Aldezabal *et al.* (2003) es similar a la de De Rijk (1969): SOV es el orden de palabras más frecuente en euskera y, por tanto, es el orden básico de palabras.

Hidalgo (1995a, 1995b) y Aske (1997), sin embargo, sostienen que el orden básico de palabras en euskera es SVO. Hidalgo (1995a, 1995b) lleva a cabo un estudio con un corpus escrito de euskera de mayor tamaño que los anteriores. Aunque no detalla el número total de oraciones que componen su corpus, sí puede calcularse el total de oraciones que usa para la distribución de los órdenes de palabras. Su corpus consta de oraciones obtenidas de 19 textos de diferentes autores de entre los siglos XVI-XX. Para el cómputo de oraciones Hidalgo excluyó menos tipos de oraciones de las requeridas. Solamente excluyó las oraciones sin verbo, con verbo no conjugado, sin complementos, oraciones interrogativas, imperativas y negativas. No excluyó las oraciones ditransitivas e intransitivas, que constituyen una parte significativa de las oraciones a computar. Sus resultados muestran que SVO (54%) es el orden de palabras más frecuencia en euskera, seguido por el orden SOV (24%) (TABLA 4, GRÁFICO 3). Hidalgo (1995b), basándose en los datos, concluye que es un error decir que el euskera sea una lengua SOV.

Tabla 4

Frecuencias absolutas de los órdenes de palabras en euskera
sumando todos los corpus de Hidalgo (1995a, 1995b)

	TOTAL
SOV	482
SVO	1.081
OVS	204
OSV	70
VSO	136
VOS	40
TOTAL	2.013

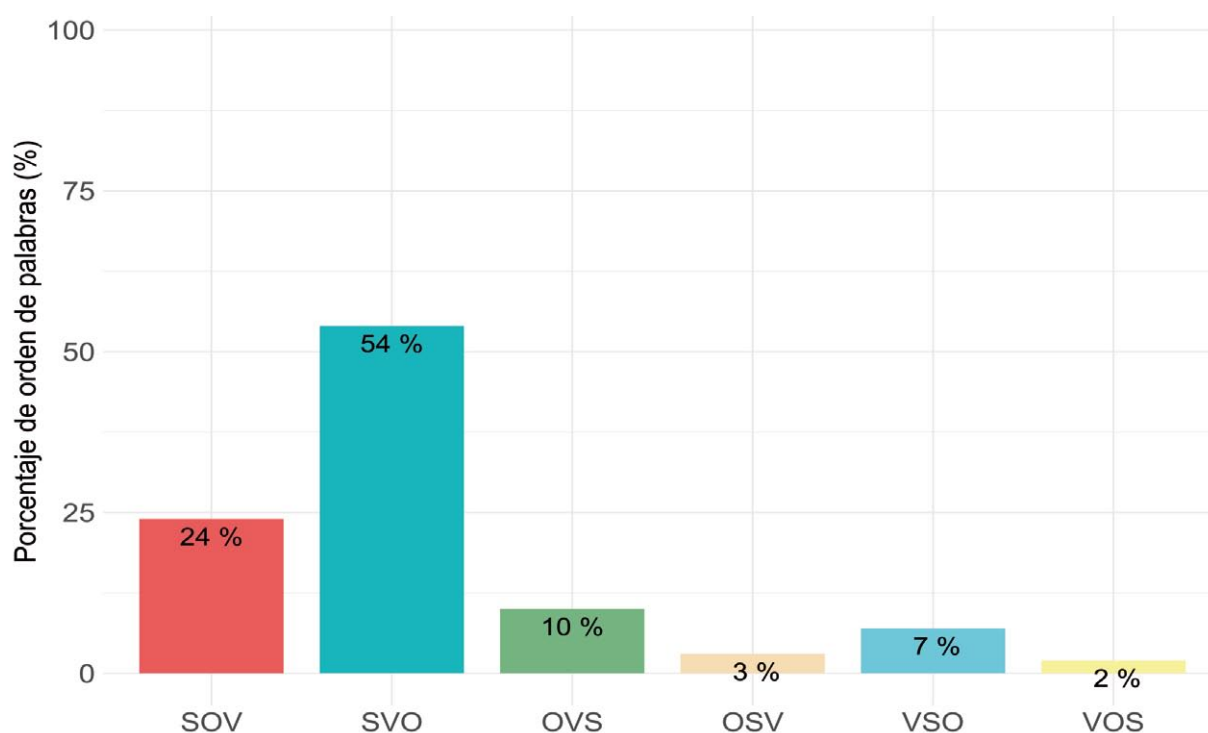


GRÁFICO 3

Porcentajes de los órdenes de palabras en euskera en el corpus de Hidalgo (1995a, 1995b)

Al igual que Hidalgo (1995a, 1995b), Aske (1997) encuentra que el orden SVO es el más frecuente en euskera. Mediante dos estudios de corpus, uno oral y otro escrito, compara únicamente la frecuencia de los órdenes de palabras SOV, SVO, VSO y VOS en euskera. No reporta la frecuencia de los órdenes OSV y OVS. El corpus oral consta de relatos de 46 participantes (niños y adultos) de dos películas mudas *Pearl Film* y *Modern Times* y el corpus escrito constaba del primer artículo de las novelas en euskera *Behi euskaldun baten memoriak* y *Kuba triste dago*. Aske concluye que, en conjunto, SVO es el orden de palabras más frecuente (61%) (TABLA 5, GRÁFICO 4). Separándolos por tipo, el corpus escrito muestra esta misma predominación del orden SVO (66%), pero en el corpus oral no hay diferencias entre los órdenes SOV (50%) y SVO (50%). Aske (1997) sugiere que esta preferencia por SVO puede deberse a la influencia del castellano de los participantes.

Tabla 5

Frecuencias absolutas de los órdenes de palabras en euskera en los corpus oral y escrito (y la suma de estos) de Aske (1997)

	Corpus oral	corpus escrito	TOTAL
SOV	12	14	26
SVO	12	46	57
VSO	0	9	9
VOS	0	1	1
TOTAL	24	70	94

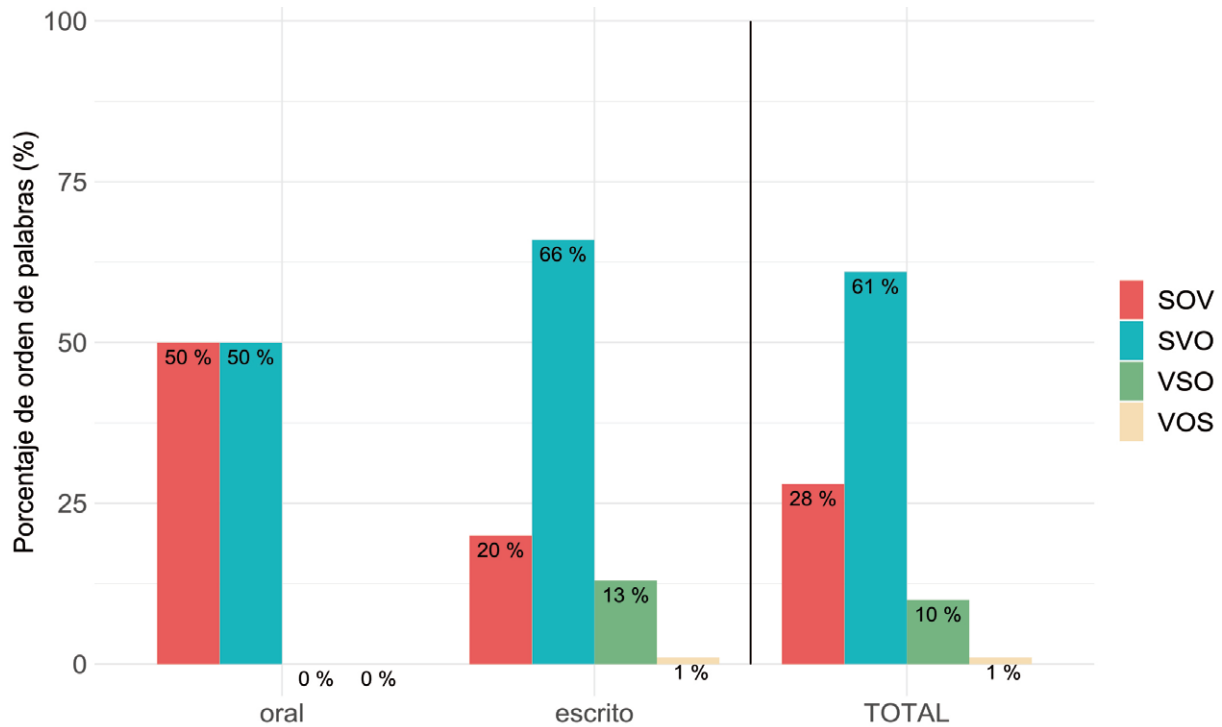


GRÁFICO 4

Porcentajes de los órdenes de palabras en euskera en los corpus oral, escrito y la suma de estos (Total) de Aske (1997)

Así pues, las distribuciones observadas por Hidalgo (1995a, 1995b) y Aske (1997) sugerirían que SVO es el orden básico en euskera, en contra de la distribución encontrada por De Rijk (1969) y Aldezabal *et al.* (2003), quienes observan que SOV es el orden más frecuente. Esta discrepancia en las frecuencias encontradas se debe a la metodología utilizada por Hidalgo (1995a, 1995b) y Aske (1997) frente a los otros dos estudios. Mientras que De Rijk (1969) y Aldezabal *et al.* (2003) etiquetan correctamente como objeto (O) solo los objetos directos, Hidalgo (1995a, 1995b) y Aske (1997) incluyen también en esta categoría cualquier otro tipo de complemento verbal (e.g., objetos indirectos, complementos circunstanciales, etc.), desvirtuando así el cómputo de la categoría O. Esta cuestión la abordo con más detalle en la discusión (sección 1.3).

A continuación, presento el diseño y los resultados de un nuevo estudio de corpus escrito en euskera que he llevado a cabo. Este nuevo corpus es más heterogéneo que los anteriores, ya que incluye textos de diferentes géneros textuales. A su vez, comparo los resultados obtenidos con los de los estudios previos y hago un análisis de todos los datos para determinar cuál es la frecuencia de uso de los posibles órdenes de palabras en euskera.

2. Estudio de corpus: orden de palabras en euskera

En el presente estudio de corpus examino la frecuencia de distribución de los seis posibles órdenes de palabras en euskera. Los propósitos de este nuevo estudio de corpus en euskera son dos: por un lado, usar un corpus amplio, similar en tamaño a los de Hidalgo (1995a, 1995b) y Aldezabal *et al.* (2003), pero que incluye diferentes géneros para ser lo más representativo posible

de los diferentes estilos discursivos en euskera; por otro lado, analizar las frecuencias encontradas en el corpus mediante análisis estadísticos para testear la significatividad de las diferencias observadas. Los estudios de corpus en euskera previos utilizan estadísticas descriptivas básicas, i.e., la frecuencia de aparición de un orden de palabras en el corpus; pero no analizan si las frecuencias observadas son simplemente debidas al azar o no.

2.1. *Materiales*

El corpus escrito de euskera utilizado en este artículo consta de 4000 oraciones. He obtenido las oraciones del corpus *EPG-Ereduzko Prosa Gaur* (Sarasola, Salaburu, Landa y Zabaleta, 2009) y he seleccionado diferentes géneros para tener una muestra heterogénea. El criterio de selección ha sido el siguiente: 1750 oraciones del periódico *Berrria*, 1300 oraciones de diferentes libros, 300 oraciones de los guiones de la serie de televisión *Goenkale*. A estas he añadido 600 oraciones de la revista de divulgación *Elhuyar* y 50 oraciones de la revista científica *Uztaro*. Con la intención de tener un corpus aún más heterogéneo, he obtenido las oraciones de diferentes subapartados en cada género (salvo las oraciones de los guiones de la serie televisiva *Goenkale* y la revista científica *Uztaro*), y con un número de oraciones igual en cada una de ellos. En el periódico de lengua vasca *Berrria* he utilizado las secciones de Economía, Sociedad, Mundo, Deportes, Cultura, Política y Nacional (250 oraciones $\times$ 7 secciones = 1.750 oraciones). En libros, he seleccionado libros de cuatro géneros diferentes: Comedia, Misterio, Histórica y Ensayo (No-ficción) (325 oraciones $\times$ 4 géneros = 1.300 oraciones). Y en la revista divulgativa *Elhuyar* las secciones Historia, Cultura, Naturaleza, Salud, Tecnología y Ciencia (100 oraciones $\times$ 6 secciones = 600 oraciones).

2.2. *Procedimiento*

He etiquetado las oraciones transitivas manualmente y las he clasificado según el orden de palabras, i.e., teniendo en cuenta el orden lineal en el que aparecen el sujeto (S), el objeto directo (O) y el verbo (V):

- | | |
|---|-----------------|
| (1) a. <u>Nik</u> _[S] <u>albaniarra</u> _[O] <u>defenditu</u> <u>ditut</u> _[V]
«Yo he defendido a los albaneses» | SOV [libros] |
| b. <u>Gaizka</u> _[S] <u>irekitzen</u> <u>du</u> _[V] <u>atea</u> _[O]
«Gaizka abre la puerta» | SVO [guiones] |
| c. <u>Lana</u> _[O] <u>sei ikerketa-talde</u> _[S] <u>egin dute</u> _[V]
«El trabajo lo han hecho seis grupos de investigación» | OSV [revista] |
| d. <u>Lau txanda</u> _[O] <u>aurreikusi</u> <u>ditu</u> _[V] <u>batzordeak</u> _[S]
«Cuatro turnos ha previsto el comité» | OVS [periódico] |
| e. <u>Berehala</u> <u>irentsi</u> <u>zuen</u> _[V] <u>amua</u> _[O] <u>katxaloteak</u> _[S]
«En seguida se ha tragado el anzuelo el cachalote» | VOS [libros] |
| f. <u>Eugin</u> <u>jaso</u> <u>zuen</u> _[V] <u>Azkue</u> _[S] <u>kanta</u> <u>hau</u> _[O]
«En Eugi recogió Azkue esta canción» | VSO [periódico] |

Del total de oraciones que componen el corpus, solo he tenido en cuenta para el análisis estadístico las 1.054 oraciones transitivas declarativas, descartando las oraciones intransitivas, las di-

transitivas y las transitivas negativas, interrogativas, con subordinadas como objeto (objetos CP) y aquellas con argumentos omitidos (NP omitidos) (TABLA 6).

Tabla 6

Distribución del tipo de oraciones del corpus escrito de euskera

Intransitivas	Transitivas					Ditransitivas	TOTAL
989	declarativas	negativas	interrogativas	objetos CP	NP omitidos	122	4.000
	1.054	80	21	233	1.501		

A la hora de analizar los datos del corpus he usado los análisis estadísticos prueba de bondad de ajuste chi-cuadrado (*chi-square goodness of fit test*) y el modelo de regresión logística multinomial. La prueba de bondad de ajuste chi-cuadrado la he usado para analizar si la distribución de la frecuencia de los órdenes de palabras en el corpus se debe al azar o no. La regresión logística multinomial, por su parte, la he utilizado para analizar si el tipo de género del corpus influye en la frecuencia de uso de los diferentes órdenes de palabras. Los análisis estadísticos los he computado mediante el programa estadístico R (R Core Team, 2017) y usando el paquete *polytomous* (Arppe, 2013). He tomado como nivel de referencia la media de las medias de los cuatro géneros (periódico, libros, revistas y guiones), dado que el *intercept* es una media no ponderada, i.e., las variables tienen diferentes frecuencias. Los resultados los he considerado significativos a un nivel $p < .05$. Los gráficos han sido realizados con el paquete *ggplot2* (Wickham, 2009).

2.3. Resultados

La TABLA 7 muestra la clasificación de las 1054 oraciones transitivas que componen el corpus, según el orden de palabras y el tipo de género. De acuerdo con la prueba de bondad de ajuste chi-cuadrado (*chi-square goodness of fit test*) la distribución de las frecuencias de los seis posibles órdenes de palabras en euskera no es idéntica, i.e., no son igual de frecuentes ($\chi^2 (5, N = 1054) = 1249.9, p < .001, V = 0.48$): el orden que aparece significativamente con mayor frecuencia es SOV (52%), doblando la frecuencia de SVO (26%), que es el segundo orden más frecuente. Esta misma frecuencia mayoritaria de SOV se observa también en cada género: periódico (54%: $\chi^2 (5, N = 582) = 726.16, p < .001, V = 0.49$), libros (41%: $\chi^2 (5, N = 232) = 233.98, p < .001, V = 0.44$), revistas (50%: $\chi^2 (5, N = 151) = 175.42, p < .001, V = 0.48$) y guiones (69%: $\chi^2 (5, N = 89) = 181.81, p < .001, V = 0.63$).

TABLA 7

Distribución en euskera de las oraciones transitivas según el orden de palabras y el tipo de género.

Entre paréntesis aparecen las frecuencias absolutas

	periódico		libros		revista		guiones		TOTAL	
SOV	54%	(314)	40%	(94)	50%	(76)	69%	(61)	52%	(545)
SVO	23%	(133)	38%	(88)	26%	(39)	15%	(13)	26%	(273)
OSV	17%	(101)	16%	(37)	20%	(30)	12%	(11)	17%	(179)
OVS	3%	(15)	1%	(2)	2%	(4)	2%	(2)	2%	(23)
VSO	2%	(10)	3%	(7)	1%	(1)	1%	(1)	2%	(19)
VOS	1%	(9)	2%	(4)	1%	(1)	1%	(1)	1%	(15)
TOTAL	100%	(582)	100%	(232)	100%	(151)	100%	(89)	100%	(1054)

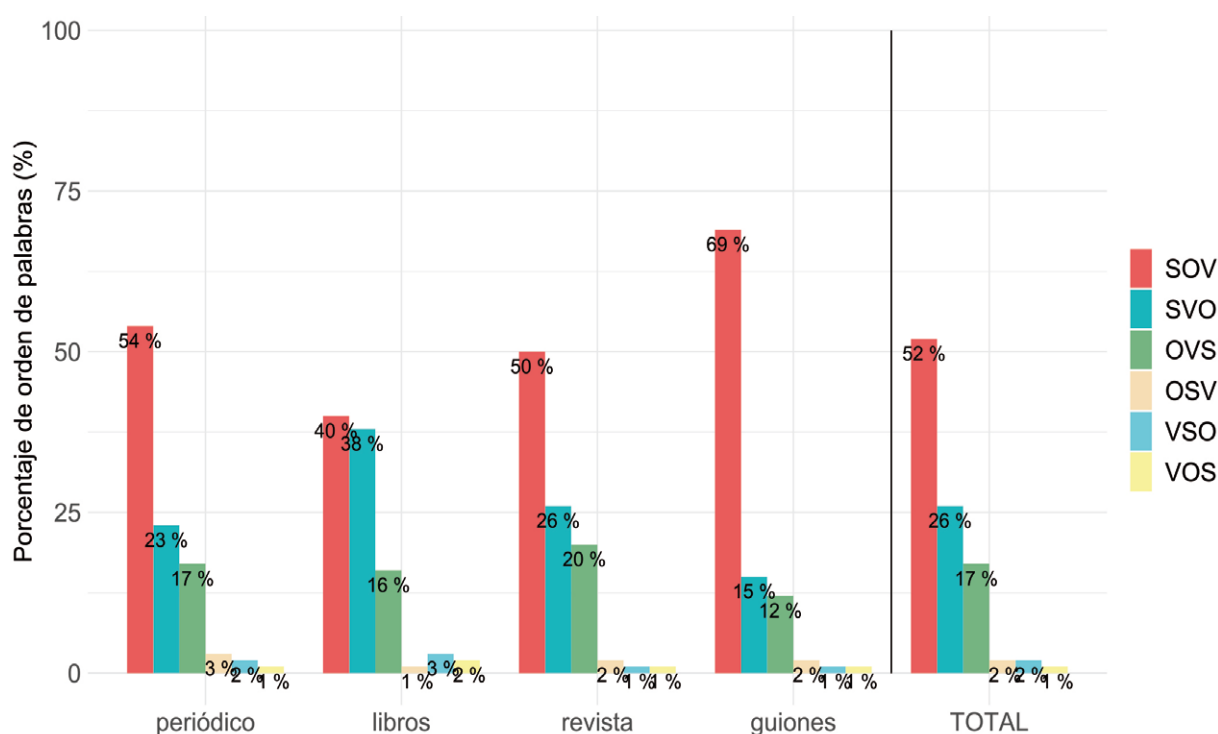


GRÁFICO 5

**Porcentajes de los órdenes de palabras en euskera
en los diferentes géneros (periódico, libros, revista, guiones) y la suma de estos (Total)**

Como vemos en el GRÁFICO 5, el orden SOV es el más frecuente en general y en todos los géneros; pero su frecuencia es mayor en los guiones (69%) que en el resto de géneros (periódico: 54%; libros: 41%; revistas: 50%). Para analizar si esta diferencia de frecuencia de SOV entre los cuatro géneros es significativa he llevado a cabo una prueba chi-cuadrado (*Pearson's chi-square test*), comparando su frecuencia con la de SVO en los cuatro géneros, que es el siguiente orden

más frecuente. Dado que los subcorpus de cada género son de diferente tamaño no se refleja con precisión las frecuencias relativas de ambos órdenes en cada género; por ello, he normalizado la frecuencia de ambos órdenes en cada género y he calculado sus frecuencias por 1.000. La prueba chi-cuadrado muestra que la frecuencia de uso de SOV es significativamente similar en cada género [χ^2 (3, N = 4000) = 220.91, $p < .001$, $V = 0.24$]. Aún así, el modelo de regresión logística multinomial muestra que el género guiones influye significativamente en que SOV se utilice con mayor frecuencia comparado con el resto de géneros [$\beta = 0.6371$, $p < .001$, odds ratio = 1.89]: la probabilidad de uso de SOV es 1,9 veces mayor en los guiones que en el resto de géneros. El orden SVO, por el contrario, se ve favorecido por el género libros [$\beta = 0.64$, $p < .001$, odds ratio = 1.89], donde es 1,9 veces más probable que se use con mayor frecuencia que en el resto de géneros (TABLA 8).

Tabla 8

Resultados del modelo de regresión logística para el orden de palabras en euskera según el tipo de género.
El Intercept es la media de las medias de los cuatro géneros (periódico, libros, revistas, guiones)

Orden de palabras por género-Coeficientes de <i>estimate</i> :						
	SOV	SVO	OVS	OSV	VSO	VOS
(Intercept)	(0.1416)	-1.132	-1.644	-3.939	-4.251	-4.421
periódico	(0.01682)	(-0.08421)	(0.08331)	(0.3062)	(0.2046)	(0.2675)
libros	-0.5255	0.64	(-0.01802)	(-0.8064)	(0.781)	(0.3781)
revistas	(-0.1283)	(0.07753)	(0.2495)	(0.3344)	(-0.7595)	(-0.5895)
guiones	0.6371	-0.6333	(-0.3148)	(0.1658)	(-0.2262)	(-0.05617)
<i>p-value</i> de los coeficientes						
(Intercept)	0.078	0.001***	0.001***	0.001***	0.001***	0.001***
periódico	0.866	0.483	0.533	0.373	0.641	0.555
libros	0.001***	0.001***	0.914	0.164	0.092	0.472
revistas	0.361	0.636	0.167	0.468	0.344	0.465
guiones	0.001***	0.007**	0.212	0.776	0.779	0.945
R2.likelihood: 0.015						

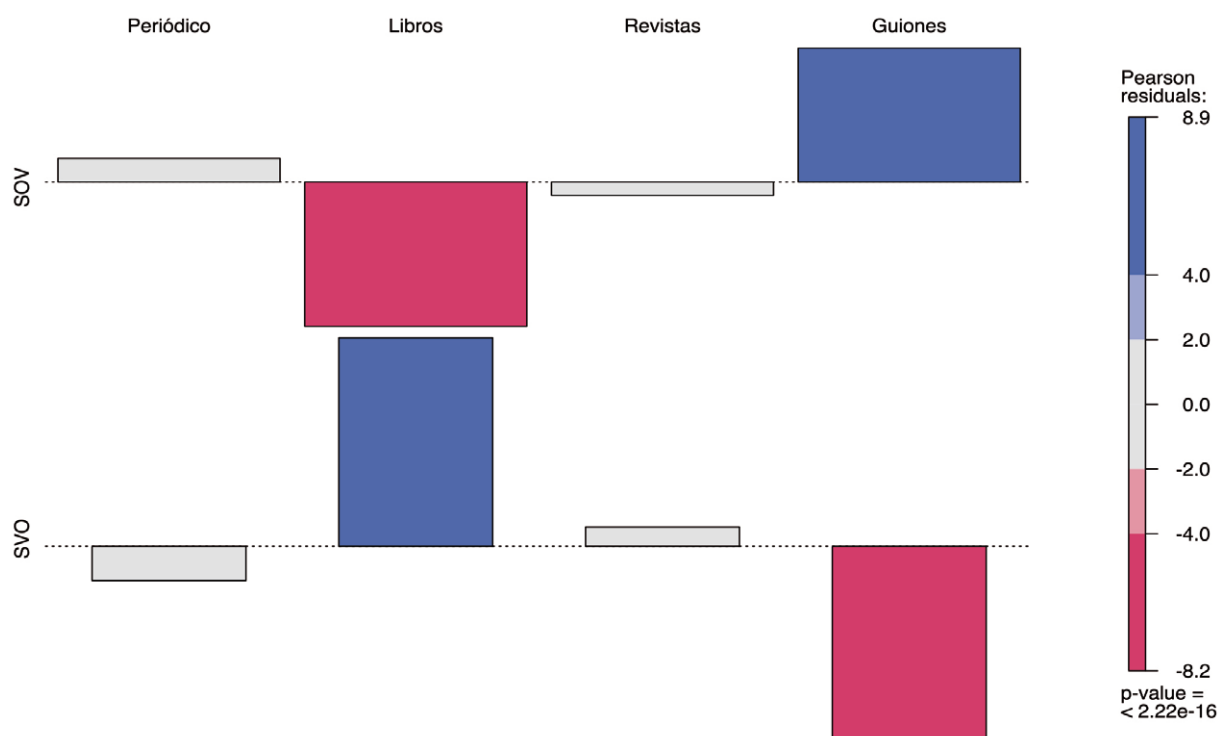


GRÁFICO 6

Usos de los SOV y SVO en los cuatro géneros: asociación de residuales.
 Las barras azules indican residuales positivos y las barras rojas residuales negativos

El GRÁFICO 6 enseña que en el género guiones el orden SOV es relativamente más frecuente al resto de los géneros, mientras que en el género libros sucede lo mismo con el orden SVO. De este dato podemos concluir que cuanto más se acerca un corpus al registro oral mayor es la frecuencia de SOV. Por el contrario, en el género libros que puede argüirse es el que más se aleja del registro oral encontramos mayor frecuencia del orden SVO.

2.3.1. COMPARACIÓN CON LOS ESTUDIOS DE CORPUS PREVIOS

El GRÁFICO 7 muestra la distribución de la frecuencia de uso de los órdenes de palabras en euskera en los cinco estudios de corpus existentes (De Rijk (1969), Hidalgo (1995a, 1995b), Aske (1997) y Aldezabal *et al.* (2003); ver sección 2.3) y los resultados del presente estudio de corpus (ver sección 1.2.3). Como puede observarse, en todos los corpus SOV y SVO son los órdenes que aparecen con mayor frecuencia: SOV en los corpus de De Rijk (1969), Aldezabal *et al.* (2003) y el presente estudio de corpus; y SVO en los corpus de Hidalgo (1995a, 1995b) y Aske (1997).

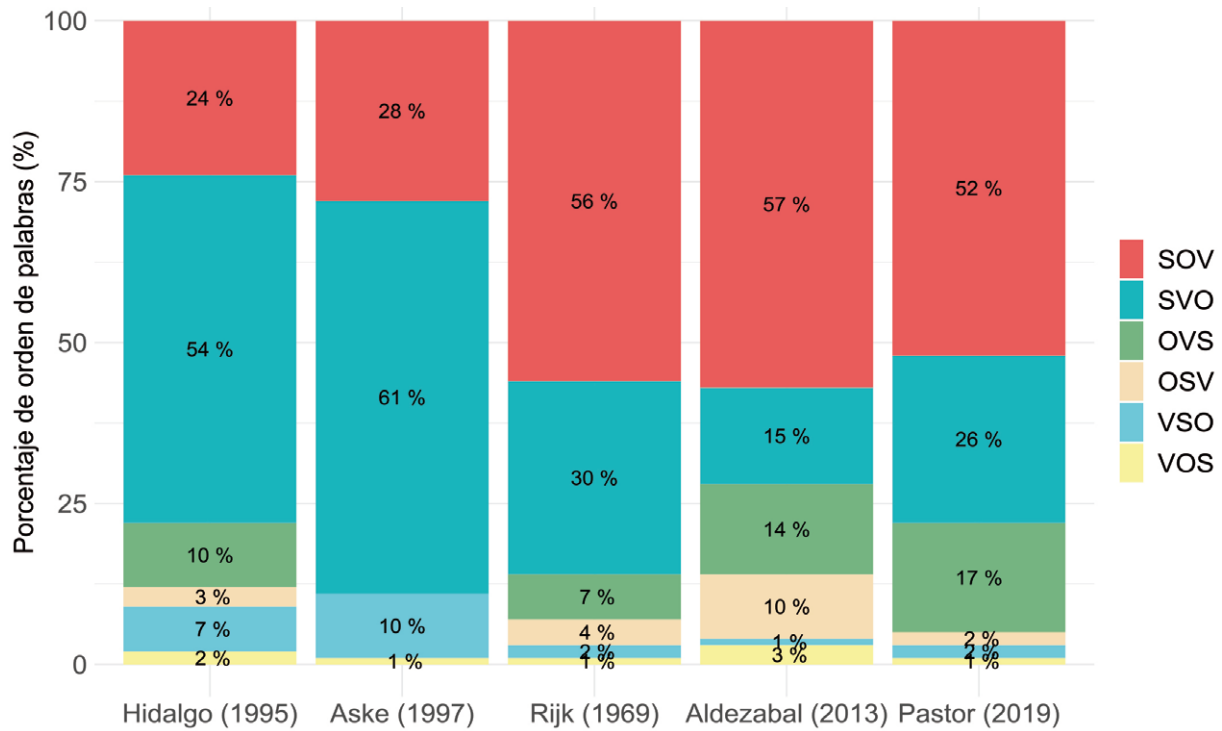


GRÁFICO 7

Distribución de la frecuencia de uso de los órdenes de palabras en euskera en los diferentes estudios de corpus (De Rijk, Hidalgo, Aske, Aldezabal, Pastor)

He llevado a cabo una regresión logística multinomial para observar si el uso de SOV y SVO se ve favorecido por los diferentes estudios de corpus (TABLA 9). Al hacer el análisis estadístico he omitido los datos de Aske (1997) ya que no reporta las frecuencias de los órdenes OVS y OSV. El análisis de regresión logística multinomial muestra que SOV se ve favorecido significativamente en los corpus de De Rijk (1969) [$\beta = 0.3974$, $p < .001$, odds ratio = 1.49], Aldezabal *et al.* (2003) [$\beta = 0.4141$, $p < .001$, odds ratio = 1.51] y el presente corpus [$\beta = 0.2053$, $p < .001$, odds ratio = 1.23]: la probabilidad de SOV es 1,4 veces mayor en estos estudios de corpus que en el de Hidalgo (1995a, 1995b). Por el contrario, el orden SVO se ve favorecido en el corpus de Hidalgo (1995a, 1995b) [$\beta = 1.026$, $p < .001$, odds ratio = 2.79] y su probabilidad de uso es 2,8 veces mayor que en el resto de corpus.

Tabla 9

Resultados del modelo de regresión logística para el orden de palabras en euskera según el autor del corpus.

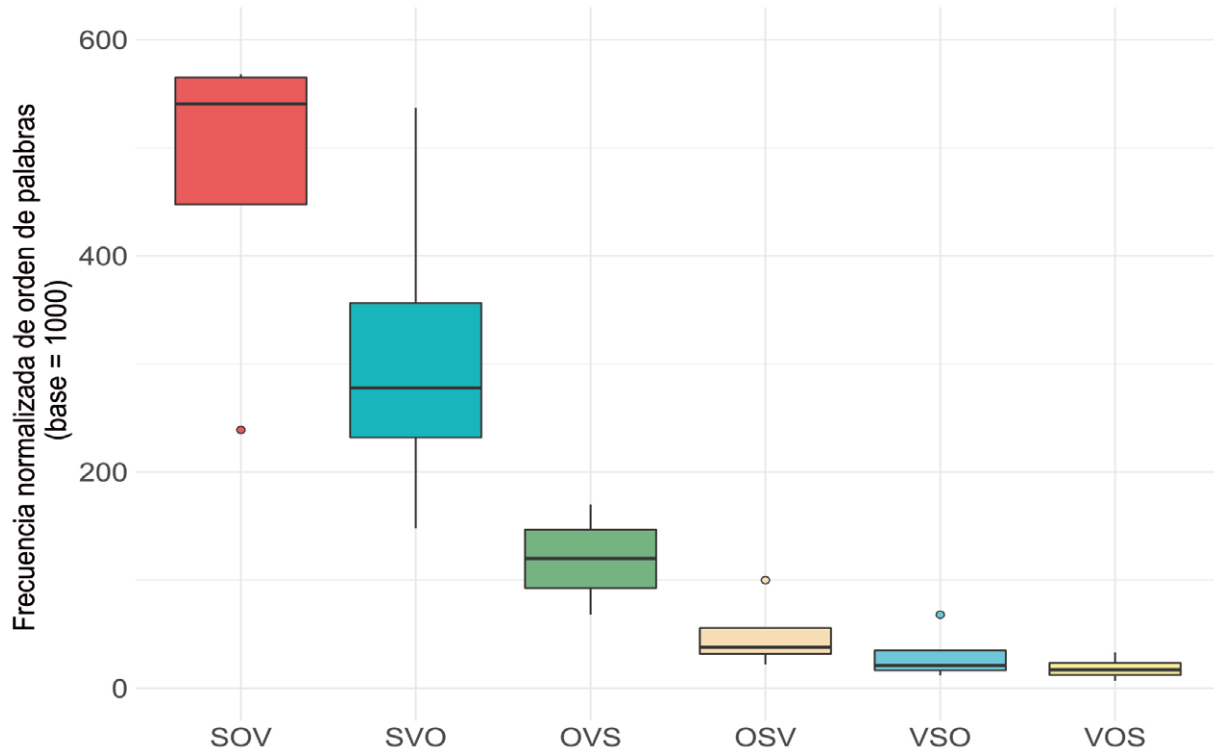
El Intercept es la media de las medias de los cuatro autores de corpus

(De Rijk, Hidalgo, Aldezabal, Pastor [esta tesis doctoral]).

Los coeficientes que aparecen entre paréntesis no son significativos (ver la tabla: p-value de los coeficientes)

Orden de palabras por autor del corpus-Coefficientes de <i>estimate</i> :						
	SOV	SVO	OVS	OSV	VSO	VOS
(Intercept)	-0.1389	-0.8778	-2.055	-3.118	-3.691	-4.133
De Rijk (1969)	0.3974	(0.01276)	-0.5697	(-0.02455)	(-0.01561)	-0.8907
Hidalgo (1995)	-1.017	1.026	(-0.1269)	(-0.2057)	1.067	(0.2347)
Aldezabal (2003)	0.4141	-0.8691	0.2291	0.9162	-0.7435	0.7618
Pastor [tesis]	0.2053	-0.1697	0.4675	-0.686	(-0.3074)	(-0.1058)
<i>p-value</i> de los coeficientes						
(Intercept)	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
De Rijk (1969)	0.001***	0.881	0.001***	0.897	0.952	0.045*
Hidalgo (1995)	0.001***	0.001***	0.120	0.103	0.001***	0.260
Aldezabal (2003)	0.001***	0.001***	0.037*	0.001***	0.021*	0.002**
Pastor [tesis]	0.001***	0.012*	0.001***	0.001***	0.156	0.677
R2.likelihood: 0.062						

Por último, he llevado a cabo un análisis con los datos de todos los corpus existentes en euskera para obtener una estimación de la frecuencia de los órdenes de palabras. La prueba de Kruskal-Wallis muestra que no hay diferencias significativas en la distribución de los diferentes órdenes de palabras, i.e., todos los corpus revelan una distribución similar en la frecuencia de uso de los órdenes de palabras ($H(3) = 0.201$, $p = .0977$). El orden SOV es el más frecuente, seguido, en orden decreciente, de SVO, OVS, OSV, VSO y VOS (GRÁFICO 8).



Distribución de la frecuencia de uso de los órdenes de palabras en euskera colapsando los estudios de corpus existentes (De Rijk, Hidalgo, Aldezabal, Pastor)

En resumen, los resultados de ese estudio de corpus revelan una preferencia por el orden básico SOV, siendo este el orden de palabras que aparece con mayor frecuencia en comparación con el resto de órdenes (SVO, OVS, OSV, VSO y VOS). Además, el orden SOV es también el orden más frecuente en cada uno de los géneros, aunque se ve más favorecido en los guiones de la serie de televisión *Goenkale*. En línea con esto, el análisis conjunto de los estudios de corpus existentes (De Rijk, 1969; Hidalgo, 1995a, 1995b; Aldezabal *et al.*, 2003) también muestran que SOV es el orden más frecuente en euskera.

3. Discusión

En este artículo he presentado un nuevo estudio de corpus escrito formado por diferentes géneros en euskera para explorar la distribución y frecuencia de uso de los seis posibles órdenes de palabras en euskera (SOV, SVO, OVS, OSV, VSO, VOS). Los resultados muestran que SOV es el orden dominante, i.e., el orden de palabras que se usa con mayor frecuencia, independientemente del género del texto. Estos resultados convergen con los estudios de corpus de De Rijk (1969) y Aldezabal *et al.* (2003), que también encuentran que SOV es el orden más frecuente en euskera. Sin embargo, estos resultados contrastan con los de Hidalgo (1995a, 1995b) y Aske (1997), que observan que SVO es el orden que aparece con mayor frecuencia seguido de SOV.

La discrepancia en la frecuencia entre los órdenes SOV y SVO entre los estudios de corpus de De Rijk (1969), Aldezabal *et al.* (2003) y esta tesis doctoral por un lado, y los de Hidalgo

(1995a, 1995b) y Aske (1997) por otro, se debe exclusivamente a los criterios de selección y etiquetado de las oraciones. Mientras que los primeros estudios solo tienen en cuenta oraciones transitivas declarativas afirmativas en las que ambos argumentos aparecen expresados, los segundos incluyen también oraciones intransitivas: «...*en vez de observar las oraciones que muestran externamente solamente el Sujeto, Objeto y Verbo conjugado, tener en cuenta junto a estas todas las oraciones principales declarativas que junto al verbo conjugado muestran cualquier otro tipo de complemento (llamémosle a cada uno «X»), y no solo los elementos S y O.*» (Hidalgo, 1995a:499). De esta manera, para Hidalgo (1995a, 1995b) y Aske (1997) las siguientes dos oraciones tienen el mismo orden de palabras (SOV): «*Sarak_[S] liburua_[O] irakurri du*» [Sara ha leído el libro] y «*Sara_[S] etxean_[O] gelditu da*» [Sara se ha quedado en casa]. Pero ambas oraciones no son iguales: la primera es una oración transitiva, porque el verbo «*irakurri*» pide dos argumentos: un sujeto (*Sarak*) y un objeto directo (*liburu*). La segunda oración, por el contrario, es una oración intransitiva porque el verbo «*gelditu*» solo pide un argumento —el sujeto (*Sara*)— de tal forma que «*etxean*» no es un objeto. Ejemplos similares se pueden encontrar en el corpus de Hidalgo (1995a, 1995b) (e.g., «*hori jin da goizan_[O]*» [Ese se ha ido a la mañana]). También etiquetan como oraciones transitivas, aunque son intransitivas, aquellas oraciones que tienen construcciones «nombre + *egin*» (e.g., «*baina nik alde egingo nuke*» [pero yo huiría] (Aske, 1997:927)). Las oraciones con construcciones formadas por «nombre + *egin*» llevan el sujeto en caso ergativo, como el sujeto de una oración transitiva; pero a pesar de ello, este tipo de construcción son inergativas, i.e., intransitivas (Laka, 1993).

En cuanto al etiquetaje de las oraciones transitivas, Hidalgo (1995a, 1995b) y Aske (1997) tampoco siguen el criterio marcado por De Rijk (1969), Aldezabal *et al.* (2003) y esta tesis doctoral, que solo computan aquellas oraciones transitivas que solo tienen objetos nominales (aunque pueden estar modificados por una oración relativa). Hidalgo (1995a, 1995b) y Aske (1997) incluyen dentro de su muestra oraciones transitivas en las que el objeto directo es una oración: «*Zuc erraiten duzu nic emaiten dudala_[O]*» [Tú dices que yo doy]. Estos autores también incluyen oraciones ditransitivas: «*Ama Virgiña erremediotakuak emanen digu osasona*» [La Virgen María de los remedios nos dará salud] (Hidalgo, 1995a:428) y «*berak adieraziko dizu aukerarik onena*» [él te indicará el mejor sitio] (Aske, 1997:926). Estas oraciones son ditransitivas, porque los verbos «*eman*» y «*adierazi*», en esos contextos, pide tres argumentos (sujeto, objeto directo e indirecto), aunque en ambas el objeto indirecto esté omitido.

Los estudios que hacen un etiquetado correcto de las oraciones de la muestra (De Rijk, 1969, Aldezabal *et al.*, 2003 y esta tesis) encuentran que el orden más frecuente es SOV; en los estudios con un etiquetado incorrecto de la muestras (Hidalgo, 1995a, 1995b y Aske, 1997), sin embargo, SVO es el orden más frecuente. Por tanto, podemos concluir y confirmar que el orden más frecuente en euskera es SOV y, por ende, el orden básico.

Además, este resultado converge con los de estudios experimentales de procesamiento oracional en euskera (Erdocia, Laka, Mestres-Misse y Rodríguez-Fornells, 2009; Erdocia, Laka y Rodríguez-Fornells, 2012; Ros, Santesteban, Fukumura y Laka, 2015). Estos estudios hallan que SOV es el orden de palabras que menor coste de procesamiento presenta en euskera. En comprensión, Erdocia *et al.* (2009) y Erdocia *et al.* (2012) observan que los hablantes de euskera muestran una negatividad N400, una positividad P600 y mayores tiempos de lectura en las oraciones con otros órdenes (SVO, OVS, OSV) en comparación con oraciones con el orden SOV. En producción, Ros *et al.* (2015) encuentran que los participantes prefieren producir el orden SOV frente al resto de órdenes, independientemente de la longitud del sujeto y el objeto. En línea con esto, es-

tudios con pacientes afásicos (Arantzeta, Webster, Laka, Martinez-Zabaleta y Howard, 2016 y Arantzeta *et al.*, 2017) también muestran una preferencia por el uso del orden SOV: las personas con afasia comprenden mejor las oraciones con el orden básico frente a las que no tiene el orden básico. Toda esta evidencia va en línea con los modelos *frequency-based accounts* (Hale, 2001; Levy, 2008), que predicen que las oraciones con órdenes más frecuentes son más fáciles de procesar que las menos frecuentes.

4. Conclusiones

El estudio de corpus de este artículo es el primero que analiza la frecuencia de los diferentes órdenes posibles en euskera en diferentes géneros. Puede considerarse también el estudio de corpus más amplio (en número de oraciones) en euskera etiquetado manualmente según la distribución del sujeto (S), objeto directo (O) y verbo (V) dados los errores de etiquetaje encontrados en el estudio de corpus de Hidalgo (1995a, 1995b), que era hasta la fecha el estudio de corpus de euskera más amplio conocido. Los resultados de este artículo sustentan la idea de que el orden de palabras más frecuente en euskera es SOV y por tanto es el orden básico, independientemente del tipo de texto y estilo.

Referencias

- Aldezabal, I., Aranzabe, M. J., Atutxa, A., Gojenola, K., Sarasola, K., y Zabala, I. (2003). *Hitz-hurrenkera-ren azterketa masiboa corpusean*.
- Arantzeta, M., Bastiaanse, R., Burchert, F., Wieling, M., Martinez-Zabaleta, M., y Laka, I. (2017). Eye-tracking the effect of word order in sentence comprehension in aphasia: evidence from Basque, a free word order ergative language. *Language, Cognition and Neuroscience*, 32(10), 1320-1343. doi: 10.1080/23273798.2017.1344715
- Arantzeta, M., Webster, J., Laka, I., Martinez-Zabaleta, M., y Howard, D. (2016). Cross-linguistic asymmetries in sentence comprehension deficits in bilingual Basque-Spanish aphasia. Evidence from eye-tracking and behavioural data. *Stem-, Spraak- en Taalpathologie*, 21(Suppl.), 28-31.
- Arppe, A. (2013). *polytomous: Polytomous logistic regression for fixed and mixed effects* [Computer software manual]. Disponible en <https://CRAN.R-project.org/package=polytomous>
- Aske, J. (1997). *Basque Word Order and Disorder Principles, Variation, and Prospects*. (PhD Dissertation), University of California.
- Comrie, B. (1989). *Language Universals and Linguistic Typology* (2 ed.). Chicago: The University of Chicago Press.
- De Rijk, R. P. G. (1969). Is Basque an SOV Language? *Fontes Linguae Vasconum*, 1, 319-351.
- Dryer, M. S. (1995). Frequency and pragmatically unmarked word order. In P. Downing & M. Noonan (Eds.), *Word order in discourse* (pp. 105-135). Amsterdam: John Benjamins.
- Dryer, M. S. (2013a). Determining Dominant Word Order. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Dryer, M. S. (2013b). Order of subject, object, and verb. In M. Haspelmath & M. S. Dryer (Eds.), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library.
- Erdocia, K., Laka, I., Mestres-Misse, A., y Rodriguez-Fornells, A. (2009). Syntactic complexity and ambiguity resolution in a free word order language: behavioral and electrophysiological evidences from Basque. *Brain and Language*, 109(1), 1-17. doi: 10.1016/j.bandl.2008.12.003

- Erdocia, K., Laka, I., y Rodríguez-Fornells, A. (2012). Processing Verb Medial Word Orders in a Verb Final Language. In M. Lamers & P. de Swart (Eds.), *Case, Word Order and Prominence* (Vol. 40). Netherlands: Springer.
- Greenberg, J. (1966). *Language Universals (with special reference to feature hierarchies)*. The Hague: Mouton.
- Greenberg, J. H. (1963). Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In J. H. Greenberg (Ed.), *Universals of Language* (pp. 40-70). Cambridge, MA: MIT Press.
- Hale, J. (2001). *A probabilistic earley parser as a psycholinguistic model*. Trabajo presentado en the Conference Namel, Conference Locationl.
- Hawkins, J. A. (1983). *Word Order Universals*. New York: Academic Press.
- Hidalgo, V. (1995a). *Hitzen ordena euskaraz*. (PhD Dissertation), Universidad del País Vasco / Euskal Herriko Unibertsitatea (UPV/EHU).
- Hidalgo, V. (1995b). Ohar estatistiko garrantzitsuak euskararen ordenaren inguru Euskara S.V.O.? *Fontes Linguae Vasconum*, 70, 401-420.
- Laka, I. (1993). Unergatives that assign ergative, unaccusatives that assign accusative. *MIT Working Papers in Linguistics*, 18(Papers on Case and Agreement), 149-172.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126-1177. doi: 10.1016/j.cognition.2007.05.006
- López, B. (1997). Aportaciones de la tipología lingüística a una gramática particular: el concepto de orden básico y su aplicación al castellano. *Verba. Anuario Galego de Filloxía*, 24, 45-82.
- R Core Team. (2017). R: A language and environment for statistical computing (Version 3.2.1). Vienna, Austria: R Foundation for Statistical Computing. Disponible en www.R-project.org
- Ros, I., Santesteban, M., Fukumura, K., y Laka, I. (2015). Aiming at shorter dependencies: the role of agreement morphology. *Language, Cognition and Neuroscience*, 30(9), 1156-1174. doi: 10.1080/23273798.2014.994009
- Sarasola, I., Salaburu, P., Landa, J., y Zabaleta, J. (2009). Ereduzko Prosa Gaur (EPG). Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU). Disponible en <http://www.ehu.eus/euskara-orria/euskara/ereduzkoa/>
- Whaley, L. J. (1997). *Introduction to typology: The unity and diversity of language*. Thousand Oaks, CA: SAGE.
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Euskararen konbinazio hiztegi baterako

IBON SARASOLA

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

JOSU LANDA

Ametzagaiña

Azken hamar-hamabost urtean, aitzinamendu ikusgarria izan du euskal hiztegi-intzak. Nahiko bat etorriko gara gehienak inpresio horretan. Euskara hizkuntza bezala deskribatzeko lexikoak barne hartzen duen arloan, gero eta deskripzio zehatzagoak eta zabalagoak daude, eta hiztegien erabiltzaileek gero eta xehetasun/zehaztasun gehiago dituzte eskura, hitz baten aplikazio zuzen edo egokia hautatzeko tenorean.

Arlo horretan gabiltzanak ez gara gutxi —lexikografia generikoaz gain terminologian edo fraseologian ere—, baina hemen aipagai dugun prozesu horretan hiru foku nabarmentzen ahalko genituzke: Euskaltzaindia, Elhuyar Fundazioa eta UPV/EHUko Euskara Institutua¹. Egindako ekarpenen norainokoa baloratzeko orduan, ez dira soilik aintzat hartu behar argitaratu izan diren hiztegiak berak, ikerketaren eremuan planteatu diren ildo berriei ere erreparatu behar diegu.

Esan liteke azken aldian nolabaiteko paradigma aldaketa bat egon dela lexikografian —ez bakarrik euskal lexikografian, fenomeno globalagoa dela esan liteke—, eta aldaketa horren ezaugarri nagusia corpusen erabilera da: baliabide informatikoei esker, eta gero eta zabalagoak diren corpusen osatzearen eta ustiatzearen ondorioz, hizkuntzaren benetako erabilera bera da ondoren datozen abstrakzioen abiapuntua. Garai batean, lexikografoek asmatu egiten zituzten hiztegi-tan ageri ziren adibide bakanak. Gaur egun, hitz bat egiazki nola erabilia izan den azaltzea da joera nagusia, normalean lexikografoak eskura izaten dituen corpusetan oinarrituta.

1. Konbinatoriaren inportantzia

Geroago konkrezio handiagoz azalduko dugu, baina hizkuntzaren sisteman bada alor deskriptibo bat (arauemailea ere bai, nahi bada), hitzen konbinatoriak soilik azal dezakeena. Eta konbinatoriak dakarrena begietara ekartzeko, modurik onena adibide zehatzak sistematizatzea eta zerrendatzea da. Garbi edukita, nolanahi ere, konbinazio hiztegi batek ez diola soilik lexikoari edo semantikari eragiten, pragmatikaren edo sintaxiaren ikuspegitik ere konsekuentzia argiak dituela.

Ipin dezagun lehen adibide bat: denok baliokidetzat jotzen ditugu «diskoa plazaratu» eta «diskoa argitaratu». Euskaltzaindiaren Hiztegian (EH) begiratzen badugu, haxe da esaten zai-guna *plazaratu* hitzari buruz:

¹ Nahiago genukeen artikulu hau Pello Salaburuk ere sinatu izan balu, gure orain arteko ia proiektu guztietan bidaide eta konplize izan baitugu.

plazaratu, plazara/plazaratu, plazaratzen

1 *da/du* ad. Plazara joan; plazara eraman.

2 *da/du* ad. Jendaurrean agertu edo azaldu. Ik. **kaleratu; argitaratu**. *Eragozpenak eragozpen, plazaratu ginen Etxeparez geroz. Plazaratu duen azken liburuan.*

Gure adibideari bigarren adiera zaio egokia: «Jendaurrean agertu edo azaldu». Pentsa genezake, beraz, *ikuskitzuna plazaratu* bezala esan genezakeela *ikuskitzuna argitaratu*. Baina ez da horrela, azken hori nekez topatu ahal izango da euskarazko testu batean. Gauza bera, *arrangura plazaratu / arrangura *argitaratu* bikoteekin. Beraz, usadioari behatuz gero, itxura batean badira gauza batzuk *plazaratu* daitezkeenak, baina ez *argitaratu*. Jakina, bigarren adiera horren definizioa alda liteke edo errestrikzio deskriptikoak gehitu, eta hori da lexikografia tradizionalan segitu ohi den bidea: adiera desberdinen arteko desberdintasun semantikoak hitzez deskribatu, horretarako definizioak sortuz.

Lexikografiaren ildo tradizionalak arreta jarri ohi dio hitz baten esanahi eta adiera desberdinen arteko bereizteari, eta berorien esanahia definizioen bitartez zehazten saiatzeari. Eta ez dago gaizki, aportazio handiak egin ditu eta egiten ari da bide hori. Baina askotan kamuts gertatu ohi da. Adibidez, EHN horrela definitzen da *ezetsi* hitza:

ezetsi, ezets, ezesten

1 *du* ad. Arbuiatu, ezertan ez hartu. Ik. **gaitzetsi**. *Lurreko zorionak ezetsirik. Giza egia, gure erabateko egia, ezetsiz gero, zer lirakeke egia eta gezurra?*

2 *du* ad. Ezezkoa eman, ez onartu. *Lege bat ezetsi.*

Bi adiera horietako zeinetan kokatuko ote genukeen adibide hau: —*Hori oso kasu berezia da —ezetsi zuen On Camillok*. Zaila, zinez. Seguru gaude *ezetsi* hitzaren konbinazio deskribapen baten informazio gehigarri eta argigarria aportatuko lukeela hori bezalako auzietan.

Ez da dena erraztasuna corpusen bidezko hiztegieta. Adibidez, kontu bat da adiera jakin barentzat adibide egoki bat ematea, eta beste bat esaldi jakin bat zein adieratan kokatu asmatzea. Horrek definizio zabalagoak egitera behartzen gaitu. Milaka adibideren artean *ezagutu*-ren kasua aipatuko dugu. Horrela dator oraingoz Eguno Euskararen Hiztegia (EEH):

ezagutu (*orobat ezautu g.er.*), ezagut(u), ezagutzen **1** *du* ad gogoan zerbaiten berri izan edo zerbaiten ideia eduki; norbaiten edo zerbaiten berri jakin; norbait nor den gogoan izan edo ohartu, norbaitekin halako harremana izan.

2 (*adizlagun eta kidekoekin*)

3 (*era burutua izenondo gisa*)

4 aitortu, onartu.

5 *ezin ezagutuzko* ezin ezagut daitekeena.

Hots, argi dago definizioaren arabera *ezagutu*-k gaztelaniazko *conocer* eta *reconocer*-en alor semantikoak betetzen dituela. Adibidez, esaldi hauetan:

Joanena beste bataiorik ezagutu ez arren.

Filosofiak ezin du erlijioa definitu gizartean duen eragina ezagutu gabe.

Proposatzen dugun hiztegi mota berri honen arabera, honek izan behar luke *ezagutu* kontzeptuaren definizioa: hitz hori agertzen den testuinguruaren multzoa, ageri diren konbinazioak sintak-

tikoki eta semantikoki sailkatuta eta multzokatuta. Gure ustez, eta eskuartean ditugun baliabideak probestuz, arrazoizkoagoa litzateke errestrikzio horiek beste modu batez ere adieraztea: hitz bat praktikan zer beste hitzekin konbinatzen den erakutsiz, konbinazio horiek corpusetan zer testuingurutan ageri diren bistaratuz.

Esan gabe doa, bi bideak ez daude kontrajarriak, eta elkarren osagarri dira. Batek ez du bestea kentzen. Lexikografoa saia daiteke, eta saiatzea komeni da, desberdintasun semantiko eta konzeptual horiek modu abstraktuan deskribatzen. Baina aldi berean, ukalezina da sarritan egiazko erabilera ikusteak askoz informazio zehatzagoa eta praktikagoa eskaintzen duela.

Adibideak ugariak izan daitezke:

gosea ase / gosea berdindu
*aldarrikapena ase / aldarrikapena *berdindu*

arriskua urrundu / arriskua uxatu
*helburua urrundu / helburua *uxatu*

gorputz nekatu / gorputz ahitu
*ibilera nekatu / ibilera *ahitu*

Baina badira ustez sinonimoak diren hitzak konbinatuz, esangura desberdinak sortzen diren kasuak, ugariak hauek ere:

gazteriaz arduratu / gazteriaz kezkatu
abesbatzaz arduratu / abesbatzaz kezkatu (bi gauza diferente)

prozesua indartu / prozesua azkartu
arnasa indartu / arnasa azkartu (bi gauza diferente)

errua onartu / errua aitortu
barkamena onartu / barkamena aitortu (bi gauza diferente)

jendetza batu / jendetza bildu
opariak batu / opariak bildu (bi gauza diferente)

Horrek garamatza «belarria»ren edo «sena»ren auzi ez beti behar bezain eztabaidatura. Garbi dago hizkuntzaren erabiltzaileek hitz segida batzuk bestetzek baino hobe hartzen dituztela, kate horietan dauden hitzak —hiztegi tradizionalen arabera— elkarren trukagarri izanagatik ere. Formula batzuek funtzionatu egiten dute, eta bestetzek ez.

Har dezagun *duda* eta *zalantza*ren arteko ustezko baliokidetza. Inor gutxik esango du sinonimo petoak ez direnik. Hitz soil bezala hartuta, erabiltze tasa handikoak dira biak. Orotariko Euskal Hiztegian (OEH) antzeko luzera dute bi sarrerek, eta baita EEHn ere. Baina, beste hitz batzuekin konbinatzen hasten direnean, aldatu egiten da kontua:

duda-mud(et)an (210)²
zalantza-balantz(et)an (9)
zalantza-malantz(at)an (3)

duda izan (13)
zalantza izan (111)

² Besterik adierazi ezean, artikuluan aipatuko diren agerpen kopuru guztiak Egungo Testuen Corpusetik (ETC) hartuak dira.

dudan jarri (339)
zalantzan jarri (42)

Ez zaigu iruditzen kopuru diferentzia horiek huskeria direnik. Agian usadioaren inertziari dagozkio kasu askotan, nork jakin gaur-gaurkoz. Baina gaude, horretarako ere balioko lukeela konbinazio hiztegi batek, hau da, «belarria» deitu izan den horrek nola funtzionatzen duen hobeki deskribatu ahal izateko.

Hiztegiek ez dute askotan hitz berririk deskubritzen, aurrez dakizkigunak gogorarazi baizik. Gauza bera, edo areago, esan daiteke balizko konbinazio hiztegi bati buruz. Balio arauemaile edo deskriptibotik harago doan funtzio bat da. Testu bat idazten edo itzultzen ari garenean, maiz gertatu ohi da hitz bat —edo hitz konbinazio bat— «mihi puntan» izatea, baina une horretan ahora edo burura ez etortzea. Horretarako ere balio dute hiztegiek, idazleen eta itzultzaileen lagungarri, eta begien bistakoa da konbinazio hiztegi batek horri buruz egingo lukeen aportazioa.

Gehienetan, hiztegiek ez diote hiztunari informazio berririk ematen, aurrez zekiena «gogorarazi» baizik. Areago, sarritan ohiko inpresioa da hiztegiek ematen duten informazio hori ez zela beharrezkoa, erabiltzaileek «natural» edo «ebidente» jotzen dutelako, areago adibideak irakurri ostean. Justu kontrakoa da, gure ustez. Hitzek beste hitz batzuekin, eta ez beste batzuekin, konbinatzeko ageri duten kasuistika deskribatzea eta sistematizatzea aportazio aparta izan liteke erabiltzen den hizkuntzaren kalitaterako.

Har dezagun sinonimoen esparru —zenbaitetan— labainkorra. Iritzi aski orokorra da, gaur egun euskararako dauden sinonimo hiztegiak aski kamutsak direla. Batik bat, euskalkien eta hizkeren araberako aldaketan oinarritu ohi dira, eta ez hainbeste kontzeptuen hurbiltasunean. Gure ustez, antzeko testuinguruetan zein hitz erabiltzen diren ikusita, beste hurbiltze efikazago bat izango genuke sinonimoen hiztegi baterako ere, hots, arrazoi estilistikoengatik edo funtzionalengatik hobe letorkigukeen hitz konbinazio bat detektatu eta erabili ahal izateko.

Hitzen definizioaz gain, hitzaren «kalitatea» edo «balioa» neurrian jartzeko ere jakingarria da hitz jakin bat nork erabiltzen duen. Ez da berdín hitz jakin bat halakok edo halakok erabiltzea. Alde horretatik, inportantea deritzogu adibide bakoitza nork eta non erabilia izan den azaltzeari.

Esan daiteke euskal hiztegi gintzaren hutsune handietako bat —eta autokritika bezala ere ulerbedi— adverbioena dela. Adibidez, hiztegi guztietan jasota dago *gertu* adverbioa, baita *gerturik* ere, eremu semantiko oso desberdinekoak bi-biak. Baina, adibidez, *gertutik* adverbioak ez du sarrrera bereizirik gure hiztegi-tan. Eta, konbinatoriak baino azaltzen ez duen bezala, semantikoki oso gauza diferenteak dira *gertu ikusi* eta *gertutik ikusi*. Lehenago aipatu dugun ustezko konbinatoria librearen kontrako adibide asko inbentaria daitezke adverbioetan ere:

galdera sakonak / gogoeta sakonak
**sakonki galdetu / sakonki gogoetatu*

Hiztegi guztiek esango digute *zehatz* eta *konkretu* sinonimoak direla, baina *zehatz adierazi* edo *zehazki adierazi* hitz konbinazioen agerpenak aski sarriak diren bezala, agerpen bakar bat ere ez dugu topatu *konkretu(ki) adierazi* konbinaziorako.

Baina hitzen konbinatoria, guztiz librea ez denez, hizkuntzaren bestelako sail batzuekin elkar lotzen da, morfo-sintaxiarekin konparazione. Kasu bat aipatzearren, «singularra/plurala» auzia dago. Aditz askok singularreko nahiz pluraleko osagaiak izan ditzakete (*amets bat izan / ametsak*

izan), baina ez da inola ere hala aditz guztietan. Har dezagun *txirikordatu* aditza; ezinbestean, bere osagai batzuk pluralean eta beste batzuk singularrean hartuko ditu:

hatzak/ideiak/gaiak/hitzak txirikordatu (pluralean)

ilea/lastoa/lokarria/artilea/sugea txirikordatu (singularrean)

2. Joerak eta planteamenduak

Konbinazio hiztegitzaren sorrerak erabat aldatu du hiztegiak eratzeko modua. Hiztegitzaren tresna berri honen bidez hitzak edo lemak konbina ditzakegu halako eran non lortu baititza-kegu unitate lexiko, semantiko eta pragmatikoak. Konbinazio lexiko horiek hiru mota nagusitan bildu ditzakegu.

Lehenengo motan lokuzioak eta modismoak ditugu. Mota konplexuak dira, horien esanahia ez baita osagaien konbinaketaren ondorio soilak. Adibide bat *hotz eta motz* izan daiteke, non erreferentziak ez baitu zer ikusirik tenperaturarekin edo luzerarekin: *Ez zen gizonaren eta herrien alderdi bat ere Lafitte hotz eta motz utziko zuenik*.

Bigarrenengo motan «kolokazio» deituak ditugu. Kolokazioak hitz multzoak dira, hiltunen erabilerak halako berezitasun bat eman dienak. Gorago aipatu ditugun *gosea berdindu* eta *arriskua uxatu*-ren kasua bete-betean kokatzen da kolokazioen arloan. Argi dago kolokazioen garrantzia: kolokazioen erabilerak erakusten digu azken batean aurrean dugun hiltunak berezkoa duen hiltzuntza ala guztiz euskaldundu gabeko baten aurrean gauden.

Hirugarren motan, «errestrikzio lexikoak» ditugu. Bistan da hitz konbinazio ugari egin egin daitekeela, baina bistan da orobat konbinazio asko daudela ezinezkoak direnak hiltzuntzaren erabileraren aldetik. Kostata esango du hiltzun batek *zeru ausarta* edo *zubi errukiorra*, ezpada poesia mota berezi batean. Baina edozein hiltzunek onartuko luke *itsasontzi baten sabelean etorri nintzen ni hona* bezalakoak.

Aski modu sistematikoan azter ditzakegun konbinazio lexikoak —hots, hiltzuntzaren osagai objektibo legez—, ez dira soilik lokuzioak edo esamolde bereziak, fraseologiaren edota terminologian alorrean aztertu ohi direnak. Sistema linguistikoak ustez ahalko liratekeen konbinazioetan errestrikzio asko eta asko ezartzen ditu, aurrez pentsa genitzakeenak baino askoz ere gehiago. Aditx trantsitibo batek zein osagai izan ditzakeen, adberbio batek zein aditx modifika ditzakeen, edota adjektibo baten eta izen baten artean zein konbinazio harreman praktikoa dagoen.

Horiek horrela, eztabaidagai izan da beti zer den kolokazio bat eta zer den errestrikzio lexiko bat. Dena dela, hurbilketa gisa esan dezakegu hauek direla kolokazioen eta errestrikzio lexikoen ezaugarriak. Lehenik, konbinazio lexikoak dira, baina lexemak esaten digu zein piezarekin konbina daitekeen. Bigarrenik, erabilerak errepikatuak, eta ez beste ezerk, eusten dio delako konbinazioari. Hirugarrenik, eta zertzelada honek garrantzia berezia du hiztegitzaren, kolokazioak ezin dira beti eta automatikoki hiltzuntza jakin batetik beste batera itzuli, eta, hortaz, lan berezia eskatzen diete hiztegitzei.

Kolokazioen eta errestrikzio lexikoen alorrean, bi joera nagusi aipa daitezke. Lehengoa maiztasunetan oinarrituriko konbinatoria da; bigarrena, deskribapen formaletan oinarriturikoa. Lehenengo joeran *The BBI combinatory dictionary of English* hiztegia izan zen lehena. Bigarren joeran

oinarrituriko lan garrantzizkoena *Dictionnaire explicatif et combinatoire du français contemporain*, lau liburukitan publikatu zena 1984 eta 1999 urteen artean. Nolanahi ere, geure gusturako, gehien segitu beharreko erreferentziak bihurtu dira Ignacio Bosqueren gidaritzapean argitaratu ziren *Redes* (2004) eta *Práctico* (2006) konbinazio hiztegiak.

3. Orain arteko pausoak

Euskal konbinatoria lexikoan dena ez da desertu. Orain artean ere, aitzinlanak egin izan dira, eta euren aportazioa aintzat hartzekoa da dudarik gabe, abiapuntu bezala besterik ez bada ere.

Tradizio handiena duen alorra, lokuzio, esapide idiomatiko eta terminologiarena izan da gure artean. Horrek orain arte eman duenaren emaitzaren isla modura, har dezagun Elhuyar-en *Euskara-gaztelania hiztegia*, orain arte egindako lanaren adibide paradigmatico bezala³. Hiztegi horretan, hitzen konbinazio lexikalizatuak dagokienez, «lexiak» (*bigarren eskuko*, *esku-pilota*, *esku-hartze...*) eta «lokuzioak» (*eskuz aldatu*, *esku beteka*, *eskutik irabazi...*) bereizten dituzte. Hauek dira gaur arte euskararako detektatu dituztenen kopuruak: 7.397 lexia eta 5.161 lokuzio.

Esapideez edo lokuzioez haragoko konbinazio libreagoak zerrendatzen egindako lehen saioa, guk dakigula, Ereduzko Prosa Gaur (EPG) corpusean egin zen, webeko kontsulta tresnan «Hitz multzoak» izeneko atalean. Bertan, corpuseko bi eta hiru hitzeko multzo guztiak erauzi ziren, inolako murrizketarik gabe. Orduko hura baliatuta txertatu izan dira EEHko zenbait sarreratan ageri diren konbinazio multzoak. Saio horrek muga eta ments nabariak ditu: soilik hitzak hartzen ditu aintzat (ez lemak), eta hitz guztiei balio bera ematen die, inolako ponderaziorik gabe.

Egungo Testuen Corpusean (ETC) beste saiakera bat egin da, oraingoan lemetan oinarrituta eta ponderazio neurri batzuk aplikatuta, kategoria gramatikala ere kontuan hartuta. 186 milioi testu hitzeko corpus baterako 3.222.771 hitz bikote bereizi dira. Kontsulta interfazeaz hitz bakoitzaren barruan ematen da hitz hori ageri deneko konbinatoria guztia, eta posible da emaitza iragaztea, kategoriaren eta kokapenaren arabera. Baina tresna horrek ere muga garbiak ditu: bi hitzetik gorako konbinaziorik ez da ageri, eta lemei soilik egiten zaie jaramon (esate batera, *esku hartu* konbinazioa ageri da, baina bere baitan bereizi gabe hauek denak bilduta: *esku hartu* bera, *eskuz hartu*, *eskutik hartu*, *eskuan hartu...*).

Azken aldiko beste ahalegin azpimarratzekoa Elhuyar-en Web-corpusen atariarena⁴ izan da. Horko «Hitz-konbinazioak» atalean, euskarazko corpus elebazarretik automatikoki erauzitako bigramak kontsulta daitezke. Hiru kategoria-konbinaziotako bigramak erauzi zituzten: «izena + aditza», «izena + izena» eta «izen + izenondoa». Neurri estatistiko batzuen arabera ordenatzen da emaitza, eta konbinazioen testuinguruak ere bistaratu daitezke. Hauek dira detektatutako hitz konbinazioen kopuruak: «izena + aditza»: 71.550, «izena + izena»: 33.759, «izen + izenondoa»: 27.931. Tresna horretako egileek gerora egin dute bestelako garapenik; adibidez, «izenlaguna + izena» konbinazioa kontsideratuta eta murrizketa morfosintaktikoak aplikatzen.

³ https://hiztegiak.elhuyar.eus/eu_es/

Eskerrak eman nahi dizkiogu Elhuyar-eko Klara Ceberio Berger-i, adiatsuki helarazitako datuengatik.

⁴ <http://webcorpusak.elhuyar.eus/>

Oraingoan Antton Gurrutxagaren eskuzabaltasunari zor dizkiogu datuak.

4. Proposamenaren azken zertzeladak

Euskararako konbinazio hiztegi batek nolakoa behar lukeen modu apur bat praktikoagoan go-goetatzera ausartuko gara, azkenik. Gure ekosistema hurbilean izandako esperientziek eta testuiguru globalagoan dauden joerak aintzat hartuta, hauexek ikusten ditugu abiapuntuko irizpide nagusiak:

- Euskarria elektronikoa izango da. Alegia, interneten soilik argitaratuko litzateke. Ez soilik horrelako lan bat paperera eramatea astakeria ekonomikoa litzatekeelako, baizik eta edukiaren muntak berak hori duelako zabaltzeko bide egokiena: behar besteko hedadura izango ahal lukete sarrerek, eta hitzen arteko harremanak eta dependentziak modu egokiagoan bistaratu ahal izango lirateke hiperesteken bitartez.
- Adibide guztiak nork edota non argitaratuak izan diren zehaztea inportantea da zeharo. Gorago ere aipatu dugu autoritatearen figurari, ereduzkotasunari, ematen diogun pisua.
- Begien bistakoa denez, gure giza ahalmena, lexikografiari dagokionez, mugatua da. Gurean, pentsaezina da horrelako proiektu bat gauzatzeko hamabi laguneko talde bat osatu ahal izatea, eta haiei soldata duin bat ordaindu ahal izatea. Alde horretatik, beharrezkoa izango da alor horretan orain arte jardundako eragileen arteko parte hartzea, adostasuna eta elkarlana.

Erreferentziak

- Benson, Morton, Evelyn Benson & Robert Ilson. *The BBI combinatory dictionary of English: a guide to word combinations*. John Benjamins Publishing Company, 1986.
- Bosque, Ignacio. *Redes. Diccionario combinatorio del español contemporáneo*. SM, 2004.
- Bosque, Ignacio. *Diccionario combinatorio Práctico del español contemporáneo*. SM, 2006.
- Elhuyar Fundazioa. *Euskara-gaztelania hiztegia*. Elhuyar, 2019 [<https://hiztegiak.elhuyar.eus/>].
- Elhuyar Fundazioa. *Web-corpusen ataria*. Elhuyar, 2013 [<http://webcorpusak.elhuyar.eus/>].
- Euskaltzaindia. *Euskaltzaindiaren Hiztegia*. Euskaltzaindia, 2019. [https://www.euskaltzaindia.eus/index.php?option=com_hiztegianbilatu&Itemid=410].
- Mel'čuk, Igor *et alii*. *Dictionnaire explicatif et combinatoire du français contemporain*. John Benjamins Publishing Company, 1984-1999.
- Mitxelena, Koldo *et alii*. *Orotariko Euskal Hiztegia*. 6. argitaraldia. Euskaltzaindia, 2019. [https://www.euskaltzaindia.eus/index.php?option=com_oehberria&Itemid=413].
- Sarasola, Ibon. *Egungo Euskararen Hiztegia (EEH)*. EHU/UPV, Euskara Institutua, 2019. [<http://www.ehu.eus/eeh/>].
- Sarasola, Ibon, Pello Salaburu & Landa, Josu. *Egungo Testuen Corpora (ETC)*. EHU/UPV, Euskara Institutua, 2013. [<http://www.ehu.eus/etc/>].
- Sarasola, Ibon, Pello Salaburu & Landa, Josu. *Ereduzko Prosa Gaur (EPG)*. EHU/UPV, Euskara Institutua, 2007. [<http://www.ehu.eus/euskara-orria/euskara/ereduzkoa/>].

Processing Preferences in an Ergative Language: Evidence from Basque Postnominal Relative Clauses

IRAIA YETANO

Independent Researcher

JON ANDONI DUÑABEITIA

Universidad de Nebrija

ITZIAR LAKA

University of the Basque Country (UPV/EHU)

Abstract

This study aimed at contributing to a better understanding of relative clause (RC) processing preferences across languages. Although a preference for subject relatives (SR) has been consistently found in various languages, ergative languages where the notion of subjecthood may be called into question have been grossly understudied. This paper addressed the question of whether word order differences can modulate the previously reported object relative (OR) advantage of prenominal RCs in Basque (Carreiras *et al.*, 2010), an ergative language that allows for prenominal as well as postnominal relatives. A self-paced reading experiment with Basque postnominal RCs showed a SR advantage. The reversed asymmetries for prenominals and postnominals in Basque are discussed considering processing strategies related to the ergative nature of Basque.

Keywords: RC processing, postnominal RCs, ergativity, Basque.

1. Introduction

One of the issues psycholinguistic research seeks to understand is what the mechanisms underlying human language processing are, and whether the processing system's strategies are universal or vary cross-linguistically. Subject preference stands as one of those possible candidates for a universal language processing strategy, given the abundant research attesting subject/object asymmetries in relative clause processing across languages. A sizeable body of research has shown that relative clauses (hereafter RCs) with gaps in subject position (subject relatives, SRs) like (1a) are easier to process than those with gaps in object position (object relatives, ORs) like (1b).

- (1) a. The reporter [_{RC} who_i e_i attacked the senator] admitted the error.
- b. The reporter [_{RC} who_i the senator attacked e_i] admitted the error.

The bulk of research on subject/object asymmetries has focused on head-initial languages with postnominal RCs, in which the RC containing the gap follows its head noun. A preference for SRs has been consistently found, inter alia, in English (Caplan, Vijayan, Kuperberg, West, Waters, Greve, & Dale, 2001; King & Just, 1991; King & Kutas, 1995; Traxler, Morris, & Seely, 2002;), German (Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Schriefers, Friederici, & Kühn, 1995), French (Cohen & Mehler, 1996; Frauenfelder, Segui, & Mehler, 1980; Holmes & O'Regan, 1981), Dutch (Frazier, 1987; Mak, Vonk, & Schriefers, 2002, 2006), Brazilian Portuguese (Gouvea, 2003), and Spanish (Betancort, Carreiras, & Sturt, 2009).

It has not been until the last decade when head-final relative constructions have become a target to study whether and how headedness impacts upon the directionality of RC processing preferences. Unlike head-initial or postnominal RCs, head-final or prenominal relative constructions appear preceding the head noun. Head position yields opposed orderings of filler (head noun) and gap: whereas in postnominal RCs forward filler-gap dependencies are created, prenominal RCs present backward gap-filler dependencies, as the syntactic gap precedes the head noun in the RC construction. The crucial question is whether the same subject/object processing asymmetries hold even in those languages in which head-position is reversed. Prenominal RCs have primarily been studied in East Asian languages like Chinese, Japanese, and Korean. In such languages there is usually no relativizer marking the beginning of the RC, thus triggering a temporary structural ambiguity between a main clause and an embedded clause reading for a sentence-initial clause.

A consistent subject advantage has been found in Japanese both in behavioral experiments (Ishizuka, 2005; Ishizuka, Nakatani, & Gibson, 2003; Miyamoto & Nakamura, 2003) and electrophysiological studies (Ueno & Garnsey, 2008). An alternative explanation for the SR advantage might be that ORs have greater structural ambiguity, and therefore, are more likely to be misread as a main clause. Arguing along these lines, Ishizuka, Nakatani, and Gibson (2006) presented an exception to the previous subject advantage processing pattern, as ORs were reported to be more easily processed after a contextual manipulation that reduced the main clause/RC temporary ambiguity. However, their results have been called into question for several experimental reasons, and indeed, they were not later replicated in follow-up studies (refer to Kwon, Lee, Gordon, Kluender, & Polinsky, 2010 for a detailed discussion of the results). Results for the RC processing asymmetries in Korean are unequivocal: both self-paced reading (Kwon, Polinsky, & Kluender, 2006), eye-tracking (Kwon *et al.*, 2010), and ERP experiments (Kwon, Kluender, Kutas, & Polinsky, under revision) have confirmed a subject preference, both with and without supportive contexts. Contrary to the latter, the status of RC processing in Chinese is somewhat controversial, with some studies reporting an OR preference (Gibson & Wu, 2011; Hsiao & Gibson, 2003; Lin & Garnsey, 2007), and others finding a SR processing preference (Lin & Bever, 2006a,b; Kuo & Vasishth, 2006; Yang, Perfetti, & Liu, 2010).

Leaving aside the unresolved status of RC processing in Chinese, the results on subject/object asymmetries so far seem to tip the scales in favor of a general or universal subject preference in RC processing, despite the cross-linguistic variation regarding head position in RC constructions. This universal preference for subjects was first brought into scene by Keenan and Comrie (1977) in the *Noun Phrase Accessibility Hierarchy* (hereafter, AH). The AH was proposed to account for the cross-linguistic variation regarding the relativizability of different NP positions. Based on data from about fifty languages, they proposed the universal implicational scale in (2), according to which «subjects» are the highest in the hierarchy, and are thus the easiest to relativize.

- (2) Subject > Object > Indirect Object > Oblique > Genitives > Object of Comparison

The authors further claimed that the AH has a psychological rationale, so that the hierarchy «directly reflects the psychological ease of comprehension» (Keenan & Comrie, 1977: 88). This way, NPs relativized at low positions are harder to understand than NP positions that are higher in the hierarchy. RCs with subject gaps like (1a) are thus hypothesized to be easier to process than RCs with object gaps like (1b) across languages and irrespective of head position.

Nonetheless, a claim based on universal processing preferences needs support from a wide variety of languages. Ergative languages, around 25% of the world languages (Dixon, 1994), have received little attention in the language processing literature. Ergative languages mark intransitive subjects and objects of transitive verbs with absolutive case, differently from subjects of transitive predicates, which bear ergative case (Dixon, 1994), as illustrated in example (3) from Basque.

- (3) a. Irakasle-a-k ikasle-a-ø zoriondu du eskola-n.
Teacher-sg-Erg student-sg-Abs congratulated has school-at.
«The teacher congratulated the student at school».
- b. Ikasle-a-ø berandu iritsi da eskola-ra.
Student-sg-Abs late arrived is school-to.
«The student arrived late to school».

Ergativity can prove to be a really fruitful and enlightening pathway for further studying the validity of a universal subject preference. All the languages mentioned so far are nominative-accusative, and thus align sentential arguments in such a way that subjects are nominative no matter the transitivity of the verb, and objects always bear accusative case. By contrast, ergative languages allow us to differentiate between a preference for grammatical functions (subject vs. object) and a preference for case (ergative vs. absolutive). In order to provide a full picture of universal processing strategies in general and RC processing asymmetries in particular, the compelling question to be answered is whether a subject preference still holds in languages like Basque, where subjects are not unequivocally marked with a single case marking. If a general preference for subject extractions were to be a universal processing mechanism, the subject/object processing asymmetry in an ergative language should go in the same direction as in nominative-accusative languages.

2. Subject/object asymmetries in an ergative language

In a recent study of processing asymmetries in prenominal RCs, Carreiras, Duñabeitia, Vergara, De la Cruz-Pavía, and Laka (2010) conducted two self-paced reading and one ERP experiment in Basque, a head-final ergative language. Basque is an SOV language but word order variation is allowed and very frequent. Basque is a three-way pro-drop language, allowing not only for null subjects, but also for object omission (both direct and indirect). This is the same as to say that Basque can drop arguments bearing ergative, absolutive, and dative case. Prenominal RCs have no Wh element heading the relative construction, just as in Chinese, Japanese and Korean. The complementizer *-(e)n* attached to the subordinate verb in clause-final position marks the clause as a RC. Note that the properties of Basque give rise to a quite common temporary ambiguity in head-final RCs: A prenominal RC can be initially interpreted as a main clause with dropped or scrambled arguments, at least until the parser reaches the verb to which the RC marker is attached to.

A morphological ambiguity arises when the ending *-ak* is attached to an NP, which can be interpreted either as singular ergative (i.e., subject of transitives) or as plural absolutive (i.e., subject of intransitives or object of transitives). When interpreted as singular ergative, the sequence consists of *-a*, the singular determiner *the_{sg}*, and *-k*, the ergative case marking. When interpreted as absolutive plural, its morphological structure corresponds to the single morpheme *-ak*, the plural determiner *the_{pl}*, since absolutive case has no overt marking. This way, both *irakasleak* («teacher») and *ikasleak* («student») in (4) are ambiguous in the two conditions until the inflected verb *dira/ditu* is reached. In the case of a subject relative like (4a), the processor necessarily has to (re)interpret the first NP as the subject of a transitive when facing the verb *dira* («are»). The opposite is true for (4b), in which the verb *ditu* («has») disambiguates the sentence towards an OR interpretation.

(4) a. (SR)

Irakasleak aipatu dituen ikasleak lagunak ditu orain.

[e_i irakasle-ak aipatu ditu-en] ikasle-a-k_i lagun-ak ditu orain.

[e_i teacher-pl mentioned has-Comp] student-sg-erg_i friend-pl has now.

«The student that mentioned the teachers has friends now.»

b. (OR)

Irakasleak aipatu dituen ikasleak lagunak dira orain.

[irakasle-a-k e_i aipatu ditu-en] ikasle-ak_i lagun-ak dira orain.

[teacher-sg-erg e_i mentioned has-Comp] student-pl_i friend-pl are now.

«The students that the teacher mentioned are friends now.»

Taking advantage of this morphological ambiguity, Carreiras *et al.* (2010) constructed pairs of sentences containing a subject and an object relative clause each. As a result, sentences in each pair were ambiguous between a subject and an object relative reading until the inflected verb (*dira/ditu*) in the main clause was reached. Prenominal ORs were read faster than SRs at the disambiguating region, and SRs produced larger amplitudes in the P600 window, just after reading the disambiguating word *ditu*. The results pointed to a disadvantage for prenominal SRs, thus disconfirming universal processing mechanisms based on a general preference for subject gaps in RCs. Overall, neither a general saliency of subjects nor the head noun position alone offer a comprehensible general picture of subject/object asymmetries in RC processing across languages. In the event of the results obtained for Basque, processing of RCs appears to be case sensitive. Ergativity is indeed a plausible candidate, but how exactly the ergative nature of Basque is affecting RC processing is not clear-cut. One could hypothesize that in ergative languages like Basque the processing system prefers the morphological case of the gapped element to be absolutive over ergative. The processing disadvantage for ergative DPs/NPs is usually explained in terms of markedness. Morphological unmarkedness is believed to provide a processing advantage in language, such that unmarked cases are preferred in processing and order of acquisition over marked ones (Otsuka, 2006). This was already suggested for Basque by Carreiras *et al.* (2010) and Laka (2013). Ergative case is the marked case in Basque, whereas the absolutive case is the unmarked. This becomes clear from the *-k* ending for ergative, contrasted with the zero marking for absolutive. In nominative-accusative languages, by contrast, the nominative case is the unmarked case, as opposed to the markedness of accusative case, thereby predicting a subject (i.e., nominative) preference. Morphological markedness could thus explain why ORs are preferred in an ergative language like Basque but dispreferred in nominative-accusative languages irrespective of head noun position.

However, the RC processing asymmetries in Basque might also stem from processing preferences made for a sentence-initial NP. It is generally assumed that especially in verb final languages, where the verb is usually not available until the very end of the sentence, comprehenders rely much on morphological information such as case marking on NPs to interpret *Who-did-what-to-whom* (Bornkessel & Schlesewsky, 2006; Yamashita, 1997). Yet, morphology sometimes does not help in the form-to-meaning mapping. Cross-linguistic evidence suggests that, upon encountering an ambiguous sentence-initial argument, the language processing system shows a tendency to interpret the initial NP as the subject of the sentence (De Vincenzi, 1991; Demiral, Schlesewsky, & Bornkessel-Schlesewsky, 2008; Schriefers, Friederici, & Kühn, 1995; Wang, Schlesewsky, Bickel, & Bornkessel-Schlesewsky, 2009). Again, given ergativity, the question arises as to whether initial interpretive parsing decisions are universal, respond to grammatical function, or are instead sensitive to case form. In view of the object advantage found for prenominal RC processing by Carreiras *et al.* (2010), one could argue that Basque speakers preferably interpret an initial ambiguous NP as ergative-marked rather than as absolutive-marked. Let us consider this possibility in greater depth. When confronted with an ambiguous *-ak* ending in initial position like in (4), participants could interpret it as:

- (i) a transitive singular subject (ergative marked)
- (ii) an intransitive plural subject (absolutive marked)
- (iii) a plural object (absolutive marked)

The authors found that the main verb was read faster when it disambiguated the sentence towards an OR. These processing preferences suggest that the disambiguation towards a SR required a more costly reanalysis than the OR reading, which in turn implies that the parser initially committed to a singular ergative interpretation of the initial NP rather than the plural absolutive interpretation. Should initial choices favor a plural absolutive interpretation like in (ii) or (iii), a garden-path effect would be expected at the disambiguating verb in the OR condition. Yet, no revision effects were found at the prenominal OR condition.

This same preference in ergative languages is explained in terms of dependent and independent case marking by Polinsky, Gómez-Gallo, Graft, and Kravtchenko (2011) for Avar. In an ergative language, the dependent case is ergative, since it can usually only appear in structures where the absolutive is present. The opposite is true for accusative languages, where the dependent case corresponds to the accusative. RCs with an object absolutive gap should show a processing advantage in ergative languages, because, once the parser has encountered an initial ergative marked (dependent) argument, the gap in the absolutive position should be easier to project than upon encountering an initial absolutive case-marked argument.

These two processing forces, namely a preference for unmarked (absolutive) gaps together with a preference to interpret an initial ambiguous NP as ergative, would both favor an OR advantage for Basque prenominal RCs like (4): upon reading the first word *irakasleak*, participants possibly interpreted it as singular ergative, and thereby, as a transitive subject of a main clause (Frazier, 1987); later in the sentence subjects identified they were reading a RC, and consequently posited an absolutive gap; these incremental interpretive choices created a garden-path at the disambiguating verb window in the SR condition but no revision was required in the OR condition. The critical issue here is whether these processing strategies related to ergativity work in tandem, or rather, one is overwhelmingly more powerful than the other. Unfortunately, this is something we cannot answer based on results from Basque prenominal RCs like (4).

3. Processing postnominal Relative Clauses in Basque: a self paced reading experiment

Crucially, Basque also has postnominal RCs. Unlike prenominals, postnominals are headed by a *wh*-pronoun, and they are preferably interpreted as appositive or non-restrictive. Though mainly used in high registers, they have been used in written Basque since the 16th century (Oyharçabal, 2003), and are currently used in standard language. The study of postnominal RC processing in Basque provides a unique path to examine in a single language the challenges word order differences in RC constructions push to the language processing system. This matter is rendered particularly interesting by the fact that morphological markedness and sentence-initial processing choices pull in opposite directions in the case of Basque postnominal RCs. Consider the sentences in (5), which exhibit the postnominal versions of the SR and OR in (4).

(5) a. (SR)

Ikasleak, zeinak irakasleak aipatu baititu, lagunak ditu orain.

Ikasle-a-k_i, [zein-a-k_i e_i irakasle-ak aipatu bait-ditu,] lagun-ak ditu orain.

Student-sg-erg_i, [who-sg-erg_i e_i teacher-pl mentioned Comp-has,] friend-pl has now.

«The student, who mentioned the teachers, has friends now.»

b. (OR)

Ikasleak, zeinak irakasleak aipatu baititu, lagunak dira orain.

Ikasle-ak_i, [zein-ak_i irakasle-a-k e_i aipatu bait-ditu,] lagun-ak dira orain.

Student-pl_i, [who-pl_i teacher-sg-erg e_i mentioned Comp-has,] friend-pl are now.

«The students, who mentioned the teacher, are friends now.»

A preference for morphologically unmarked gaps would predict the same processing asymmetry in prenominal as well as postnominal RCs in Basque, under which an OR advantage is expected. Sentence-initial processing choices favoring a singular ergative interpretation over a plural absolutive one would instead predict a SR advantage for postnominal RCs, contrary to what is expected for prenominals given the same processing strategy. Here we report a self-paced reading experiment with subject- or ergative-gapped and object or absolutive-gapped postnominal RCs in Basque, such as those in (5a) and (5b) respectively. Should we find a reading slowdown at the critical disambiguating verb in the OR condition, this would indicate a processing advantage for SRs, and consequently, an ergative preference for initial ambiguous NPs would be argued to be the main factor affecting RC processing asymmetries. If, by contrast, an OR preference holds even with postnominal RCs, then markedness-based accounts of RC processing would gain force.

3.1. *Methods and materials*

PARTICIPANTS

Fourty students at the University of the Basque Country (UPV/EHU) participated in the experiment as paid volunteers after giving informed consent (36 female; age range 18-31, mean age 19.72). All were native speakers of Basque with normal or corrected-to-normal vision and no history of language-related deficits. A linguistic background questionnaire attested to the participants' proficiency, age of acquisition, and exposure to Basque.

MATERIALS

Twenty-six experimental sets like (5) were constructed, each with two RC conditions: a postnominal SR and a postnominal OR. RCs were fully ambiguous between a SR and an OR reading until the main verb (REGION 7, the critical region). A post-critical word was included to avoid effects associated with sentence-final integration processes (Aaronson & Ferres, 1984). Except for the penultimate word, the two versions in each pair involved the same words displayed in the same order, so lexical frequency was controlled overall. The disambiguating verbs (*ditu* -has- and *dira* -are-) were controlled for length (four letters long each), frequency (3.25 and 3.81 respectively), and orthographic neighborhood (15 and 16 orthographic neighbors, respectively) using E-Hitz (Perea, Urkia, Davis, Agirre, Laseka, & Carreiras, 2006). In order to control for case (mis)match effects (Sauerland & Gibson, 1998), the case of the *wh*-pronoun and the antecedent NP matched in each condition: ORs all had the antecedent NP and the *wh*-pronoun in plural absolutive (-ak), whereas SRs had both in singular ergative (-a-k) (both being formally indistinguishable).

Table 2

Experimental sample set

Condition	Example
(a) SR	Ikasleak, zeinak irakasleak aipatu baititu, lagunak ditu orain. Ikasle-a-k, [zein-a-k irakasle-ak aipatu bait-ditu], lagun-ak ditu orain. Student-sg-S, [who-sg-S ₁ teacher-pl mentioned COMP-has], friend-pl has now. «The student, who mentioned the teacher, has friends now.»
(b) OR	Ikasleak, zeinak irakasleak aipatu baititu, lagunak dira orain. Ikasle-ak ₁ , [zein-ak irakasle-a-k aipatu bait-ditu], lagun-ak dira orain. Student-pl ₁ , [who ₁ teacher-sg-S mentioned COMP-has], friend-pl are now. «The students, who the teacher mentioned, are friends now.»

Participants were randomly assigned to one out of two counterbalanced lists, such that each read an equal number of SRs and ORs, but only one version of each pair. Experimental items were interspersed among 74 unambiguous fillers of similar length, some of which contained postnominal RCs. Overall, around 25% of the sentences were ambiguous RCs. Trial order repetition effects were avoided randomizing the presentation order for each participant.

PLAUSIBILITY SURVEY

To guarantee that reading materials were all equally plausible as SRs or ORs, thirty-six Basque informants (who did not participate in the experiment) rated a total of seventy-two sentences on a scale ranging from 1 (totally implausible) to 7 (totally plausible) in a separate off-line norming survey. Forty-eight postnominal SR and OR pairs were constructed and intermixed with twenty-four filler sentences of various linguistic structures —with and without postnominal RCs— and of various degrees of plausibility. Experimental items were divided in two counterbalanced lists, and each of the lists was then divided in two lists that differed in order of presentation.

The twenty-two sentence pairs with the largest within-pair differences were discarded. Ratings for the remaining twenty-six pairs did not differ significantly ($p > .5$), indicating that the

RC heads in the experimental sentences were equally plausible as the subject (mean rating 5.10; $\pm .53$) or the object (mean rating 5.18; $\pm .42$) of the RCs.

PROCEDURE

A word-by-word self-paced-reading experiment was conducted using the moving window paradigm (Just, Carpenter, & Wooley, 1982). Participants were tested individually in a silent and correctly illuminated room. Stimuli presentation and data collection was performed with Linger (version 2.94) software by Doug Rohde (see <http://tedlab.mit.edu/~dr/Linger/>) on a computer with a 15.4 inch CRT monitor. The reading pace was set by the participants' key pressing. Participants were initially presented with a computer screen filled with a string of dashes and spaces. Each time the space bar was pressed one word was unmasked, keeping the rest masked. Reading times (henceforth, RTs) were measured as time lapses between the display of each word and the key press. After each experimental sentence a yes/no comprehension question was displayed all at once on a new screen. Participants pressed «A» if they believed the answer was *yes*, and «L» otherwise. Correct answers were balanced, and only feedback to incorrect answers was provided. Written instructions and a practice session preceded the experimental trial. Sentences were presented in a single line using a 25-point courier font. Participants had breaks every twenty-five items. The whole session lasted about 20 minutes.

3.2. Data Analysis and Results

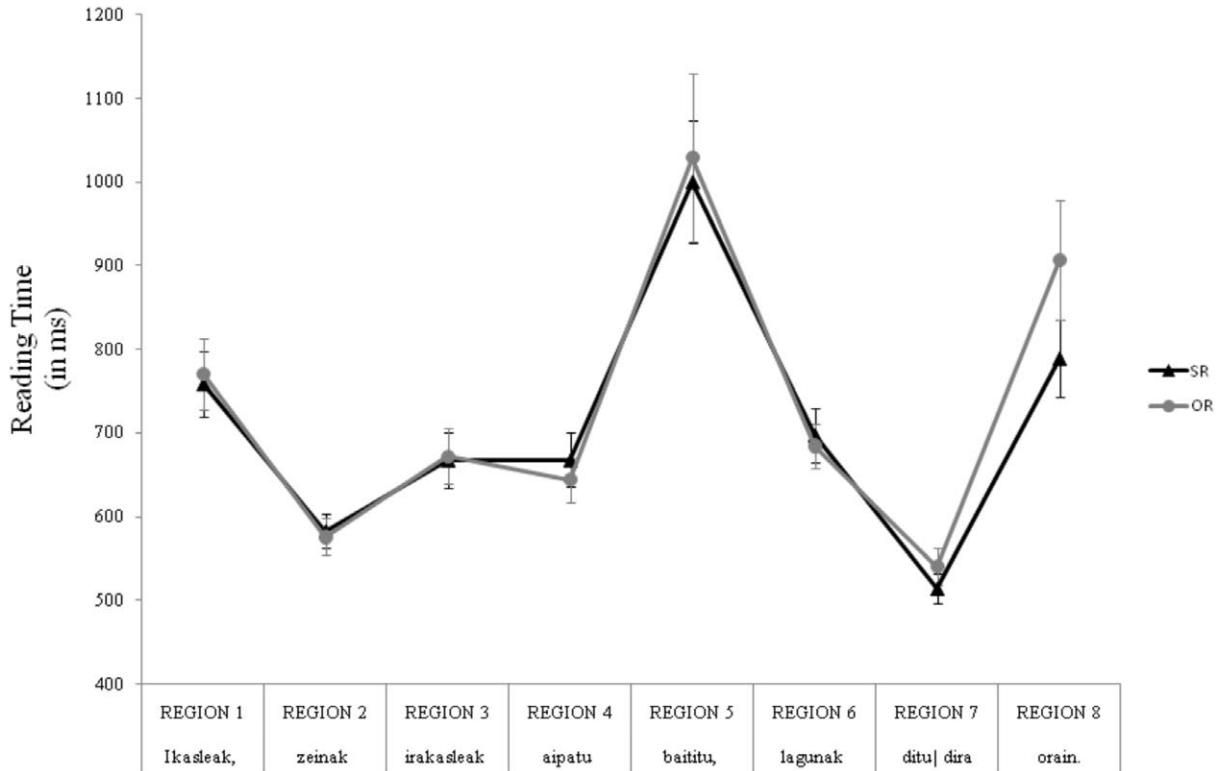
Participants' accuracy in comprehension questions and RTs for each of the eight regions was recorded. All RTs above or below 2.5 standard deviations from the mean were discarded for the analysis. Overall, less than 3% of the data was excluded. Mean RTs for each region were entered into pairwise *t*-tests for the factor RC type (SR vs. OR). Statistical analyses were performed over participants' and items' data (reported as t_1 and t_2 , respectively).

COMPREHENSION PERFORMANCE

A main effect of RC type was found exclusively in the analysis by participants: the mean comprehension accuracy was higher for SRs than for ORs (85.96% and 80.77% respectively), $t_1(39) = 2.63, p < .02, t_2(24) = .48, p > .60$.

READING TIMES

Figure 1 below plots the mean RTs per region for SR and OR clauses, together with the corresponding standard error bars. Results suggest a processing advantage for postnominal SRs over ORs: reading times were shorter for the SR condition at the critical disambiguating region (REGION 7) $t_1(39) = 2.03, p < .05, t_2(24) = 1.86, p = .07$, and this effect persisted into the post-verbal region (REGION 8), where the effect was largest, $t_1(39) = 2.18, p < .04, t_2(24) = 2.39, p < .03$. It is frequent in self-paced reading experiments that an effect is found in the region immediately following the critical word, especially when the critical word is short as is the case with the disambiguating verbs *dira/ditu* (Mitchell, 1984). As expected, statistical comparisons on the previous regions revealed no significant difference between subject and object RCs (all t s < 1.1 and all p s > .30).



4. Discussion

Here we report a self-paced reading experiment with Basque postnominal RCs, in which reading times showed a significant effect of relative clause type: SRs were read faster than ORs. The current study aimed at contributing to a better understanding of RC processing preferences across languages, and addressed the question of whether word order modulates the OR advantage found for prenominal RCs in Basque, an ergative and head-final language. To this end, a self-paced reading experiment evaluated how the position of the RC relative to the head noun might influence the RC processing asymmetries in Basque, a language that allows not only for prenominal but also for postnominal RCs. The results showed opposed subject and object processing asymmetries for prenominal and postnominal RCs in Basque.

The data cannot be explained just appealing to a general advantage for unmarked (i.e., absolutive) case-marking of the gapped element, for this processing strategy would equally favor an OR advantage both in prenominal as well as in postnominal Basque RCs. Instead, the SR advantage of postnominals is compatible with the idea that Basque speakers preferably interpret a sentence-initial argument as ergative rather than as absolutive. As suggested by Laka (2013), this can be understood as a strong subject-first preference in ergative languages, as ergative case unequivocally marks subjects.

This is not to say that a preference for unmarkedness should be ruled out based on the current data from Basque postnominal RCs. In this respect, future work should investigate processing asymmetry effects with RCs containing unambiguous initial arguments, and thus determine whether the same subject/object asymmetries obtain for prenominal RCs in the absence of garden-

path effects provoked by initial disambiguation choices. If in that case the OR advantage for prenominal RCs becomes diluted, psycholinguistic research should start considering that the same processing asymmetries could be extrapolated to simple declarative sentences, and that consequently, a simple clause might be equally complex as regards processing when there is competition between two animate arguments. This is, of course, something that should be further explored.

Appendix: experimental items

1. Ikasleak, zeinak irakasleak aipatu baititu, lagunak dira/ditu orain.
 «The students, who the teacher mentioned, are friends now». (OR)
 «The student, who mentioned the teachers, has friends now». (SR)
2. Zuzentzaileak, zeinak itzultzaileak kontratatu baititu, profesionalak dira/ditu dirudienez.
 «The reviewers, who the translator hired, are professionals apparently». (OR)
 «The reviewer, who hired the translators, has professionals apparently». (SR)
3. Ilobak, zeinak osabak maite baititu, medikuak dira/ditu ospitalean.
 «The nephews, who their uncle loves, are doctors at the hospital». (OR)
 «The nephew, who loves his uncle, has doctors at hospital». (SR)
4. Gaixoak, zeinak medikuak bisitatu baititu, adiskideak dira/ditu gaztetandik.
 «The patients, who the doctor visited, are friends since they were young». (OR)
 «The patient, who visited the doctors, has friends since s/he was young». (SR)
5. Notarioak, zeinak abokatuak lasaitu baititu, kolaboratzaileak dira/ditu oraindik.
 «The solicitors, who the lawyer calmed down, are still collaborators». (OR)
 «The solicitor, who calmed down the lawyers, still has collaborators». (SR)
6. Artistak, zeinak antzezleak aurkeztu baititu, amoranteak dira/ditu aspalditik.
 «The artists, who the actor introduced, have been lovers for a long time». (OR)
 «The artist, who introduced the actors, has lovers since long ago». (SR)
7. Gizonak, zeinak emakumeak laztandu baititu, maitaleak dira/ditu diotenez.
 «The men, who the woman caressed, are lovers, so they say». (OR)
 «The man, who caressed the women, has lovers, so they say». (SR)
8. Neskatzak, zeinak andreak apaindu baititu, laguntzaileak dira/ditu ileapaindegian.
 «The girls, to whom the woman did the haircuts, are assistants at the hairdresser's». (OR)
 «The girl, who did the haircut to the women, has assistants at the hairdresser's». (SR)
9. Informatikariak, zeinak teknikariak ezagutu baititu, langileak dira/ditu enpresan.
 «The computer technicians, who the engineer met, are employees at the company». (OR)
 «The computer technician, who met the engineers, has employees at the company». (SR)

10. Kantariak, zeinak dantzariak ikusi baititu, irakasleak dira/ditu eskolan.
«The singers, who the dancer saw, are teachers at school». (OR)
«The singer, who saw the dancers, has teachers at school». (SR)
11. Bizilagunak, zeinak etorkinak salbatu baititu, pisukideak dira/ditu orain.
«The neighbours, who the immigrant saved, are flat mates now». (OR)
«The neighbour, who saved the immigrants, has flat mates now». (SR)
12. Psikiatrak, zeinak psikologoak kritikatu baititu, lankideak dira/ditu ospitalean.
«The psychiatrists, who the psychologist criticized, are colleagues at the hospital». (OR)
«The psychiatrist, who criticized the psychologists, has colleagues at the hospital». (SR)
13. Suhiltzaileak, zeinak poliziak babestu baititu, anaiak dira/ditu nonbait.
«The firefighters, who the policeman protected, are brothers apparently». (OR)
«The firefighter, who protected the policemen, has brothers apparently». (SR)
14. Haurrak, zeinak mutilak engainatu baititu, bizilagunak dira/ditu ziurrenik.
«The children, who the boy fooled, are neighbors probably». (OR)
«The child, who fooled the boys, has neighbors probably». (SR)
15. Zinegileak, zeinak aktoreak txalotu baititu, jarraitzaileak dira/ditu iaztik.
«The directors, who the actor praised, are followers since last year». (OR)
«The director, who praised the actors, has followers since last year». (SR)
16. Zientzialariak, zeinak ikertzaileak nahi baititu, ikasleak dira/ditu unibertsitatean.
«The scientists, who the researcher wants, are students at the university». (OR)
«The scientist, who wants researchers, has students at the university». (SR)
17. Jokalariak, zeinak atezainak bultzatu baititu, aurkakoak dira/ditu futbol-zelaian.
«The players, who the goalkeeper pushed, are opponents on the soccer field». (OR)
«The player, who pushed the goalkeepers, has opponents on the soccer field». (SR)
18. Saltzaileak, zeinak erosleak kontsolatu baititu, senitartekoak dira/ditu agian.
«The sellers, who the buyer comforted, are relatives maybe». (OR)
«The seller, who comforted the buyers, has relatives maybe». (SR)
19. Palestinarrak, zeinak israeldarrak tirokatu baititu, ezagunak dira/ditu bestaldean.
«The Palestinians, who the Israeli shot, are well-known on the other side». (OR)
«The Palestinian, who shot the Israelis, has acquaintances on the other side». (SR)
20. Agureak, zeinak gazteak zaindu baititu, senideak dira/ditu beharbada.
«The elderly men, for whom the young person took care, are relatives maybe». (OR)
«The elderly man, who took care of the young people, has relatives maybe». (SR)
21. Mutikoak, zeinak neskatilak izutu baititu, begiratzzaileak dira/ditu udalekuetan.
«The boys, who the girl scared, are tutors at the summer camp». (OR)
«The boy, who scared the girls, has tutors at the summer camp». (SR)

22. Neskak, zeinak mutilak lotsatu baititu, lagunak dira/ditu benetan.
 «The girls, who the boy embarrassed, are friends really». (OR)
 «The girl, who embarrassed the boys, has friends really». (SR)
23. Umeak, zeinak lehengusinak zipriztindu baititu, anai-arrebak dira/ditu antza.
 «The children, who their cousin splashed (with water), are siblings seemingly». (OR)
 «The child, who splashed his/her cousins (with water), has siblings seemingly». (SR)
24. Zuzendariak, zeinak arduradunak salatu baititu, entzuleak dira/ditu prentsasurekoan.
 «The directors, who the manager accused, are listeners at the press conference». (OR)
 «The director, who accused the managers, has listeners at the press conference». (SR)
25. Manifestatzaileak, zeinak bakezaleak besarkatu baititu, bazkideak dira/ditu erakundean.
 «The demonstrators, who the pacifist hugged, are members at the organization». (OR)
 «The demonstrator, who hugged the pacifists, has members at the organization». (SR)
26. Idazleak, zeinak kazetariak hunkitu baititu, irakurleak dira/ditu jakina.
 «The writers, who the journalist moved, are readers of course». (OR)
 «The writer, who moved the journalists, has readers of course». (SR)

Acknowledgements

This work was funded by the Basque Government's Department of Education, Universities, and Research (IT1169-19).

References

- Aaronson, D., & Ferres, S. (1984). The word-by-word reading paradigm: An experimental and theoretical approach. In D. Kieras, & M. Just (Eds.), *New methods in reading comprehension research*. Hillsdale, NJ: Erlbaum.
- Betancort, M., Carreiras, M., & Sturt, P. (2009). The processing of subject and object relative clauses in Spanish: An eye-tracking study. *The quarterly Journal of Experimental Psychology*, 62 (10), 1915-1929.
- Bornkessel, I. & Schlesewsky, M. (2006). The Extended Argument Dependency Model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, 113, 787-821.
- Caplan, D., Vijayan, S., Kuperberg, G., West, C., Waters, G., Greve, D., & Dale, A. M. (2001). Vascular responses to syntactic processing: event-related fMRI study of relative clauses. *Human Brain Mapping*, 15, 26-38.
- Carreiras, M., Duñabeitia, J. A., Vergara, M., de la Cruz-Pavía, I., & Laka, I. (2010). Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition* 115 (1), 79-92.
- Cohen, L., & Mehler, J. (1996). Click monitoring revisited: An on-line study of sentence comprehension. *Memory and Cognition*, 24, 94-102.
- De Vincenzi, M. (1991). *Syntactic parsing strategies in Italian*. Dordrecht: Kluwer.
- Demiral, S.B., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2008). On the universality of language comprehension strategies: Evidence from Turkish. *Cognition*, 106, 484-500.

- Dixon, R. M. W. (1994). *Ergativity*. Cambridge: Cambridge University Press.
- Frauenfelder, U., Segui, J., & Mehler, J. (1980). Monitoring around the relative clause. *Journal of Verbal Learning and Verbal Behavior*, 19, 328-337.
- Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistics Theory*, 5, 519-559.
- Gibson, E. & Wu, I. (2011). Processing Chinese relative clauses in context. *Language and Cognitive Processes*. First published on: 07 March 2011 (iFirst).
- Gouvea, A. C. (2003). Processing Syntactic Complexity: Cross-Linguistic Differences and ERP Evidence. University of Maryland, College Park.
- Holmes, V. M., & O'Regan, J. K. (1981) Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior*, 20, 417-430.
- Hsiao, F. & Gibson, E. (2003) Processing relative clause in Chinese. *Cognition*, 90, 3-27.
- Ishizuka, T. (2005). Processing Relative Clauses in Japanese. *UCLA Working Papers in Psycholinguistics*, 2, 135-157.
- Ishizuka, T., Nakatani, K., & Gibson, E. (2003). Relative clause extraction complexity in Japanese. Poster presented at the 16th annual CUNY conference on human sentence processing, March 27-29, Massachusetts Institute of Technology, Cambridge, MA.
- Ishizuka, T., Nakatani, K., & Gibson, E. (2006). Processing Japanese relative clauses in context. Paper presented at the 19th Annual CUNY Conference on Human Sentence Processing, CUNY, New York, March, 2006.
- Just, M. A., Carpenter, P., & Wooley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology, General*, 3, 228-238.
- Keenan, E. L. & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8, 63-99.
- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580-602.
- King, J., & Kutas, M. (1995). Who did what and when? Using word- and cause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, 7, 376-395.
- Kuo, K. & Vasisht, S. (2006) Processing Chinese relative clauses: Evidence for the universal subject preference. Manuscript, University of Potsdam.
- Kwon, N., Kluender, R., Kutas, M., & Polinsky, M.. Under revision. *Subject/Object processing asymmetries in Korean relative clauses: Evidence from ERP data*.
- Kwon, N., Lee, Y., Gordon, P., Kluender, R., & Polinsky, M. (2010). Cognitive and linguistic factors affecting of the subject/object asymmetry: An eyetracking study of pre-nominal relative clauses in Korean. *Language*, 83 (3), 546-583.
- Kwon, N., Polinsky, M., & Kluender, R. (2006) Subject preference in Korean. In D. Baumer, D. Montero, and M. Scanlon (Eds.) *Proceedings of the 25th West Coast Conference on Formal Linguistics* (pp. 1-14). Somerville, MA: Cascadilla Proceedings Project.
- Laka, I. (2013). Merging from the temporal input: on subject-object asymmetries and an ergative language. In M. Piattelli-Palmarini & R. Berwick (eds.) *Rich Languages from Poor Inputs*, Oxford: Oxford University Press, pp. 127-145.
- Lin, C.-J. C. & Bever, T. G. (2006b). Subject Preference in the Processing of relative Clauses in Chinese. In D. Baumer, D. Montero, and M. Scanlon (Eds.) *Proceedings of the 25th West Coast Conference on Formal Linguistics* (pp. 254-260). Somerville, MA: Cascadilla Proceedings Project.
- Lin, C.-J. C. & Bever, T.G. (2006a). Chinese is no exception: Universal subject preference of relative clause processing. Paper presented at *The 19th Annual CUNY Conference on Human Sentence Processing*, CUNY Graduate Center, New York, NY.

- Lin, Y. B. & Garnsey, S.M. (2007). Plausibility and the resolution of temporary ambiguity in relative clause comprehension in mandarin. In *Proceedings of the CUNY Sentence Processing Conference*.
- Mak, W. M., Vonk, W., & Schriefers, H. (2002). The Influence of Animacy on Relative Clause Processing. *Journal of Memory and Language*, 47, 50-68.
- Mak, W. M., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, 54, 466-490.
- Mecklinger, A., Schriefers, H., Steinhauer, K. & Friederici, D. (1995) Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory and Cognition*, 23, 477-494.
- Mitchell, D. C. 1984. An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In *New methods in reading comprehension research*, ed. D. Kieras and M.A. Just, 6989. Hillsdale, NJ: Earlbaum.
- Otsuka, Y. (2006). Syntactic ergativity in Tongan. In A. Johns, D. Massam, & J. Ndayiragije (Eds.), *Ergativity: emerging issues* (79-108). Studies in Natural Language and Linguistic Theory, 65. Springer, Netherlands: Dordrecht.
- Oyharçabal, B. (2003) Relatives. In I. Hualde, & J. Ortiz de Urbina (Eds.), *A Grammar of Basque* (pp. 762-823). Berlin: Mouton de Gruyter.
- Perea, M., Urkia, M., Davis, C.J., Agirre, A., Laseka, E., & Carreiras, M. (2006). E-Hitz: A word-frequency list and a program for deriving psycholinguistic statistics in an agglutinative language (Basque). *Behavior Research Methods*, 38, 610-615.
- Polinsky, M., Gómez-Gallo, C., Graff, P., & Kravtchenko, E. (2011). Subject preference and ergativity. *Lingua*, 121.
- Sauerland, U. & Gibson, E. (1998). How to predict the relative clause attachment preference. Paper presented at the 11th CUNY sentence processing conference. New Brunswick NJ: Rutgers University.
- Schriefers, H., Friederici, A. D., & Kühn, K. (1995) The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language*, 34, 499-520.
- Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: evidence from eye movements. *Journal of Memory and Language*, 47, 69-90.
- Ueno, M., & Garnsey, S. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes*, 23, 646-688.
- Wang, L., Schlesewsky, M., Bickel, B., & Bornkessel-Schlesewsky, I. (2009). Exploring the nature of the «subject»-preference: Evidence from the online comprehension of simple sentences in Mandarin Chinese. *Language and Cognitive Processes*, 24 (7-8), 1180-1226.
- Yamashita, H. (1997). The effects of word order and case marking information on the processing of Japanese. *Journal of Psycholinguistic Research*, 26, 163-188.
- Yang, C. L., Perfetti, C. A., & Liu, Y. (2010). Sentence integration processes: An ERP study of Chinese sentence comprehension with relative clauses. *Brain and Language*, 112(2), 85-100.

Sentence processing in Andean and Peninsular Spanish: word order and animacy

NATALIA YOZA, KEPA ERDOCIA, ITZIAR LAKA

Linguistics and Basque Studies Department
University of the Basque Country (UPV/EHU)

Abstract

We report pioneering experimental evidence for sentence processing in Andean Spanish. Studies report that preverbal objects are more frequent in the Andean variety (A. Escobar, 1978; Lujan, Minaya and Sankoff, 1981; Muysken, 1984; Cerron-Palomino, 1990; A.M. Escobar, 1990, 2000; Klee, 1996; Zavala 1999; Muntendam, 2013). We compared the comprehension and reading times for SVO and OVS sentences by Andean and Peninsular Spanish speakers. SVO was fastest in both varieties, replicating previous studies on Peninsular Spanish (Basilico *et al.*, 1995; Casado *et al.*, 2005 and Del Rio *et al.*, 2012) and providing novel evidence that, despite the relatively higher frequency of preverbal objects, canonical word order in Andean Spanish is also SVO. We also considered the impact of object animacy in sentence processing: in both Andean and Peninsular varieties, object animacy results in a general slowdown and poorer comprehension, replicating previous findings. We provide the first online evidence involving Andean Spanish, and show that, despite their higher frequency, preverbal objects are processed with difficulty and SVO is canonical also in this variety.

Keywords: Andean Spanish, preverbal object, word order, animacy, sentence processing, reading times.

1. Introduction

Many languages allow variations of sentence word order, but each has a canonical order which is the most frequent and corresponds to the linearization of the simplest sentence structure generated by the grammar (Chomsky, 1981; Kaan, 1997; Lee & Kim, 2012). Most languages have either Subject-Verb-Object (SVO) or Subject-Object-Verb (SOV) as canonical (Dryer 2013). Spanish and English are examples of SVO languages, while Quechua and Japanese are SOV (Greenberg, 1963).

In Peninsular Spanish, canonical SVO sentences (1a) are the most frequent (Muysken, 1984; Ocampo, 1994; Klee, 1996; Muntendam, 2013), whereas OVS sentences (1b) are infrequent and both syntactically and semantically more complex (Laka 1990; Zubizarreta, 2000).

- (1) a. María compró pocas manzanas.
María bought few apples.
«Mary bought few apples»
- b. Pocas manzanas compró María.
Few apples bought María
«Few apples is what Mary bought»

Experimental studies show that canonical sentences incur less processing costs; in Self-Paced Reading studies, shorter reading times have been reported for canonical word orders in languages like German (Weyerts, Penke, Münte, Heinze & Clahsen, 2002; Bahlmann, Münte, Rodríguez-Fornells & Rotte, 2007), Finnish (Kaiser & Trueswell, 2004), Japanese (Wolff, Schlesewsky, Hirotani & Bornkessel-Schlesewsky, 2008), Basque (Erdocia, Laka, Mestres-Misse & Rodríguez-Fornells, 2009), Russian (Sekerina, 2003; Levy, Feodorenko & Gibson, 2013), and Hindi (Vasishth, 2002), among others.

Why are canonical sentences processed fast and easy? Three non-mutually exclusive theories explain it. According to the *Derivational Theory of Complexity* (Chomsky & Miller, 1963; Erdocia, Laka & Rodríguez-Fornells, 2012; Phillips, 2013) sentences with longer syntactic derivations involve more computational steps and this is why they incur greater processing demands; since object initial sentences in VO languages involve fronting of the object, this constituent must be held in memory until integrated with its base position (Wolff *et al.*, 2008). The *Subject First* preference (Bever, 1970; Bornkessel & Schlesewsky, 2009) also predicts that sentences that are not subject initial incur in greater processing costs because the comprehension system prefers to process Agent initial sequences. Finally, frequency based theories (Hale, 2001; Levy, 2008) show there is a correlation between frequency and processing ease, which in the case of sentence word order predicts an advantage for canonical word orders.

Accordingly, in Spanish, experimental studies reveal that SVO is the easiest word order to process: In an ERP experiment comparing SVO and OVS sentences, Casado, Fernández-Frías, Martín-Loeches & Muñoz (2005) found that SVO sentences were easiest to process. Similarly, Del Rio, Lopez-Higes & Martín-Aragoneses (2012), measuring reading times, found that object (OVS order) relative clauses required longer reading times than subject (SVO order) relative clauses, further indicating that canonical word order in Spanish is SVO.

Casado *et al.* (2005) measured the electrophysiological brain responses of Spanish speakers while reading SVO and OVS sentences with animate and inanimate objects. When comparing O-inanimate SVO and OVS conditions, the OVS condition elicited a negative component at sentence initial and final positions and a P600 component at verb position, signaling greater memory storage and reintegration cost for the displaced object. When comparing O-animate SVO and OVS conditions, a P600 component was observed at the second DP of sentence-final S as compared to O, also revealing that canonical word order in Spanish is SVO.

Andean Spanish is spoken in social and geographic contexts with a presence or influence of Quechua, Aymara or some other Andean language (Godenzzi, 1991). Studies of this variety report that preverbal objects are more frequent than in other varieties of Spanish (Muysken, 1984; Klee 1996; Muntendam, 2013), as summarized in the following table:

Table 1

Percentages of VO/OV in Non-Andean Spanish and Andean Spanish

	Non-Andean Spanish	Andean Spanish		
	Ocampo (1994)	Muysken (1984)	Klee (1996)	Muntendam (2013)
VO	92.1%	80%	81.7%	81.5%
OV	7.9%	20%	18.3%	18.5%

Sources: For Ocampo (1994): Spanish monolinguals from Buenos Aires (Argentina); for Muysken (1984): Quechua-Spanish bilinguals from Ecuador; for Klee (1996): Quechua-Spanish bilinguals from Calca (Peru); and for Muntendam (2009): Quechua-Spanish bilinguals from Tarata (Bolivia) and Juncal (Ecuador). Adapted from Muntendam (2013).

Preverbal objects can express repetition, summary, agreement or explanation (Klee, 1996; Camacho, 1999; Sanchez, 2003). An example is provided in (3):

- (3) A la chica he visto en misa.
 DOM the girl have-I seen at mass.
 «I have seen the girl at mass» (from A. Escobar, 1978)

The higher frequency of preverbal objects reported for Andean Spanish could indicate: (a) that speakers of this variety process sentences with preverbal objects with greater ease than speakers of Peninsular Spanish (where preverbal objects are rare), or (b) that canonical word order in Andean Spanish is not strictly SVO and is heavily modulated by the OV nature of indigenous contact languages. To the extent of our knowledge, experimental methodology has not been used to study Andean Spanish. Here, we aim to broaden our knowledge about sentence processing in Spanish across different varieties providing online processing measures. Specifically, we seek to test whether Andean and Peninsular Spanish differ in the cognitive cost involved in processing preverbal objects and whether object animacy modulates this cost. To this end, we compared reading times of SVO and OVS sentences with and without animate objects by means of a Self-Paced Reading task with two groups of monolingual speakers: Peninsular and Andean.

2. The experiment

We originally took the experimental design in Casado *et al.* (2005) about canonical word order in (Peninsular) Spanish as the base of our materials. We hypothesized that SVO will be processed faster and with less comprehension errors in comparison to OVS in the Peninsular Spanish group, thus replicating the results obtained by Casado *et al.* (2005). We also expect larger processing cost in sentences with animate objects as compared to sentences with inanimate objects because object animacy has been reported to make thematic role assignment more taxing cross-linguistically (Grewe, Bornkessel-Schlesewsky, Zysset, Wiese, von Cramon & Schlesewsky, 2007; Nieuwland, Martin, Carreiras, 2013). As for the Andean Spanish group, given the higher frequencies consistently reported for preverbal objects in this variety, we expect to find that object initial sentences incur in less processing costs than in Peninsular Spanish, or alternatively, if the greater frequency of object initial sentences reported for Andean Spanish is not significant enough to induce changes in sentence representation, we should find similar reaction times for both

groups of participants across conditions. Finally, since the O-animate condition carries a specific Differential Object Marking (DOM) ‘*a*’, we particularly expect increasing processing cost for O_{animate} VS sentences, since their marking signals them unambiguously as object initial from the start of the sentence.

2.1. Method

2.1.1. PARTICIPANTS

32 monolingual speakers of Peninsular Spanish (mean age 21.75; 21 women) all of them undergraduate students at the University of the Basque Country (EHU/UPV) in Vitoria-Gasteiz (Spain) and 32 monolingual speakers of Andean Spanish (mean age 21.06; 22 women) all undergraduate students at the National University of San Agustín in Arequipa (Perú) participated in the experiment. All of them consented to participate and were given some money for their time and effort.

2.1.2. MATERIALS

Stimuli consisted of 180 transitive sentences in SVO and OVS orders with animate and inanimate objects. There were 4 conditions:

- (4) a. *Condition 1: SVO sentence with animate object* (SVO_(an))
El inquilino alarmó al casero.
The tenant alarmed DOM.the landlord.
«The tenant alarmed the landlord».
- b. *Condition 2: SVO sentence with inanimate object* (SVO_(in))
El perro saltó la valla.
The dog jumped the fence.
«The dog jumped the fence».
- c. *Condition 3: OVS sentence with animate object* (O_(an)SV)
Al casero alarmo el inquilino.
DOM.the landlord alarmed the tenant
«The tenant alarmed the landlord».
- d. *Condition 4: OVS sentence with inanimate object* (O_(in)SV)
La valla saltó el perro.
The fence jumped the dog.
«The dog jumped the fence».

Conditions 3 and 4 were generated from conditions 1 and 2 respectively by reversing the order of Os and Ss. We created two lists in order to counterbalance the experimental material across participants and to prevent them from reading the same version of a sentence twice. Both lists contained 45 experimental sentences from each condition. We also included the same 50 filler sentences in both lists. All in all, each participant read 230 sentences.

Our experimental material was originally generated based on the materials used in Casado *et al.* (2005), but we had to introduce a few changes we discuss here: Their materials contained animate objects without the DOM marker *a*, yielding ungrammatical sentences in Spanish (for

instance, **el poeta desafió el novelista*) as discussed already by Demestre (2012). Accordingly, all our animate objects had the required DOM *a* marker (for instance, *al poeta desafió el novelista*). We substituted ditransitive and psychological verbs (like *obligar* «to force» or *disgustar* «dislike»), which involve experiencers and goals and hence more complex argument structures with simple transitive verbs involving agents and themes in their argument structure. A few lexical items used exclusively in Spain were substituted for words that are used also in Perú. For example, instead of *marido* «husband» (only Peninsular Spanish) we used *esposo* (both Peninsular and Peruvian Spanish)¹. In total we changed 46 items (30 nouns and 16 verbs). We also measured the length and the frequency of the lexical items in our experiment using *EsPal* (Duchon, Perea, Sebastián-Gallés, Martí & Carreiras, 2013). The verbs in the animate and inanimate conditions differ in length ($SVO_{(in)} = 6.61$; $SVO_{(an)} = 7.68$; $t(89) = -5.443$, $p < 0.001$) and frequency ($SVO_{(in)} = 19.91$, $SVO_{(an)} = 6.10$; $t(89) = 4.958$, $p < 0.001$). Inanimate objects were significantly shorter than animate objects ($SVO_{(in)} = 6.08$, $SVO_{(an)} = 7.71$; $t(89) = -6.962$, $p < 0.001$). In inanimate conditions, subjects contained more letters than objects ($S = 7$, $O = 6.1$; $t(89) = 3.618$, $p < 0.001$), and in animate conditions, the opposite ($S = 7.12$, $O = 7.71$, $t(89) = -2.469$, $p = 0.015$).

In order to keep participants' attention and determine how they interpreted the sentences, we included a comprehension question after half of the experimental sentences². The comprehension question was related to the previously read sentence and the answers of half of the questions of each condition was *correct* (see example 9 for *correct* answers). To answer the question, participants had to press the button corresponding to P and Q on the keyboard. We manipulated the word order of the comprehension questions in order to prevent any bias generated by the higher proportion of a given word order during the experiment. Thus, half of the comprehension questions of each experimental condition were SVO and the other half OVS. Participants received feedback on their answers. We recorded the accuracy and the reaction times of their responses. Half of the filler sentences were also followed by a comprehension question.

- (5) a. *SVO_(in) condition*
 El chico perdió un juguete.
 «The boy lost a toy».
- b. *SVO question*
 ¿Es verdad que el chico perdió un lápiz?
 Is (it) true that the boy (S) lost a pencil (O)?
 «Is it true that the boy lost a pencil?»
- c. *OVS question*
 ¿Es verdad que un lápiz perdió el chico?
 Is (it) true that a pencil (O) lost the boy (S)?
 «Is it true that a pencil lost the boy?»

2.1.3. PROCEDURE

Experimental materials and comprehension task were presented to the participants through Self Paced Reading (Just, Carpenter & Woolley 1982) using the *Linger* program (Rohde,

¹ All changes introduced in the materials are reported in the appendix.

² The 45 experimental items could not be divided by 2, thus we distributed the CT as follows: 22 questions were about inanimate object conditions (SVO and OVS) and 23 were about animate object conditions (SVO and OVS).

2001). We presented the sentences constituent by constituent by means of the Moving Window Technique (Just *et al.*, 1982). Participants read in a normal and comfortable pace. In order to move from one constituent to the next, participants pressed the space bar of the computer keyboard. The time they took to press the space bar was recorded by the program. Self-Paced Reading provides accurate reaction time measures of the participants' reading process, allowing us to infer the computations that happen in the participants' minds while they are processing each condition.

Before starting the experiment, participants performed a short training of 10 trial sentences with comprehension questions for half of them. For the Peninsular Spanish group, the experiment was conducted in the *Experimental Linguistics Laboratory* at the University of the Basque Country (EHU/UPV) in Vitoria-Gasteiz, Spain. For the Andean Spanish group, the experiment was conducted in Arequipa, Peru, at the Faculty of Economy of National University of San Agustín.

2.2. Results

2.2.1. WHOLE SENTENCE READING TIMES

For the Peninsular Spanish group, a two way repeated measures ANOVA was carried out to analyze whole sentence processing. The *Constituent Order* factor had a significant effect ($F(1,31) = 13.291$, $P(HF) = 0.001$), signaling that SVO sentences were read faster than OVS. Sentences with animate objects were read significantly slower than the sentences inanimate objects (*Animacy*, $F(1,31) = 33.875$, $P(HF) < 0.001$). Interaction between *Constituent Order* and *Animacy* factors did not reach significance ($F(1,31) = 3.051$, $P(HF) = 0.091$). As it can be seen in Figure 1 and Table 2, the $SVO_{(in)}$ condition was read the fastest. The pairwise comparison of that condition with the rest of conditions is significant ($SVO_{(in)}$ vs. $SVO_{(an)}$: $t(31) = -5.192$, $p < 0.001$; $SVO_{(in)}$ vs. $O_{(in)}$ VS: $t(31) = -2.804$, $p = 0.009$; $SVO_{(in)}$ vs. $O_{(an)}$ VS: $t(31) = -5.829$, $p < 0.001$). The slowest condition, as expected, was $O_{(an)}$ SV. The comparison between canonical $SVO_{(an)}$ vs. non canonical $O_{(an)}$ VS shows a canonical order advantage ($t(31) = -3.177$, $p = 0.03$). The comparison between $O_{(in)}$ VS and $O_{(an)}$ VS ($t(31) = -5.191$, $p < 0.001$) reveals a penalty for object animacy, as expected.

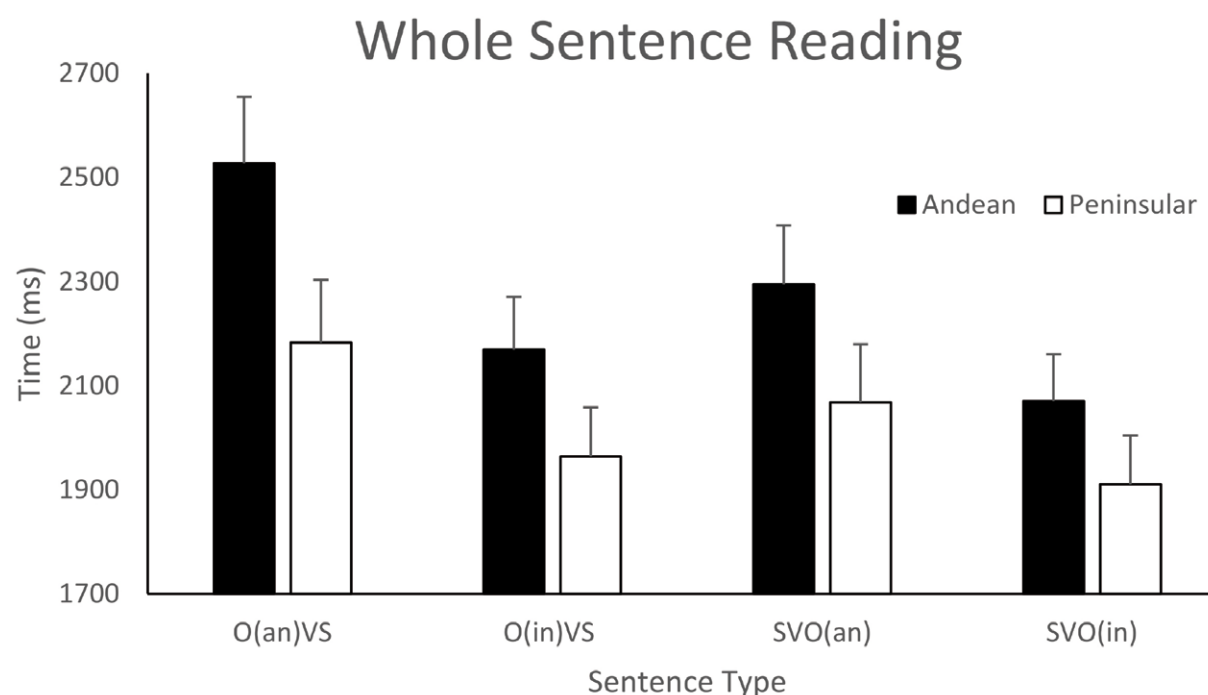
For the Andean Spanish group, we carried out two way repeated measures ANOVA to analyze whole sentence processing (see Table 2 and Fig. 1). The main result is that SVO sentences were read faster than OVS, as in the Peninsular group (*Constituent Order*, $F(1,31) = 33.120$, $P(HF) < 0.001$). Comparing the inanimate conditions, we observed that SVO was read faster than OVS ($SVO_{(in)}$ vs. $O_{(in)}$ VS: $t(31) = -3.149$, $p = 0.004$), and the same happened when we compared the two conditions with animate objects ($SVO_{(an)}$ vs. $O_{(an)}$ VS: $t(31) = -5.142$, $p < 0.001$). *Animacy* also resulted significant ($F(1,31) = 62.819$, $P(HF) < 0.001$), signaling that sentences with animate objects were read slower than sentences with inanimate objects. The interaction between the two factors reached significance ($F(1,31) = 6.655$, $P(HF) = 0.015$). Pairwise comparisons for the factor *Animacy* also showed that animate conditions took longer than inanimate conditions ($SVO_{(in)}$ vs. $SVO_{(an)}$: $t(31) = -5.110$, $p < 0.001$; $O_{(in)}$ VS vs. $O_{(an)}$ VS: $t(31) = -7.735$, $p < 0.001$; $SVO_{(in)}$ vs. $O_{(an)}$ VS: $t(31) = -8.317$, $p < 0.001$), as expected.

Table 2

Mean values of the reading times of the experimental sentences condition by condition.

Reaction times are provided in milliseconds (ms) and standard errors in parenthesis

	O _(an) VS	O _(in) VS	SVO _(an)	SVO _(in)
Peninsular Spanish	2182 ms (± 120)	1964 ms (± 94)	2067 ms (± 112)	1909 ms (± 94)
Andean Spanish	2527 ms (± 128)	2168 ms (± 102)	2294 ms (± 114)	2070 ms (± 89)

**FIGURE 1**

Whole sentence reading times of Andean Spanish speakers (black bars) and Peninsular Spanish speakers (white bars)

We also carried out a threeway ANOVA with the two groups, Andean Spanish vs. Peninsular Spanish as a between subject factor. The three way interaction between *Group*, *Constituent Order* and *Animacy* was not significant. We observed main effects for *Constituent Order* ($F(1,62) = 45.815$, $P(\text{HF}) < 0.001$) and for *Animacy* ($F(1,62) = 95.943$, $P(\text{HF}) < 0.001$), and the interaction between those two factors ($F(1,62) = 9.705$, $P(\text{HF}) = 0.003$) resulted significant, revealing that both groups of speakers processed the experimental conditions similarly. On the other hand, we observed two way significant interactions between *Animacy* and *Group* ($F(1,62) = 4.431$, $P(\text{HF}) = 0.039$) and *Constituent Order* and *Group* ($F(1,62) = 4.812$, $P(\text{HF}) = 0.032$). Importantly, the difference between animate and inanimate conditions was bigger for the Andean group than for the Peninsular group. In the same vein, the Andean group took longer to process O-initial conditions as compared to S initial conditions than did the Peninsular group.

2.2.2. COMPREHENSION TASK

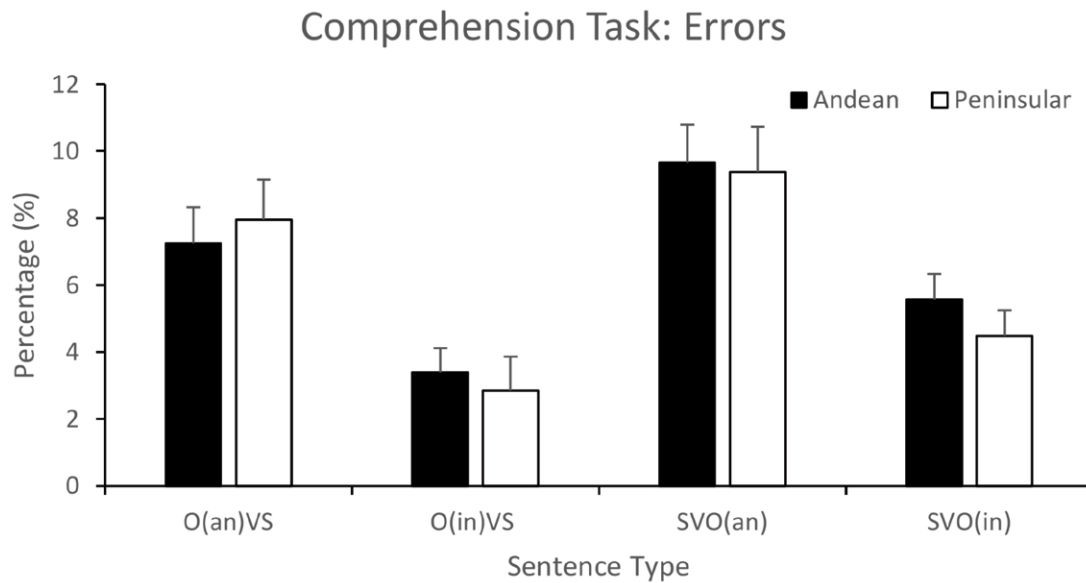
When comparing the error rates of the Peninsular group, we observed a main effect of *Animacy* ($F(1,31) = 23.881$, $P(\text{HF}) < 0.001$), revealing that sentences with animate objects generated more errors than those with inanimate objects (see table 3 and Fig. 2). The pairwise comparison showed that only animate-inanimate comparisons resulted significant ($\text{SVO}_{(\text{in})}$ vs. $\text{SVO}_{(\text{an})}$: $t(31) = 3.341$, $p = 0.002$; $\text{O}_{(\text{in})}\text{VS}$ vs. $\text{O}_{(\text{an})}\text{VS}$: $t(31) = 4.609$, $p < 0.001$; $\text{SVO}_{(\text{in})}$ vs. $\text{O}_{(\text{an})}\text{VS}$: $t(31) = -2.712$, $p = 0.011$; $\text{SVO}_{(\text{an})}$ vs. $\text{O}_{(\text{in})}\text{VS}$: $t(31) = 4.418$, $p < 0.001$). As for reaction times (see table 3 and Fig. 3), a similar scenario obtained, that is, a main effect of *Animacy* ($F(1,31) = 17.584$, $P(\text{HF}) < 0.001$). The pairwise comparisons also revealed that only animate-inanimate comparisons resulted significant ($\text{SVO}_{(\text{in})}$ vs. $\text{SVO}_{(\text{an})}$: $t(31) = 2.535$, $p = 0.017$; $\text{O}_{(\text{in})}\text{VS}$ vs. $\text{O}_{(\text{an})}\text{VS}$: $t(31) = -4.399$, $p < 0.001$; $\text{SVO}_{(\text{in})}$ vs. $\text{O}_{(\text{an})}\text{VS}$: $t(31) = -3.68$, $p = 0.004$; $\text{SVO}_{(\text{an})}$ vs. $\text{O}_{(\text{in})}\text{VS}$: $t(31) = 3.529$, $p = 0.001$).

Table 3

Mean values of the comprehension task (percentage of errors and reaction times, RT)
for Andean and Peninsular Spanish Speakers

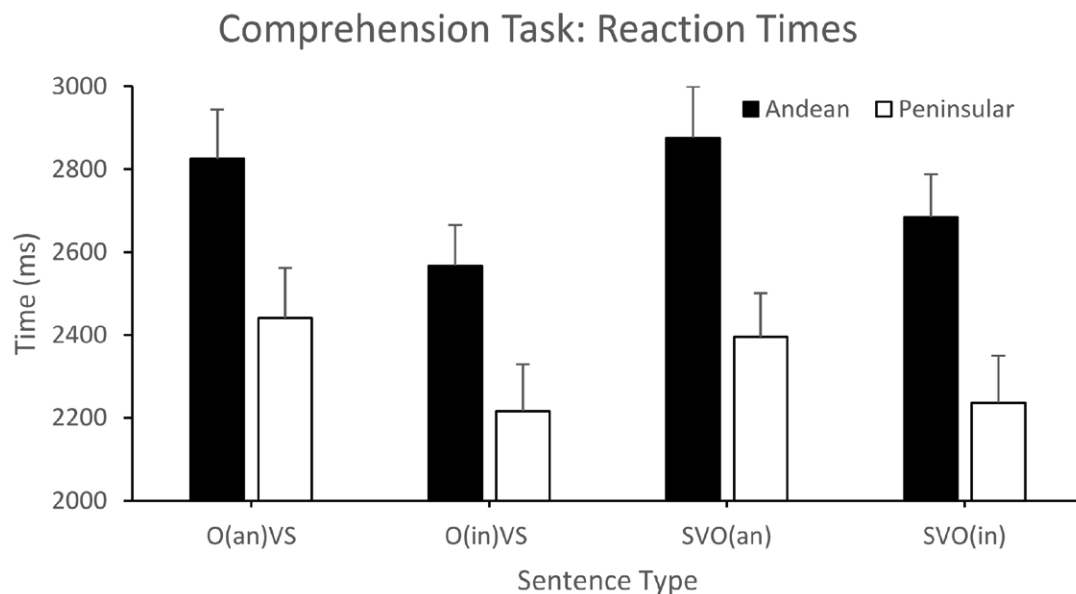
	$\text{O}_{(\text{an})}\text{VS}$		$\text{O}_{(\text{in})}\text{VS}$		$\text{SVO}_{(\text{an})}$		$\text{SVO}_{(\text{in})}$	
	Errors	RT	Errors	RT	Errors	RT	Errors	RT
Peninsular Spanish	7.9% (± 1.2)	2441 ms (± 121)	2.8% (± 1)	2216 ms (± 113)	9.4% (± 1.3)	2395 ms (± 106)	4.5% (± 0.8)	2236 ms (± 113)
Andean Spanish	7.2% (± 1)	2826 ms (± 127)	3.4% (± 0.7)	2567 ms (± 98)	9.7% (± 1.1)	2876 ms (± 123)	5.6% (± 0.8)	2684 ms (± 103)

Comparing the error rates of the Andean group, we observed a main effect of *Animacy* ($F(1,31) = 19.847$, $P(\text{HF}) < 0.001$) and another main effect of *Constituent order* ($F(1,31) = 8.097$, $P(\text{HF}) = 0.008$), but we did not observe any interaction between those factors (see table 3 and Fig. 2). The pairwise comparison showed that word order has an impact when objects were inanimate ($\text{SVO}_{(\text{in})}$ vs. $\text{O}_{(\text{in})}\text{VS}$: $t(31) = 2.184$, $p = 0.037$) but not when the objects were animate, due to the amount of errors observed in $\text{SVO}_{(\text{in})}$. Indeed, the animate-inanimate comparison that involved $\text{SVO}_{(\text{in})}$ and $\text{O}_{(\text{an})}\text{VS}$ conditions resulted non significant, while the rest of animate-inanimate comparisons were significant ($\text{SVO}_{(\text{in})}$ vs. $\text{SVO}_{(\text{an})}$: $t(31) = 3.205$, $p = 0.003$; $\text{O}_{(\text{in})}\text{VS}$ vs. $\text{O}_{(\text{an})}\text{VS}$: $t(31) = 3.253$, $p = 0.003$; $\text{SVO}_{(\text{an})}$ vs. $\text{O}_{(\text{in})}\text{VS}$: $t(31) = 5.062$, $p < 0.001$). For reaction times, a main effect of *Animacy* ($F(1,31) = 11.318$, $P(\text{HF}) = 0.002$) was observed (see table 3 and Fig. 3). However, apart of the animate-inanimate comparisons ($\text{SVO}_{(\text{in})}$ vs. $\text{SVO}_{(\text{an})}$: $t(31) = 2.019$, $p = 0.052$; $\text{O}_{(\text{in})}\text{VS}$ vs. $\text{O}_{(\text{an})}\text{VS}$: $t(31) = 3.893$, $p < 0.001$; $\text{SVO}_{(\text{in})}$ vs. $\text{O}_{(\text{an})}\text{VS}$: $t(31) = -2.062$, $p = 0.048$; $\text{SVO}_{(\text{an})}$ vs. $\text{O}_{(\text{in})}\text{VS}$: $t(31) = 3.446$, $p = 0.002$), $\text{SVO}_{(\text{in})}$ vs. $\text{O}_{(\text{in})}\text{VS}$ comparison also resulted significant ($t(31) = 2.194$, $p = 0.036$).

**FIGURE 2**

Percentage of errors in the comprehension task of Andean Spanish speakers (black bars) and Peninsular Spanish speakers (white bars)

Regarding the errors in the comprehension task, the group analysis revealed a main effect of *Constituent Order* ($F(1,62) = 9.641$, $P(\text{HF}) = 0.003$) and another main effect of *Animacy* ($F(1,62) = 43.707$, $P(\text{HF}) < 0.001$). No interaction was observed. Finally, the group analysis of the reaction times generated a main effect of *Animacy* ($F(1,61) = 26.459$, $P(\text{FH}) < 0.001$) without any significant interaction. The lack of interactions with *Group* factor signals that Peninsular and Andean Spanish speakers performed similarly in the comprehension task.

**FIGURE 3**

Reaction times of the comprehension task of Andean Spanish speakers (black bars) and Peninsular Spanish speakers (white bars)

2.2.3. CONSTITUENT BY CONSTITUENT ANALYSIS

Sentences were presented constituent by constituent in the experiment. Hence, we carried out pairwise comparisons of all experimental conditions at each sentence position (DP1, V, DP2) for both groups of participants.

For the Peninsular Spanish group, at sentence initial position (DP1), we observed that animate objects took longest, indicating they were hardest to process. They were slower than subjects and slower than inanimate objects ($O_{(in)}VS$ vs. $O_{(an)}VS$: $t(31) = -4.167$, $p < 0.001$; $SVO_{(an)}$ vs. $O_{(an)}VS$: $t(31) = -3.337$, $p = 0.002$; $SVO_{(in)}$ vs. $O_{(an)}VS$: $t(31) = -4.519$, $p < 0.001$; see Fig.4 and Table 4). Recall that animate objects in Spanish carry a DOM marker *a* (*María vio [o a la niña]* «María saw the girl»), which rules out a temporary parsing of this initial DP as a subject. But inanimate objects are not morphologically marked (*María vio [o la luz]* «María saw the light»), and they can be temporarily interpreted as subjects at sentence initial position (e.g. *la luz cegó a María* «The light blinded Mary»). Hence, the observed similar reading times are expected. Unexpectedly, the comparison of sentence initial inanimate objects with the subject in the $SVO_{(an)}$ condition was significant ($O_{(in)}VS$ vs. $SVO_{(an)}$: $t(31) = -2.683$, $p = 0.012$). We think this result is due to the fact that the inanimate objects in our materials were shorter than the subjects in the $SVO_{(an)}$ condition (see the discussion in the materials subsection above). Another unpredicted result was observed when comparing reading times of S in $SVO_{(in)}$ vs. $SVO_{(an)}$ conditions ($t(31) = -3.226$, $p = 0.003$), which differ in reading times though a revision of the items shows they did not differ in length or frequency.

At V (sentence second position), Peninsular Spanish speakers are slower in the O-animate condition than in the O-inanimate one ($SVO_{(in)}$ vs. $SVO_{(an)}$: $t(31) = -4.008$, $p < 0.001$; $O_{(in)}VS$ vs. $O_{(an)}VS$: $t(31) = -6.383$, $p < 0.001$; $O_{(in)}VS$ vs. $SVO_{(an)}$: $t(31) = -3.728$, $p = 0.001$). Further comparing reading times of V in the two O-animate conditions ($SVO_{(an)}$ vs. $O_{(an)}VS$), a statistically significant results was observed ($t(31) = -2.394$, $p = 0.023$), showing that sentence initial animate objects correlated with longest reading times at the verb.

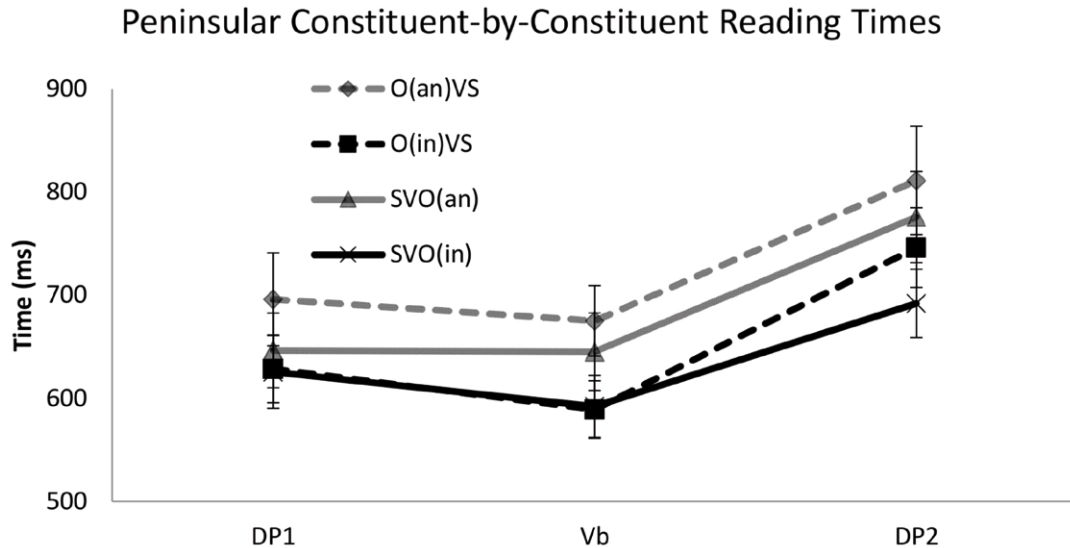


FIGURE 4

Constituent-by-constituent reading times of the Peninsular Spanish speakers at each constituent position in every experimental condition. DP1 refers to the first Determiner Phrase of the sentences, the subject of SVO conditions and the object of OVS conditions. DP2 refers to the last Determiner Phrase of the sentences, the object of SVO conditions and the subject of OVS conditions

At sentence final position (DP2), the largest processing cost was observed for the subject in the O-inanimate condition ($SVO_{(in)}$ vs. $O_{(in)}$ VS: $t(31) = -4.143$, $p < 0.001$). Further, comparisons between O-animate and O-inanimate conditions resulted significant at DP2 position ($SVO_{(in)}$ vs. $SVO_{(an)}$: $t(31) = -4.465$, $p < 0.001$; $O_{(in)}$ VS vs. $O_{(an)}$ VS: $t(31) = -2.609$, $p = 0.014$), with the exception of $O_{(in)}$ VS vs. $SVO_{(an)}$ comparison that was not significant, probably due to the increasing processing cost of the subject in final position in the O-inanimate condition (see Table 4 and Fig. 4).

Table 4

Constituent-by-Constituent mean reading times of the experimental conditions for both groups of participants (Peninsular and Andean). Time is provided in milliseconds, and standard errors in brackets.

DP1 refers to the first Determiner Phrase of the sentences, the subject of SVO conditions and the object of OVS conditions. DP2 refers to the last Determiner Phrase of the sentences, the object of SVO conditions and the subject of OVS conditions

	$O_{(an)}$ VS		$O_{(in)}$ VS		$SVO_{(an)}$		$SVO_{(in)}$	
	Penins.	Andean	Penins.	Andean	Penins.	Andean	Penins.	Andean
DP1	696 (± 45)	808 (± 43)	628 (± 32)	698 (± 33)	646 (± 36)	712 (± 34)	625 (± 35)	683 (± 30)
Verb	675 (± 34)	785 (± 42)	589 (± 27)	654 (± 28)	645 (± 37)	717 (± 34)	592 (± 30)	636 (± 28)
DP2	811 (± 52)	932 (± 51)	746 (± 38)	815 (± 47)	775 (± 44)	864 (± 49)	692 (± 33)	750 (± 36)

For the Andean Spanish group, comparisons at first constituent position (DP1) also revealed that the animate object was hardest to process ($O_{(in)}$ VS vs. $O_{(an)}$ VS: $t(31) = -5.365$, $p < 0.001$; $SVO_{(an)}$ vs. $O_{(an)}$ VS: $t(31) = -5.250$, $p < 0.001$; $SVO_{(in)}$ vs. $O_{(an)}$ VS: $t(31) = -6.303$, $p < 0.001$). As reported for the Peninsular group, we observed an unexpected difference when comparing sentence initial subjects ($SVO_{(in)}$ vs. $SVO_{(an)}$: $t(31) = -2.3568$, $p = 0.025$).

At V (sentence second) position, O-animate conditions required longer reading times than O-inanimate conditions ($SVO_{(in)}$ vs. $SVO_{(an)}$: $t(31) = -6.083$, $p < 0.001$; $O_{(in)}$ VS vs. $O_{(an)}$ VS: $t(31) = -7.192$, $p < 0.001$; $O_{(in)}$ VS vs. $SVO_{(an)}$: $t(31) = -5.532$, $p < 0.001$). We also observed a carry over effect at verb position in the $O_{(an)}$ VS condition when we compared it with the $SVO_{(an)}$ condition ($t(31) = -4.948$, $p < 0.001$), similar to the effect observed in Peninsular Spanish speakers (see Table 4 and Fig. 5).

At sentence final position (DP2), Andean speakers, like Peninsular ones also read objects faster than subjects in the O-inanimate condition ($SVO_{(in)}$ vs. $O_{(in)}$ VS: $t(31) = -3.335$, $p = 0.002$), an expected finding, signaling the extra processing effort required by sentence final subjects. Also in the O-animate condition, subjects took longer than objects ($SVO_{(an)}$ vs. $O_{(an)}$ VS: $t(31) = -2.635$, $p = 0.013$). In general, sentence final constituents in O-animate condition were read slower than in O-inanimate condition ($SVO_{(in)}$ vs. $SVO_{(an)}$: $t(31) = -4.057$, $p < 0.001$; $O_{(in)}$ VS vs. $O_{(an)}$ VS: $t(31) = -4.980$, $p < 0.001$). The only exception is the $O_{(in)}$ VS vs. $SVO_{(an)}$ comparison that was not significant, same as the Peninsular group (see Table 4 and Fig. 5).

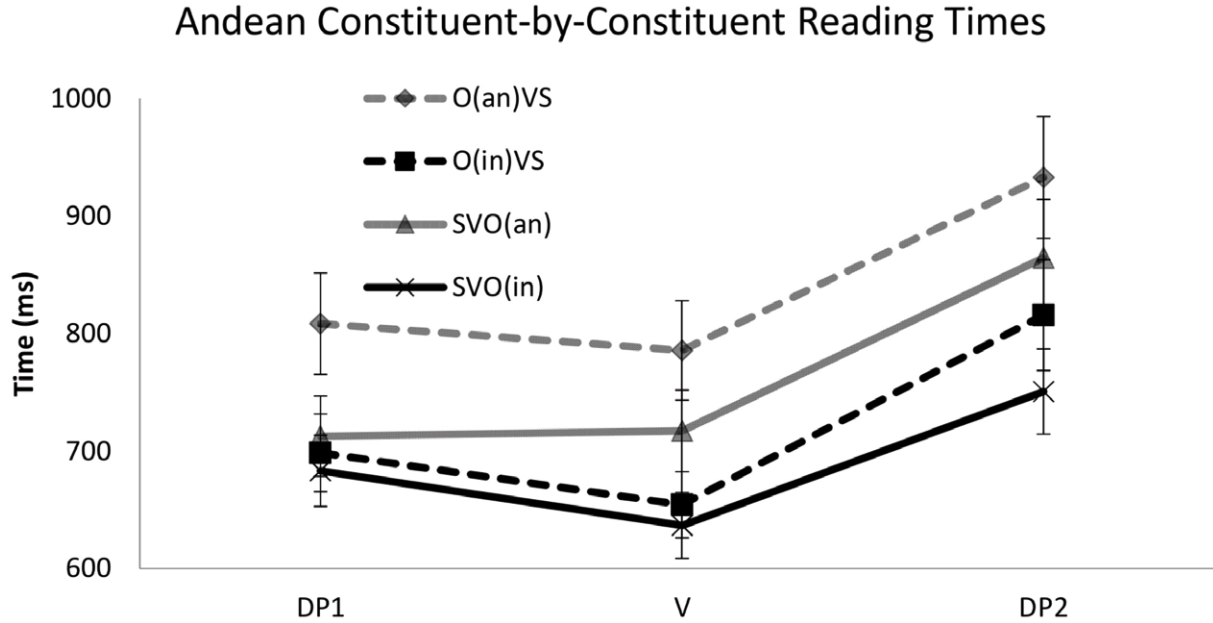


FIGURE 5

Constituent-by-constituent reading times of the Andean Spanish speakers at each constituent position in every experimental condition. DP1 refers to the first Determiner Phrase of the sentences, the subject of SVO conditions and the object of OVS conditions. DP2 refers to the last Determiner Phrase of the sentences, the object of SVO conditions and the subject of OVS conditions

In order to compare Peninsular and Andean Spanish in the constituent-by-constituent analysis, three way ANOVAs were carried out for all three constituent positions (DP1, Verb, DP2) with groups of participants as a between subject factor.

At DP1 position, results revealed main effects of *Constituent order* ($F(1,62) = 44.477$, $P(\text{HF}) < 0.001$) and *Animacy* ($F(1,62) = 54.823$, $P(\text{HF}) < 0.001$). Further, an interaction between *Constituent order* and *Group* ($F(1,62) = 5.713$, $P(\text{HF}) = 0.020$) was observed, signaling that processing preverbal objects was harder for the Andean group than for the Peninsular group. The interaction of *Group* with *Animacy* and the three way interaction were not significant.

For the verb, at second position in the sentence, results revealed main effects of *Constituent order* ($F(1,62) = 21.972$, $P(\text{HF}) < 0.001$) and *Animacy* ($F(1,62) = 98.939$, $P(\text{HF}) < 0.001$). The interactions between *Group* and *Constituent order* ($F(1,62) = 6.034$, $P(\text{HF}) = 0.017$) and *Group* and *Animacy* ($F(1,62) = 4.246$, $P(\text{HF}) = 0.044$) were significant, signaling that the verbs of O-initial and O-animate conditions took longer to process for the Andean group than for the Spanish group. The triple interaction was not significant at verb position.

Finally, for DP2, at sentence final position, the ANOVA did not reveal any significant interaction. Only main effects of *Constituent order* ($F(1,62) = 26.096$, $P(\text{HF}) < 0.001$) and *Animacy* ($F(1,62) = 48.041$, $P(\text{HF}) < 0.001$) were observed.

3. Discussion

3.1. *Whole sentence processing*

Our results reveal that word order and object animacy have a significant impact on sentence processing speed and accuracy in both varieties of Spanish, Peninsular and Andean, as expected. Specifically, both in Peninsular as well as in Andean Spanish: (a) SVO sentences are processed faster and with less errors than their OVS counterparts, and (b) sentences with animate objects take longer to process than those with inanimate ones. We thus replicate the findings previously reported by Casado *et al.* (2005) and Del Rio *et al.* (2012) showing SVO to be the easiest sentence word order to processing Peninsular Spanish. Moreover, our results reveal this to be the case also in Andean Spanish, a variety for which no processing data were available before. This entails that, despite the higher frequency of preverbal objects reported as a characteristic feature of Andean Spanish (Escobar, 1978; Lujan, Minaya & Sankoff, 1981; Cerrón-Palomino, 1990; A.M. Escobar, 1990, 2000; Zavala 1999), this feature is not significant enough to alter the basic processing preferences: SVO is the canonical word order also in this Andean variety, as revealed by shorter reading times and errors, in comparison to OVS.

Both groups of Spanish speakers, Andean and Peninsular, respond similarly when processing sentence word order and object animacy manipulations. The Andean group revealed an interaction between *Constituent order* and *Animacy* that did not reach significance in the Peninsular group (though it almost did, see result section). The direct comparison between groups did not reveal any three way interaction (*Group* by *Constituent order* by *Animacy*) and main effects of *Constituent order* and *Animacy* were observed, signaling that both groups processed experimental conditions in the same fashion. The interactions between *Group* and *Animacy* factors and *Group* and *Constituent order* factors revealed that the processing cost of sentence initial animate Objects, as compared to the rest of the conditions, was greater for the Andean group. This is an interesting finding, because it indicates less tolerance to morphologically marked preverbal objects in the Andean variety than in the Peninsular one, despite reports of higher rates of preverbal objects in Andean Spanish (Muntendam, 2013).

3.2. *Comprehension task*

The comprehension task also sought to measure the comprehension cost of the various experimental conditions. Results reveal that word order and animacy modulate the correct comprehension of the sentences in both groups. Conditions with animate objects elicited more errors and required longer reading times than the conditions with inanimate objects. Regarding word order, object initial sentences elicited more errors and required longer reading times than subject initial counterparts. These results provide additional evidence that canonical word order in Spanish is SVO (Casado *et al.*, 2005; Del Rio *et al.*, 2012). The lack of differences between Andean and Peninsular speakers in the direct group comparison indicates that both groups deploy equivalent representation and processing of sentence word order and animacy traits. In particular, no evidence emerges for a preverbal object processing advantage in Andean Spanish.

3.3. *Reading times constituent by constituent*

The fastest constituents to process in both groups, as expected, were the sentence initial S and sentence initial O in the O-inanimate condition. Given that inanimate objects are not morphologically marked in Spanish, this result shows that a sentence initial inanimate object is initially processed as a subject, deploying the *Subject First strategy* (Bever, 1970). The slowest constituent to process, also for both groups, was the sentence initial animate object, as expected. Recall that animate objects in Spanish carry a *Differential Object Marker* (DOM), the preposition *a*; this marking rules out the possibility of temporarily interpreting the sentence initial object as a subject. Comparing reading times for subjects at sentence initial position ($SVO_{(in)}$ vs. $SVO_{(an)}$), we found longer reading times for condition $SVO_{(an)}$ in both groups of participants. This result is unexpected and we have no explanation for it; we ruled out to differences in frequency or length in the experimental material, for which we tested. Direct comparison between groups revealed that processing preverbal objects was harder for the Andean Spanish speakers than for the Peninsular ones, as suggested by the larger difference in reading times between OVS and SVO sentences for Andeans at DP1 position. As seen in the discussion on reading times for the whole sentence, here again we find that Andean speakers incur in greater costs than Peninsulars when processing OVS word orders.

We compared the reading times for the verbs in the various conditions and found that, in both groups of participants, the verbs in the O-inanimate condition are processed faster than the verbs in the O-animate condition, an expected result. The verb in condition $O_{(an)}VS$ had the longest reading times in both groups of participants, plausibly due to the fact that subjects incur in greater cognitive costs when processing the differentially marked object, which already signals a non canonical sentence word order; we thus interpret it as a *spillover effect* (Just *et al.*, 1982) whereby the processing load of the animate object continues until the next element, in this case the verb. These comparisons must be taken carefully because the verbs in the O-inanimate condition are more frequent and shorter than verbs in the O-animate condition, which could also subtly contribute to faster reading times in the former. When comparing the reading times for verbs in conditions $SVO_{(in)}$ and $SVO_{(an)}$, we found that in $SVO_{(in)}$ condition the verb is processed faster by both groups of participants, which was not expected. We speculate that it results from a small difference in frequency and length of the verb in each condition (see the discussion in the materials subsection above). Direct comparison between groups at verb position revealed that the Andean group was more sensitive to word order variations and object animacy than the Peninsular group.

We measured and compared the reading times for DP2 at the end of the sentence. The S of $O_{(in)}VS$ condition was read significantly slower than the O of the $SVO_{(in)}$ condition in both groups of participants. The comparison of O and S of animate conditions ($SVO_{(an)}$ vs. $O_{(an)}VS$) at sentence final position (DP2) reached significance in the Andean group, but not in the Peninsular group. However, we did not observe any group effect for those O-animate conditions. The greater processing cost of S at sentence final position arises because this is not its canonical position and provides further evidence favoring that Spanish canonical word order is SVO (Casado *et al.*, 2005; del Rio *et al.*, 2012); although the asymmetries in length could overdraw the results.

All in all, constituent-by-constituent analysis revealed that Andean and Peninsular Spanish speakers are similarly affected by word order and animacy variations in the sentence, revealing that canonical word order is SVO in both varieties and animate objects increased the processing cost of sentences, also in both varieties. When differences between both groups of speakers were

observed, they signaled that OVS orders were harder to process for the Andean group. Thus previously reported higher frequencies for OV structures in Andean Spanish (Escobar, 1978; Lujan, Minaya & Sankoff, 1981; Cerrón-Palomino, 1990; A.M. Escobar, 1990, 2000; Zavala 1999) do not appear to sufficiently impact on the sentence processing strategies of Andean Spanish speakers, which remain similar to that of Peninsular speakers.

4. Conclusion

In the present study, we provide pioneering experimental evidence regarding sentence processing in Andean Spanish. Our findings reveal that both in Andean and Peninsular Spanish varieties: (a) sentence canonical word order is SVO, as revealed by the fast reading times and small amount of comprehension errors, and (b) object animacy slows down sentence processing regardless of its position in the sentence, but do it more when sentence initial. Interestingly, our results do not show signs of facilitation for processing animate object initial sentences in the Andean group, despite the higher frequency of preverbal objects reported as a characteristic feature of Andean Spanish (Muntendam, 2013).

Acknowledgements

We are grateful to the National University of San Agustín (Arequipa, Peru) for allowing us to carry out this research in their institution. Specially, we thank Jose Luis Vargas and Clara Calla, from the University Area of Cooperation, Agreements and Internationalization for their kind efforts. This research has been supported by grants from the Spanish Ministerio de Ciencia e Innovación (FFI2015-64183-P and PGC2018-096380-B-I00) and the Basque Government (IT1169-19).

References

- Alameda, J.R. & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Oviedo, University of Oviedo.
- Bahlmann, J.; Münte, T.; Rodríguez-Fornells, A. & Rotte, M. (2007). An fMRI study of canonical and noncanonical word order in German. *Human Brain Mapping* 28, 940-949.
- Basilico, D.; Pinar, P. & Anton-Menendez, I. (1995). *Canonical word order and the processing of verbal traces in Spanish*. Poster presented on the Eighth Annual CUNY Conference on Human Sentence Processing, Tucson, Arizona.
- Bever, T. (1970). The cognitive basis for linguistic structures. In: Hayes, R. (ed.). *Cognition and language development*. New York: Willey & Sons, 279-362.
- Bornkessel, I.; Schlesewsky, M., & Friederici, A. (2002). Grammar overrides frequency: Evidence from online processing of flexible word order. *Cognition*, 85, B21-B30.
- Bornkessel, I., & Schlesewsky, M. (2009). The role of prominence information in the real-time comprehension of transitive constructions: A cross linguistic approach. *Language and Linguistic Compass*, 3, 19-58.
- Camacho, J. 1999. From SOV to SVO: the grammar of interlanguage word order. *Second Language Research*, 15(2), 115-132.

- Casado, P. (2004). *Procesos implicados en la identificación del orden no canónico de oraciones transitivas: una aproximación electrofisiológica*. PhD dissertation. Madrid, Complutense University of Madrid.
- Casado, P.; Fernández Frías, C.; Martin-Loeches, M. & Muñoz, F. (2005). Are semantic and syntactic cues inducing the same processes in the identification of word order? *Cognitive Brain Research*, 24, 526-543.
- Cerrón-Palomino, R. (1990). Aspectos sociolingüísticos y pedagógicos de la motosidad en el Perú. In Cerron-Palomino, R. & Solis Fonseca, G. (eds.) *Temas de lingüística amerindia*. Lima, CONCYTEC / GTZ, 153-180.
- Cerrón-Palomino, R. (2003). *Castellano andino: aspectos sociolingüísticos, pedagógicos y gramaticales*. Lima, PUCP.
- Chomsky, N. (1981). *Lectures on government and binding: The Pisa Lectures*. Holland: Foris Publications.
- Chomsky, N. & Miller, G. (1963). Introduction to the formal analysis of natural languages. In Luce, R., Bush, R. and Galanter, E. (eds.) *Handbook of Mathematical Psychology*, vol. 2. New York: Wiley, 269-323.
- Comrie, B. (1981). *Language universals and linguistic typology: syntax and morphology*. The University of Chicago Press, Chicago.
- Contreras, H. (1978). *A theory of word order with special reference to Spanish*. Amsterdam, North Holland.
- Del Rio, D.; Lopez-Higes, R. & Martin-Aragoneses, M. (2012). Canonical word order and interference-based integration costs during sentence comprehension: The case of Spanish subject – and object relative clauses. *The Quarterly Journal of Experimental Psychology*, 1-21.
- Demestre, J. (2012). Deleting the object marker renders the sentence ungrammatical: comment on Casado, Martin-Loeches, Muñoz, and Fernandez-Frias (2005). *Language and Cognitive Processes*, 27, 682-692.
- Dryer, M.S. 2013. Order of Subject, Object and Verb. In: Dryer, Matthew S. & Haspelmath, Martin (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <<http://wals.info/chapter/81>, accessed on 2017-07-03>.
- Duchon, A.; Perea, M., Sebastián-Gallés, N.; Martí, A. & Carreiras, M. (2013). EsPal: One-stop Shopping for Spanish Word Properties. *Behavior Research Methods*, 45, 1246-1258. <<http://www.bcbl.eu/databases/espal/>>.
- Erdocia, K.; Laka, I.; Mestres-Misse, A. & Rodriguez-Fornells, A. (2009). Syntactic complexity and ambiguity resolution in a free word order language: behavioral and electrophysiological evidences from Basque. *Brain and Language*, 109, 1-17.
- Erdocia, K.; Laka, I. & Rodriguez-Fornells, A. (2012). Processing verb medial word order in a verb final language. In Lamers, M. & de Swart, P. (eds.), *Case, word order and prominence. Studies in theoretical psycholinguistics*. Springer, Nijmegen, 40, 217-237.
- Erdocia, K. & Laka, I. (2018). Negative transfer effects on L2 word order processing. *Frontiers in Psychology*, 9, 337. <<https://doi.org/10.3389/fpsyg.2018.00337>>.
- Escobar, A. (1978). *Variaciones sociolingüísticas del castellano en el Perú*. Lima, Instituto de Estudios Peruanos.
- Escobar, A.M. (1990). *Los bilingües y el castellano en el Perú*. Lima, Instituto de Estudios Peruanos.
- Escobar, A.M. (2000). *Contacto social y lingüístico. El español en contacto con el quechua en el Perú*. Lima, PUCP.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47, 164-203.
- Frisch, S. & Schlesewsky, M. (2001). The N400 indicates problems of thematic hierarchizing. *NeuroReport*, 12, 3391-3394.
- Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68, 1-76.

- Gibson, E., & Fedorenko, E. (2013). The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes*, 28(3): 229-240.
- Godenzzi (1991). Discordancias gramaticales del castellano andino en Puno (Peru). *Lexis*, 25 (1), 107-118.
- Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, J. (ed.). *Universals of language*. Cambridge: MIT Press, 58-90.
- Grewe, T., Bornkessel-Schlesewsky, I., Zysset, S., Wiese, R., von Cramon, D. Y., & Schlewsky, M. (2007). The role of the posterior superior temporal sulcus in the processing of unmarked transitivity. *NeuroImage*, 35(1), 343-352.
- Hale, J. (2001). A probabilistic early parser as a psycholinguistic model. In *Proceedings of NAACL*, volume 2, 159-166.
- Just, M.A.; Carpenter, P.A., & Woolley, J.D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology*, vol. 111 (2), 228-238.
- Kaan, E. (1997). *Processing subject-object ambiguities in Dutch*. PhD Dissertation. University of Groningen.
- Kaiser, E. & Trueswell, J. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94, 113-147.
- Klee, C. 1996. The Spanish of the Peruvian Andes: The Influence of Quechua on Spanish Language Structure. In Roca, A. and Jensen, J. (eds.) *Spanish in Contact, Issues in Bilingualism*, 73-92.
- Laka, I. (1990). *Negation in syntax: On the nature of functional categories and projections*. PhD Dissertation, Massachusetts Institute of Technology.
- Lambrecht, K (1994). *Information structure and sentence form: Topic, focus and the mental representation of discourse referents*. Cambridge Studies in Linguistics, 71. Cambridge, Cambridge University Press.
- Lee, H. & Kim, N. (2012). Non-canonical word order and subject-object asymmetry in Korean case ellipsis. In Müller, S. (ed.). *Proceedings of the 19th International Conference on Head Driven Phrase Structure Grammar, Chungnam National University Daejeon*, 427-442. Stanford, CSLI Publications.
- Lujan, M., Minaya, L. & Sankoff, D (1981). El principio de consistencia universal en el habla de niños bilingües peruanos. *Lexis*, 5 (2), 95-110.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106, 1126-1177.
- Levy, R., Fedorenko, E. & Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *Journal of Memory and Language*, 69, 461-495.
- Mak, W. (2001). *Processing of relative clauses: Effects of pragmatic, semantic and syntactic variables*. Doctoral dissertation. University of Nijmegen, Nijmegen.
- Mak, W.; Schriefers, H.; & Vonk, W. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language*, 47, 50-68.
- Mak, W.; Schriefers, H. & Vonk, W. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, 54, 466-490.
- Marantz, A. (2005). Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review* 22, 429-445.
- Matzke, M.; Heinke, M.; Nager, W.; Rösler, J. & Münte, T. (2002). The costs of freedom: an ERP study of non canonical sentences. *Clinical Neurophysiology*, 113, 844-852.
- Muntendam, Antje (2009). *Linguistic transfer in Andean Spanish: Syntax or Pragmatics?* PhD Dissertation, University of Illinois at Urbana-Champaign.
- Muntendam, Antje (2013). On the nature of cross-linguistic transfer: A case study of Andean Spanish. *Bilingualism, Language and Cognition*, 16, 111-131.

- Muysken, P. 1984. The Spanish that Quechua speakers learn: L2 learning as norm governed behavior. In Andersen, R. (ed.), *Second languages: a crosslinguistic perspective*. Rowley, Mass., Newbury House, 101-119.
- Nieuwland, M. S., Martin, A. E., Carreiras, M. (2013). Event-related brain potentials evidence for animacy processing asymmetries during sentence comprehension. *Brain and Language* 126, 151-158.
- Phillip, M.; Bornkessel-Schlesewsky, I.; Bisang, W. & Schlewsky, M. (2008). The role of animacy in real time comprehension of Mandarin Chinese: Evidence from auditory event-related brain potentials. *Brain and Language*, 105, 112-133.
- Phillips, C. (2013). Parser-grammar relations: We don't hear everything twice. In Sanz, M., Laka, I. & Tanenhaus, M. (eds). *Language down the garden path: the cognitive and biological basis for linguistic structure*. Oxford, Oxford University Press, 294-315.
- Prat-Sala, M. & Branigan, H. (2000). Discourse constraints on syntactic processing in Language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language*, 42, 168-182.
- Rohde, D.L.T. (2001). *Linger* [Computer Software]. Cambridge, MA: MIT TedLab.
- Sanchez, L. (2003). *Quechua-Spanish bilingualism: Interference and convergence in functional categories*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Schlewsky, M.; Fanselow, G.; Kliegl, R. & Krems, J. (2000). The subject preference in the processing of locally ambiguous wh-questions in German. In Hemforth, B. & Konieczny, L. (eds.). *German Sentence Processing*, 65-93.
- Sekerina, I. (2003). Scrambling and processing: Dependencies, complexity and constraints. In Karimi, S. (ed.). *Word order and scrambling*. Malden, WA, Blackwell Publishers, 1-18.
- Sprouse, J. & Almeida, D. (2012). Assessing the reliability of textbook data in syntax: Adger's Core Syntax. *Journal of Linguistics*, 48, 609-652.
- Valdez, L. (2002). *Clitics in the speech of monolingual Andean Spanish speakers*. PhD Dissertation, University of Pittsburgh.
- Vasishth, S. (2002). *Working memory in sentence comprehension: Processing Hindi center embeddings*. Doctoral Dissertation, Ohio State University.
- Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology*, 36, 559-570.
- Weyerts, H.; Penke, M.; Münte, T.; Heinze, H-J. & Clahsen, H. (2002). Word order in sentence processing: An experimental study of verb placement in German. *Journal of Psycholinguistic Research*, vol. 31, 211-268.
- Wolff, S.; Schlewsky, M.; Hirotani, M. & Bornkessel-Schlewsky, I. (2008). The neural mechanisms of word order processing revisited: electrophysiological evidence from Japanese. *Brain and Language*, 107, 133-157.
- Zavala, V. (1999). Reconsideraciones en torno al español andino. *Lexis*, 24 (1), 25-85.
- Zubizarreta, M.L. (2000). Las funciones informativas: tema y foco. In Bosque, I. & Demonte, V. (eds.). *Gramática descriptiva de la lengua española*, Madrid, Espasa Calpe, 4215-4244.

UNIBERTSITATEKO ESKULIBURUAK
MANUALES UNIVERSITARIOS

INFORMAZIOA ETA ESKARIAK • INFORMACIÓN Y PEDIDOS

UPV/EHUko Argitalpen Zerbitzua • Servicio Editorial de la UPV/EHU
argitaletxea@ehu.eus • editorial@ehu.eus
1397 Posta Kutxatila - 48080 Bilbo • Apartado 1397 - 48080 Bilbao
Tfn.: 94 601 2227 • www.ehu.eus/argitalpenak

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea