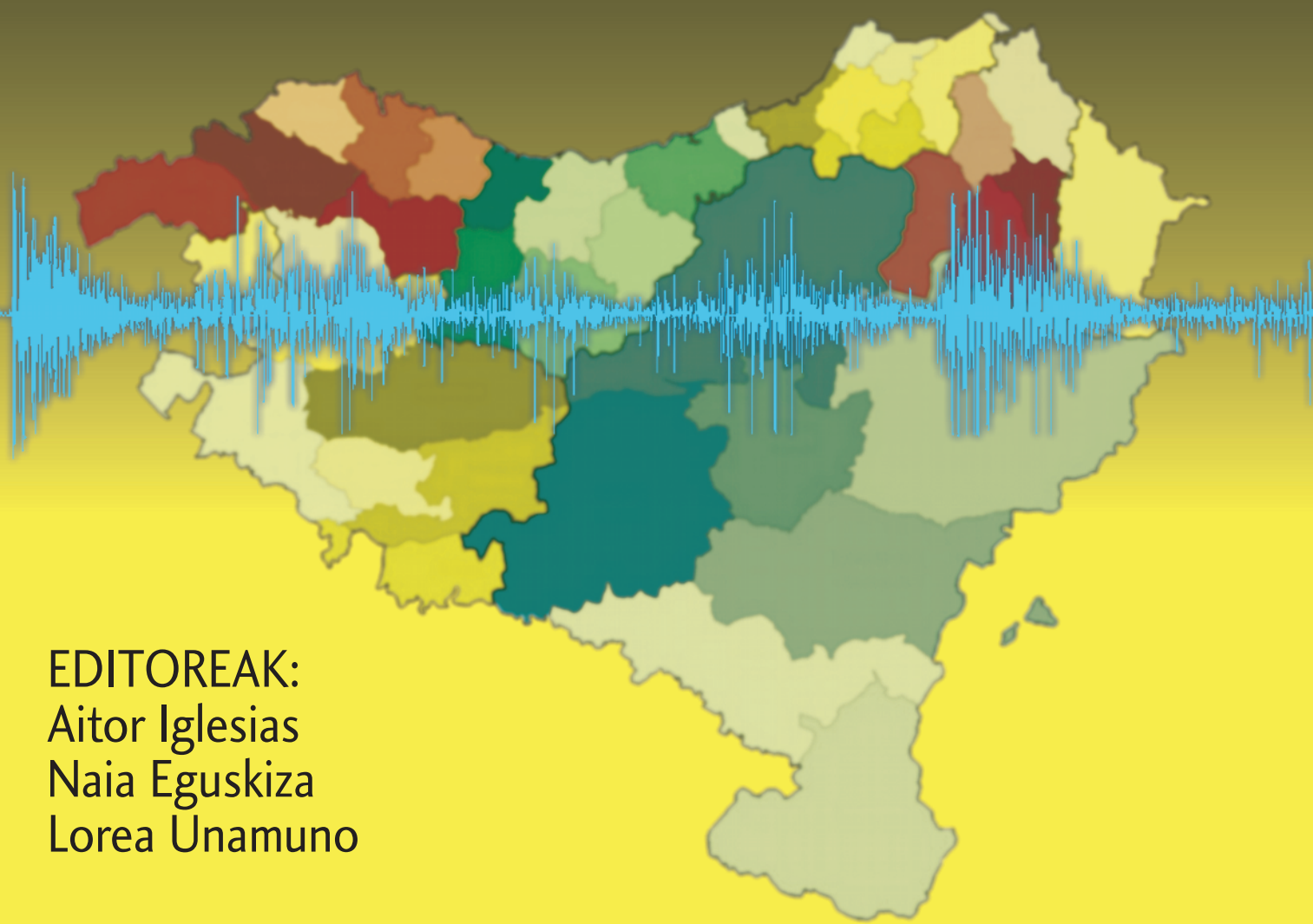


# Linguistic Variation in Basque language and Education-IV

## Euskararen bariazioa eta bariazioaren irakaskuntza-IV



EDITOREAK:  
Aitor Iglesias  
Naia Eguskiza  
Lorea Unamuno



Universidad del País Vasco Euskal Herriko Unibertsitatea

**Linguistic** variation in Basque language and education - IV [Recurso electrónico] = Euskararen bariazioa eta bariazioaren irakaskuntza -IV /editoreak, Lorea Unamuno ...[et al.]. – Datos. - Bilbao : Universidad del País Vasco / Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua = Servicio Editorial, D.L. 2019.

1 disco compacto CD-Rom (PDF, 103 p.)

Requisitos del sistema: Adobe Acrobat Reader.

Textos en euskara e inglés.

D.L.: BI-374-2019. - ISBN: 978-84-1319-004-4.

1. Euskara (Lengua) - Dialectos. 2. Lenguaje y lenguas - Variación. 3. Prosodia (Lingüística). I. Unamuno, Lorea, coed. II. Título: Euskararen bariazioa eta bariazioaren irakaskuntza -IV.

(0.034)809.169-087



UPV-EHUko EUDIA Ikerketa-taldea

Euskal Herriko Unibertsitateko ikerketa taldea (GIU 22/16)

Euskal Unibertsitate Sistemako ikerketa taldea (IT 1028/16)

## AURKIBIDEA

|  |    |
|--|----|
| Informatzaileen sailkapena bat-bateko testuen lexikoaren erabileraren arabera.....                                   | 4  |
| <i>Iñaki Gaminde</i>   |    |
| Gestures synchronized with speech in infant language.....  | 20 |
| <i>Alfonso Igualada</i>  |    |
| Aditz jokatuan non ditu zailtasunak euskara H2 ikastun berantiarrak? .....   | 27 |
| <i>Juan Abasolo</i>  |    |
| Euskalkien sailkapen zientifikoa: lexikoa (2) .....  | 41 |
| <i>Gotzon Aurrekoetxea, Jose Luis Ormaetxea, Xarles Videgain</i>   |    |
| Bideo-podcasta: erronketan oinarritutako ikaskuntza .....  | 60 |
| <i>Aitor Iglesias, Aintzane Etxebarria, Aintzane Pagadigorria, Lorea Unamuno</i>                                     |    |
| Gazte euskaldunen bai/ez galderen intonazioaz: euskaraz eta gaztelaniaz .....  | 71 |
| <i>Naia Eguskiza, Irati de Pablo, Iñaki Gaminde</i>  |    |
| Euskararen erabilera neurtzeko kale-neurketa Bilboko Hezkuntza Fakultatean.....                                      | 89 |
| <i>Aiora Aginaga, Ibai Bilbao, Irati Etxebarria, Amaia Izarzugaza, Eider Larrabeiti, Ibai Urrutia, Lorea Zulaika</i> |    |
| Helduen prosodiaren eragina irakurritako testuetan haurren arretan .....   | 98 |
| <i>Irati de Pablo, Asier Romero, Iñaki Gaminde, Naia Eguskiza</i>  |    |

# INFORMATZAILEEN SAILKAPENA BAT-BATEKO TESTUEN LEXIKOAREN ERABILERAREN ARABERA

**Iñaki Gaminde**  
**inaki.gaminde@gmail.com**

## **Laburpena**

Gure lan honen helburua bikoitza da; batetik, gai berberaren gainean informatzaileek ekoitzi dituzten bat-bateko testu laburren lexikoaren erabileran oinarrituta, ikusi nahi dugu informatzaileak geografikoki koherenteki egituratzen diren zonalde konpaktuetan ala ez eta, bestetik, era honetako azterketak egiteko teknikak eta urratsak deskribatu ere egin nahi dira beste ikertzaile batzuk animatzekotan. Corpora lortzeko metodologia azaldu ondoren, datuen prozesaketa xeheki deskribatzen da. Informatzaileak sailkatzeko hiru teknika erabiltzen dira: multzokatze hierarkikoa, K-batezbestekoak, eta logika lausoan oinarritutako FCM algoritmoa. Teknika hauen bidez lortutako multzoen koherentzia aztertze hiru neurri erabili dira:  $\alpha$  dibertsitatea, Shannonen entropia eta Yule-ren K.

**Gako hitzak:** lexikoa, multzokatze azterketa, informatzaileen sailkapenak

## **1. Sarrera**

Lexikoaren aldakortasuna aztertzean, metodologia bi erabili izan dira nagusiki orain arte gure artean. Batetik, kontzeptu edo hitz zerrenda itxiak erabili izan dira informatzaileei galdetzeko eta horien bidez datuak jasotzeko edo, bestetik, herrietako hizkeren deskripzioetan batez ere jaso direnak, gai libreak eta elkarrizketa gidatuak erabili izan dira aurreko metodologiarekin konbinatuta. Lehenengoetan, Euskaltzaindia argitaratzen dabilen Euskararen Herri Hizkeren Atlas (Euskaltzaindia, 2008) handiena eta garrantzitsua da; badira era honetako beste lan batzuk, gainera kuantitatiboki landu direnak hizkerak sailkatzekotan (Aurrekoetxea, 2008; Ormaetxea, 2008; Gaminde et. al. 2016b, 2016c, 2017). Bigarrengean artean, herrietako monografiak (Gaminde, 1997, 1999, 2000, 2001) eta gai zehatzen gainean egindako lanak (Barrutia, 1996, 2000; Basauri eta Sarasua, 2003) kokatzen dira, beste askoren artean.

Batzuetan zein besteetan ez da kontuan hartzen lexikoaren erabileraren maiztasuna ez lexiko aktiboaren eta pasiboaren artean egiten den bereizkuntza (Sisková, 2012). Era honetako lanak oso ugariak dira lexiko aberastasunaren ikuspuntutik (Hout eta Vermeer, 2007; López Morales, 2011; Kim, 2014) eta baita erazko lexikoarenetik ere (Euskal Herrian badago bat eginda Etxebarria, 1996).

Gure lan honetarako ideia sortzen da lexiko aberastasuna kuantifikatzeko garatu diren indizeen azterketatik (Vermeer, 2000; Gregori-Signes eta Clavel-Arroitia, 2015; Hout eta Vermeer, 2007) eta baita biodibertsitatea kuantifikatzeko sortu direnetatik ere. Honelakoetan, konplexutasun lexikoa aztertze indizeek hiru arlo nagusi aztertzen dituzte (Sisková, 2012; Kim, 2014): dentsitatea, sofistikazioa eta dibertsitatea. Indizeok esparru

askotan erabiltzen dira, hala nola ikas-prozesuen garapena aztertzeko idatzian zein ahozkoan (Vermeer, 2000; Xiaofei, 2012; Sisková, 2012; Gregori-Signes eta Clavel-Arroitia, 2015; Bonvin eta Lambelet, 2017), aldagai soziolinguistikoen araberrako azterketak egiteko (Van Gijssel et. al., 2005; Roberto et. al., 2012), hizkuntzaren jabekuntzaren progresioa aztertzeko (Kim, 2014; Bonvin eta Lambelet, 2017), eta abar.

Honenbestez, esan dezakegu gure lan honen helburua bikoitza dela; batetik, gai berberaren gainean informatzaileek ekoitzi dituzten bat-bateko testu laburren lexikoaren erabileran oinarrituta, ikusi nahi dugu informatzaileak geografikoki koherenteki egituratzen diren zonalde konpaktuetan ala ez, eta, bestetik, era honetako azterketak egiteko teknikak eta urratsak deskribatu ere egin nahi dira beste ikertzaile batzuk animatzekotan.

Datu linguistikoak tratatzeko metodo kuantitatiboak erabiltzea ez da gauza berria aldakortasunaren azterbidean, dela soziolinguistikaren ikuspuntutik, dela dialektometriaren ikuspuntutik. Euskal Herrian ere metodo hauek, nekez izan bada ere, aurrera egin dute aspaldiko urteotan (Aurrekoetxea, 1992, 1995, 2003, 2004, 2005, 2008, 2009, 2016; Aurrekoetxea eta Ormaetxea, 2006, 2010; Aurrekoetxea et al., 2014a, 2014b; Gaminde, 2007, Iglesias, 2014, Ensunza, 2015, Lujanbio, 2016) metodo hauek hizkerak sailkatzeko ez ezik, informatzaileak sailkatzeko ere erabili dira (Gaminde et. al. 2012, 2016a, 2017a).

Informatzaileen sailkapenaren helburua datuen azpian datzan egituraketa geolinguistikoa edo soziolinguistikoa azalartzea da, horrela bada gure lan hau sistema konplexu dinamiko moldagarrien ikuspuntuan guztiz errotzen da (Beckner et al., 2011; Ellis, 2011; López Rivera, 2013; Moreno Cabrera, 2008; Léonard et al., 2015) eta bat egiten du aspaldion garatzen ari garen esparruan ere (Gaminde, et. al.; 2017b; Eguskiza et. al., 2017)

Lana lau ataletan banatuta aurkezten dugu: sarrera honen osteko bigarren atalean, corpusaren ezaugarriak eta berori jasotzeko eta prozesatzeko metodologia aurkezten dira; hirugarren atalean datuen azterketa egiten da; azkenik, laugarren atalean, ondorio nagusiak laburbiltzen dira.

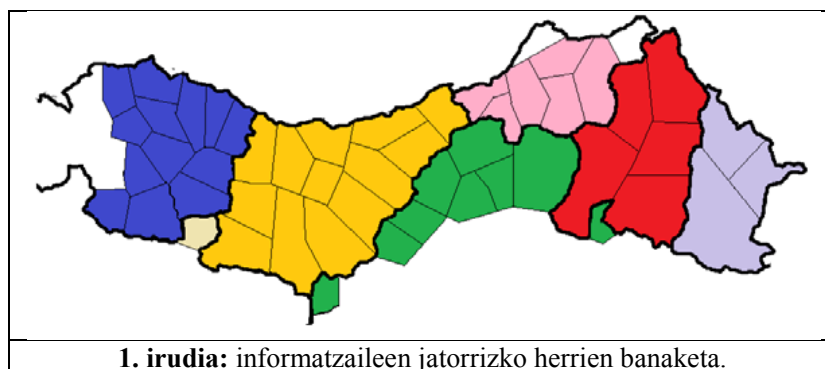
## 2. Corpora eta metodologia

Gure lan hau egin ahal izateko, XX. mende goieneko euskaldunen testuen corpora (XXTC) erabili dugu. Corpus horretan era askotako informatzaileen testuak eta materialak biltzen diren arren, gure hemengo helburuetarako bat-bateko 50 testu aukeratu ditugu.

Informatzaileen ezaugarriei dagokienez, denak andrazkoak izan dira, euren ama-hizkuntza euskara izan da, ikasketak D ereduan egin dituzte unibertsitate mailara arte, behintzat, eta XX. mende goienean jaiokoak izan dira 1980-1999 urte bitartean. Informatzaileak identifikatzeko, euren jaioterria erabiliko dugu; 1. taulan, eduki ditugun informatzaileak zehazten dira probintziaka (ikusi 1. irudiko mapa).

| Probintzia       | Informatzaileak  |
|------------------|--|
| Araba            | Aramaio  |
| Bizkaia          | Bermeo, Elorrio, Ereño, Gernika, Larrabetzu, Lekeitio, Lemoa, Mallabia, Mungia, Ondarroa, Orozko, Zeanuri eta Zornotza |
| Gipuzkoa         | Andoain, Azkoitia, Bergara, Donostia, Eibar, Errezil, Hernani, Hondarribia, Oñati, Ordizia, Segura, Tolosa eta Zarautz |
| Lapurdi          | Azkaine, Ezpeleta, Hazparne, Larresoro, Mugerre eta Ziburu   |
| Nafarroa Garaia  | Goizueta, Irurita, Leitza, Lekunberri, Lesaka, Luzaide, Sunbilla, Urdiain eta Zubieta                                  |
| Nafarroa Beherea | Aiherra, Baigorri, Donapauale, Garazi eta Izura  |
| Zuberoa          | Barkoxe, Maule eta Urdiñarbe   |

**1. taula:** lana egin ahal izateko eduki ditugun informatzaileak probintziak.



Testuak jasotzeko, 2. irudian erakusten den komikia eman zitzaien informatzaileei eta, istorioa ulertzeko eta pentsatzeko denbora laburra emandakoan, kontatzeko eskatu zitzaien. Audio seinaleak jasotzeko erabili diren tresnak Sony Minidisc, Marantz PMD620 eta ZOOMH4n izan dira. Grabazioen kalitatea bermatzeko, kanpo-mikrofonoak erabili dira kasu gehienetan, eta zuzenean “wav” formatuan jaso dira. Behin grabazioak eginda, “praat” izeneko programa erabili da (Boersma & Weenink 2016) transkribatu eta etiketatu ahal izateko.



Corpusa azterketarako prestatzeko lematizazio-prozesua egin da transkripzioak normalizatzeko. Euskalkietako ahozko laginekin jardutean, hau prozesu garrantzitsua izaten

da; izan ere, hartzen diren erabakiek lan osoaren emaitzak baldintzatu ditzakete. Jarraian, lematizatzeko hartu diren erabakirik garrantzitsuenak zehazten dira.

Lehenbizikoa izan da aldaki fonetikoak kontuan ez hartzea eta era berean morfemei jaramonik ez egitea. Honek esan nahi du “ekarri”, “karri”, “ekarten”, “ekartzen”, “ekarko” moduko guztiek “ekarri” lema bera edukiko dutela. Era berean, “aita”, “aite”, “aitta”, “atxe”, “aitxe”, “aite”, “aitze” eta abarri dagokien lema “aita” izango da eta berdin jokatu da “aitari”, “aitxeas”, “aittan” eta abarrekin.

Aditz trinko jokatuaren kasuan, partizipioaren lema erabili dugu, orduan “dijo”, “doya”, “zioan”, “doaz” eta abar “joan” lemarekin adierazi dira. Adizkia trinkoa zein laguntzailea denean, lema bi erabili dira: “dau”, “do”, “du” eta abar laguntzaileak direnean, “du” hartu da lematzat, baina, aditz trinkoa izanez gero, orduan “edun” hartu da lematzat. Gauza bera egin da “izan” aditzaren adizkiekin: laguntzaileak izanez gero, euren estandarrari dagokion adizkia erabili da lematzat, eta, trinkoa izanez gero, “izan” bera. Bestalde, adizkien kasuan ez dira kontuan hartu euren inguruan ager daitezkeen “ba”, “bait”, “la” edo “n” moduko morfemak.

Prozesu hau burutu ondoren, lemetako batzuk lema gramatikalak edo funtzio berbak izango dira, eta beste batzuk, eduki semantikodun lemak, hau da, izenak, aditzak, adjektiboak eta postposizioak. Gure lan honetan eduki semantikodun lemak baino ez ditugu erabili, denetara 2164.

Lema hauen guztien maiztasunak kalkulatu ditugu orokorrean eta informatzaile bakoitzeko. Maiztasun horien arabera informatzaileak sailkatzeko erabili dugun datu matrizea osatu dugu. Matrizea egiteko, 0,46 baino maiztasun handiagoko lemak aukeratu ditugu, edo 10 gauzapean baino gehiago izan dituztenak. Hortik beherakoak baztertu ditugu; izan ere, lexikoaren azterketan ikusiko dugun bezala, informatzaileek erabili dituzten 259 lemetan 44k baino ez daukate 0,46etik gorako maiztasuna. Lehen urratsean, informatzaile bakoitzari lema bakoitzeko erabilera kopurua esleitu zaio dagokion lekuan, 3. irudian erakusten den moduan.

| Lekukoa     | agertu | ahate | aintza | aita | argazki | argazki ka | argazki tr | argazkilan | arrian | atera | automatik | azken |
|-------------|--------|-------|--------|------|---------|------------|------------|------------|--------|-------|-----------|-------|
| Aiherra     | 1      | 1     | 0      | 3    | 3       | 0          | 0          | 0          | 0      | 0     | 0         | 0     |
| Andoain     | 0      | 2     | 0      | 0    | 5       | 0          | 0          | 0          | 2      | 4     | 0         | 0     |
| Aramato     | 2      | 1     | 0      | 3    | 3       | 0          | 0          | 0          | 0      | 3     | 0         | 0     |
| Azkaine     | 0      | 1     | 2      | 3    | 1       | 0          | 1          | 0          | 0      | 0     | 0         | 0     |
| Azkotia     | 0      | 0     | 0      | 3    | 2       | 0          | 0          | 0          | 0      | 3     | 0         | 1     |
| Baigorri    | 1      | 2     | 0      | 0    | 4       | 0          | 1          | 0          | 0      | 0     | 0         | 0     |
| Barkoxe     | 0      | 0     | 0      | 0    | 4       | 0          | 1          | 0          | 0      | 0     | 0         | 0     |
| Bergara     | 0      | 1     | 0      | 2    | 4       | 0          | 0          | 0          | 0      | 3     | 0         | 0     |
| Bermeo      | 0      | 0     | 0      | 0    | 3       | 0          | 0          | 0          | 0      | 1     | 1         | 0     |
| Donapaula   | 0      | 1     | 0      | 4    | 2       | 0          | 1          | 0          | 1      | 0     | 0         | 0     |
| Donostia    | 0      | 1     | 0      | 2    | 3       | 0          | 0          | 0          | 0      | 3     | 0         | 0     |
| Eibar       | 0      | 1     | 0      | 5    | 5       | 0          | 0          | 0          | 0      | 4     | 0         | 0     |
| Eibarri     | 0      | 0     | 0      | 3    | 3       | 0          | 0          | 0          | 0      | 2     | 0         | 0     |
| Eneño       | 1      | 0     | 0      | 0    | 0       | 0          | 0          | 0          | 0      | 2     | 1         | 0     |
| Errezil     | 0      | 3     | 0      | 4    | 3       | 1          | 0          | 0          | 0      | 3     | 0         | 1     |
| Erpeleta    | 0      | 0     | 0      | 2    | 0       | 0          | 0          | 0          | 0      | 0     | 0         | 0     |
| Garazi      | 0      | 3     | 3      | 4    | 4       | 0          | 1          | 0          | 0      | 0     | 0         | 1     |
| Gernika     | 1      | 0     | 0      | 0    | 5       | 0          | 0          | 0          | 0      | 3     | 1         | 0     |
| Gozueta     | 0      | 0     | 0      | 0    | 3       | 1          | 0          | 0          | 0      | 3     | 0         | 0     |
| Hazparne    | 3      | 0     | 0      | 0    | 4       | 0          | 0          | 0          | 0      | 0     | 0         | 1     |
| Hernani     | 0      | 2     | 0      | 3    | 3       | 0          | 0          | 0          | 1      | 2     | 2         | 0     |
| Hondarribia | 0      | 2     | 0      | 5    | 8       | 0          | 0          | 0          | 2      | 3     | 2         | 1     |
| Irurita     | 3      | 0     | 0      | 3    | 0       | 0          | 0          | 0          | 0      | 2     | 0         | 0     |
| Izura       | 0      | 2     | 0      | 5    | 6       | 0          | 3          | 0          | 1      | 0     | 0         | 1     |

**3. irudia:** informatzaileen lema bakoitzaren kopuruen arabera egineko matrizea.

Matrize hauen datuetan oinarrituta, lema bakoitzaren maiztasuna kalkulatu da informatzaile bakoitzaren lema kopuru osoaren arabera, eta sailkapenerako erabiliko den oinarritzko matrizea osatu da (4. irudia).

| lekukoa     | agertu | ahate | aintzira | aita | argazki | argazki kama | argazki tresn | argazkilarri | ari izan | atera | automatiko | azken |
|-------------|--------|-------|----------|------|---------|--------------|---------------|--------------|----------|-------|------------|-------|
| Aiherra     | 0,02   | 0,02  | 0,00     | 0,07 | 0,07    | 0,00         | 0,00          | 0,00         | 0,00     | 0,00  | 0,00       | 0,00  |
| Andoain     | 0,00   | 0,05  | 0,00     | 0,00 | 0,14    | 0,00         | 0,00          | 0,00         | 0,05     | 0,11  | 0,00       | 0,00  |
| Aramaió     | 0,05   | 0,02  | 0,00     | 0,07 | 0,07    | 0,00         | 0,00          | 0,00         | 0,00     | 0,07  | 0,00       | 0,00  |
| Azkaine     | 0,00   | 0,03  | 0,05     | 0,08 | 0,03    | 0,00         | 0,03          | 0,00         | 0,00     | 0,00  | 0,00       | 0,00  |
| Azkortia    | 0,00   | 0,00  | 0,00     | 0,08 | 0,05    | 0,00         | 0,00          | 0,00         | 0,00     | 0,08  | 0,00       | 0,03  |
| Baigorri    | 0,03   | 0,06  | 0,00     | 0,00 | 0,12    | 0,00         | 0,03          | 0,00         | 0,00     | 0,00  | 0,00       | 0,00  |
| Barkoxe     | 0,00   | 0,00  | 0,00     | 0,00 | 0,09    | 0,00         | 0,02          | 0,00         | 0,00     | 0,00  | 0,00       | 0,00  |
| Bergara     | 0,00   | 0,03  | 0,00     | 0,05 | 0,11    | 0,00         | 0,00          | 0,00         | 0,00     | 0,08  | 0,00       | 0,00  |
| Bermeo      | 0,00   | 0,00  | 0,00     | 0,00 | 0,10    | 0,00         | 0,00          | 0,00         | 0,00     | 0,03  | 0,03       | 0,00  |
| Donapauze   | 0,00   | 0,02  | 0,00     | 0,08 | 0,04    | 0,00         | 0,02          | 0,00         | 0,02     | 0,00  | 0,00       | 0,00  |
| Donostia    | 0,00   | 0,03  | 0,00     | 0,05 | 0,08    | 0,00         | 0,00          | 0,00         | 0,00     | 0,08  | 0,00       | 0,00  |
| Eibar       | 0,00   | 0,03  | 0,00     | 0,13 | 0,13    | 0,00         | 0,00          | 0,00         | 0,00     | 0,10  | 0,00       | 0,00  |
| Elorrio     | 0,00   | 0,00  | 0,00     | 0,10 | 0,10    | 0,00         | 0,00          | 0,00         | 0,00     | 0,07  | 0,00       | 0,00  |
| Ereñio      | 0,02   | 0,00  | 0,00     | 0,00 | 0,00    | 0,00         | 0,00          | 0,00         | 0,00     | 0,04  | 0,02       | 0,00  |
| Errezil     | 0,00   | 0,06  | 0,00     | 0,07 | 0,06    | 0,02         | 0,00          | 0,00         | 0,00     | 0,06  | 0,00       | 0,02  |
| Etpeleta    | 0,00   | 0,00  | 0,00     | 0,10 | 0,00    | 0,00         | 0,00          | 0,00         | 0,00     | 0,00  | 0,00       | 0,00  |
| Garazi      | 0,00   | 0,04  | 0,04     | 0,06 | 0,06    | 0,00         | 0,01          | 0,00         | 0,00     | 0,00  | 0,00       | 0,01  |
| Gernika     | 0,03   | 0,00  | 0,00     | 0,00 | 0,15    | 0,00         | 0,00          | 0,00         | 0,00     | 0,09  | 0,03       | 0,00  |
| Goizueta    | 0,00   | 0,00  | 0,00     | 0,00 | 0,08    | 0,03         | 0,00          | 0,00         | 0,00     | 0,08  | 0,00       | 0,00  |
| Hazparne    | 0,07   | 0,00  | 0,00     | 0,00 | 0,10    | 0,00         | 0,00          | 0,00         | 0,00     | 0,00  | 0,00       | 0,02  |
| Hernani     | 0,00   | 0,05  | 0,00     | 0,07 | 0,07    | 0,00         | 0,00          | 0,00         | 0,02     | 0,05  | 0,05       | 0,00  |
| Hondarribia | 0,00   | 0,03  | 0,00     | 0,08 | 0,13    | 0,00         | 0,00          | 0,00         | 0,03     | 0,05  | 0,03       | 0,02  |
| Irurita     | 0,04   | 0,00  | 0,00     | 0,04 | 0,00    | 0,00         | 0,00          | 0,00         | 0,00     | 0,02  | 0,00       | 0,00  |
| Izura       | 0,00   | 0,03  | 0,00     | 0,07 | 0,09    | 0,00         | 0,04          | 0,00         | 0,01     | 0,00  | 0,00       | 0,01  |

**4. irudia:** informatzaileen lema bakoitzaren maiztasunaren arabera egineko matrizea.

Hiru teknika sailkatzaile erabili ditugu aurkezten dugun lantxo honetan. Batetik, multzokatze-azterketa hierarkiko klasikoa egin dugu, azpiko egiturarik agertzen den jakiteko, eta erabakitzeko gainerako teknika bietan zein izan daitekeen datuekin ondoen datorren multzo kopurua. Teknika honetan, eta gero beste bietan ere, aukeratu den distantzia mota Manhattan izenekoa izan da (Pardo eta Ruiz, 2002, Pérez López, 2005) eta multzokatze metodoa Ward izenekoa (Goebel, 1992, 2010; Aurrekoetxea, 1995, 2005; Gaminde, 2007; Gaminde et al., 2016b).

Sailkatzeko erabili ditugun beste teknika biak K-batezbestekoa (KM) eta logika lausoan oinarritzen den FCM algoritmoa (fuzzy C-means, FCM) izan dira. Kasu bietan, lau multzo egin dira informatzaileekin, 1000 iterazio KM eta FCM tekniketan, eta FCM teknikan lausotasun-maila 1,2koa izan da.

FCM multzokatzea egiten dugunean, elementu guztiak geuk alde aurretik aukeratutako multzo kopuruaren arabera sailkatuta agertuko zaizkigu, baina elementu bakoitzeko adieraziko zaigu zein neurritan dagoen multzo baten barruan, haren barnekotasun-funtzioa emanik. Barnekotasun-funtzioa era honetara adierazten da:

$$\mu_A: X \rightarrow [0,1]$$

Elementu baten funtzioa “1” bada, guztiz dago multzo horren barruan, eta “0” bada, ez da multzo horretakoa, baina “0,8”koa izan daiteke multzo batean, eta “0,2”koa, beste batean. Honela, bada, hizkeren arteko trantsizio-guneak irudika daitezke. Multzo bereko elementuak haiekotzat hartu ditugu euren barnekotasun-funtzioa 0,85en eta 1en artean dagoenean, 0,85etik beherako kasuak trantsizio-guneetan kokatuko lirarteke.

Multzoen barneko koherentzia erakusteko hiru neurri multzo erabiliko dira:  $\alpha$  dibertsitatea (De Cáceres et al., 2003), Shannon-en entropia eta Yule-ren K (Roberto et al., 2012; Wojciechowski eta Gorzyski, 2016).



### 3. Datuen azterketa

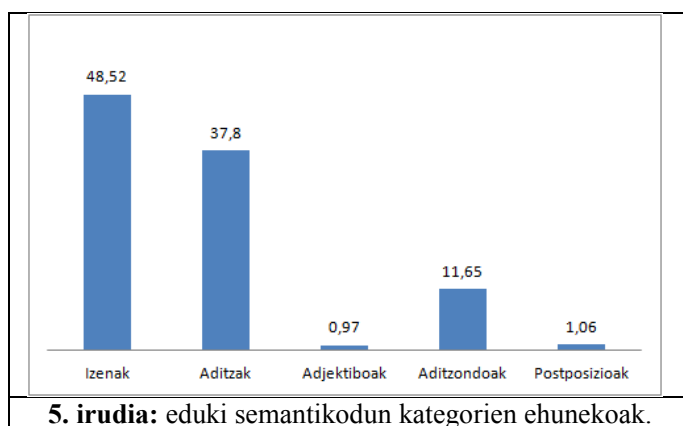
Atal hau azpialat bitan banatuta aurkezten dugu; lehen azpiatalean, testuak egiteko informatzaileek erabili duten lexikoa aztertuko dugu, eta bigarrenean, informatzaileak sailkatuko dugu testuetan erabilitako lexiko horren arabera.

#### 3.1. Lexikoaren azterketa

Gorago esaneko moduan, lan honetan, eduki semantikodun lexikoaz baino ez gara arduratuko; hauek izenak, aditzak, adjektiboak aditzondoak eta postposizioak dira, eta 2164 izan dira; 2. taulan, kategoria bakoitzeko lema kopuruak eta ehunekoak ematen ditugu. Taulan argi ikus daitekeen moduan, izenak eta aditzak dira gehien erabiltzen diren kategoriak (5. irudiko grafikoa).

| Lema mota            | kopurua | %     |
|----------------------|---------|-------|
| <b>Izenak</b>        | 1050    | 48,52 |
| <b>Aditzak</b>       | 818     | 37,80 |
| <b>Adjektiboak</b>   | 21      | 0,97  |
| <b>Aditzondoak</b>   | 252     | 11,65 |
| <b>Postposizioak</b> | 23      | 1,06  |
| <b>Guztira</b>       | 2164    | 100   |

**2. taula:** eduki semantikodun kategorien kopuruak eta ehunekoak.



3. taulan, informatzaileen eduki semantikodun lemen ehunekoaren batezbestekoa ( $\bar{x}$ ), desbideratzea (sd), minimoa (min), maximoa (max) eta heina ematen ditugu.

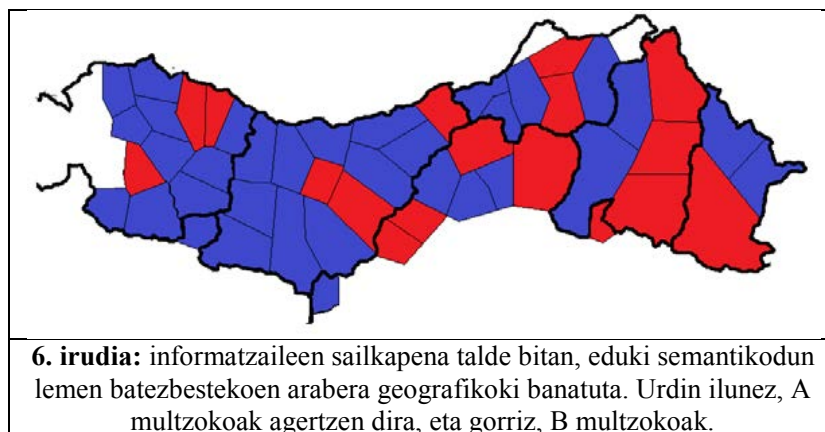
|                | $\bar{x}$ | sd    | min | max | heina |
|----------------|-----------|-------|-----|-----|-------|
| <b>SLren %</b> | 43,28     | 13,57 | 21  | 81  | 60    |

**3. taula:** eduki semantikodun lemen batezbestekoa, desbideratzea, minimoa, maximoa eta heina.

4. taulan, informatzaileak talde bitan banatuta ematen ditugu; A taldean, batezbestekoaren azpian kokatzen direnak daude, eta B taldean, oster, batezbestekoaren gainean kokatzen direnak; 6. irudiko mapan, informatzaileen hedadura ikus daiteke; mapa horretan ikus daitekeen bezala, ez dago zonalde homogeneorik irizpide honen arabera.

| Taldea | Informatzaileak   |
|--------|---|
| A      | Aiherra, Andoain, Aramaio, Azkaine, Azkoitia, Baigorri, Bergara, Bermeo, Donostia, Eibar, Elorrio, Ezpeleta, Gernika, Goizueta, Hazparne, Larrabetzu, Mallabia, Maule, Mungia, Ondarroa, Oñati, Ordizia, Orozko, Segura, Sunbilla, Urdiain, Zarautz, Zeanuri, Ziburu, Zornotza eta Zubieta. |
| B      | Donapauale, Ereño, Errezil, Garazi, Hondarribia, Irurita, Izura, Larresoro, Leitza, Lekeitio, Lekunberri, Lemoa, Lesaka, Luzaide, Mugerre, Tolosa eta Urdiñarbe.  |

**4. taula:** informatzaileen sailkapena talde bitan, eduki semantikodun lemen batezbestekoen arabera.

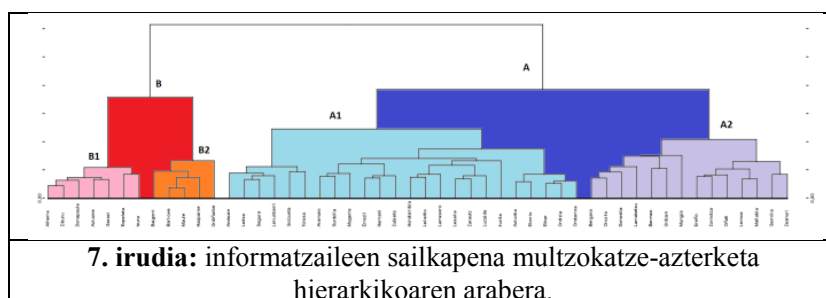


Testuak osatzeko, 259 eduki semantikodun lema erabili dira denetara; horietariko 44k (% 16,99) 0,46 baino maiztasun altuagoa daukate eta lema guztien % 74,95 biltzen dute (1622). Gainerako 215ek (% 83,01) lemen % 25,05 biltzen dute (542); euretariko 102 behin baino ez dira erabili.

### 3.2. Informatzaileen sailkapena

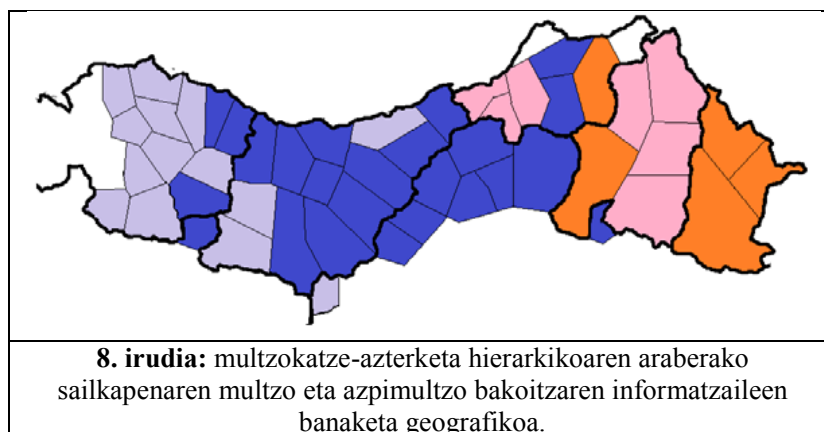
Atal honetan, bat-bateko testuen lexikoaren erabileraren arabera informatzaileekin egin ditugun hiru sailkapenen emaitzak aurkeztuko ditugu. Sailkapen horiek egiteko hiru teknika erabili ditugu multzokatze hierarkiko klasikoa, K-batezbestekoak eta FCM algoritmoa.

Multzokatze hierarkikoa erabili dugunean, 7. irudian erakusten den dendograma lortu dugu. Bertan ikusten den moduan, multzo nagusi bi agertzen dira, A eta B, eta bakoitzaren barruan beste azpimultzo bi: A1 eta A2, eta B1 eta B2; 5. taulan, multzo eta azpimultzo bakoitzeko agertzen diren informatzaileak zehazten dira.

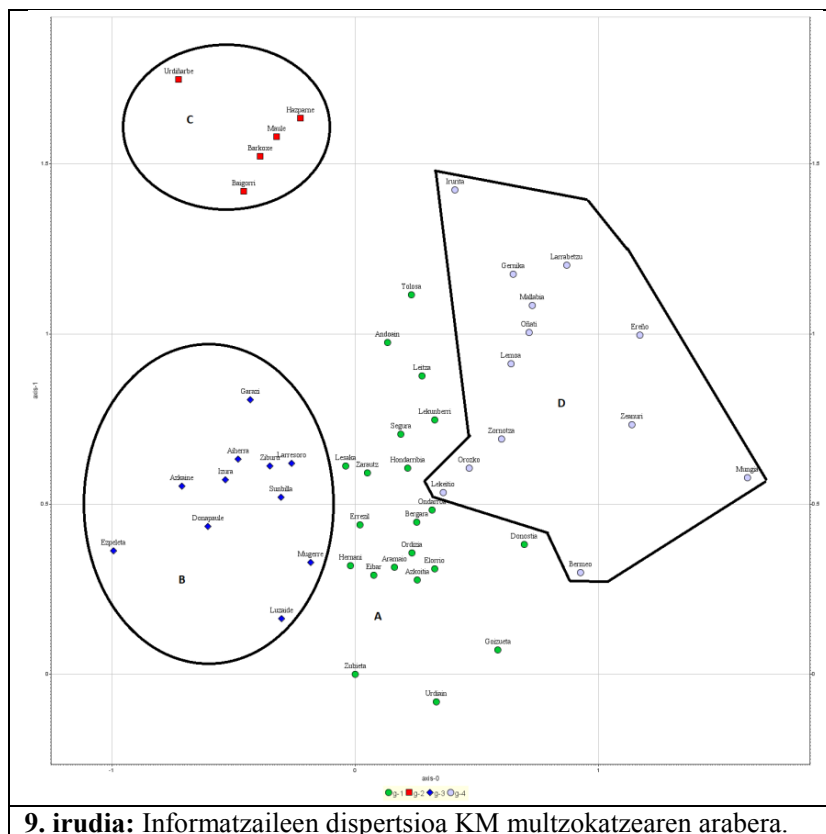


| Multzoa   | Azpi-multzoa | Informatzaileak  |
|---|--------------|--|
| A   | A1           | Andoain, Aramaio, Azkoitia, Eibar, Elorrio, Errezil, Goizueta, Hernani, Hondarribia, Irurita, Larresoro, Leitza, Lekeitio, Lekunberri, Lesaka, Luzaide, Mugerre, Ondarroa, Ordizia, Segura, Sunbilla, Tolosa, Zarautz eta Zubieta. |
|   | A2           | Bergara, Bermeo, Donostia, Ereño, Gernika, Larrabetzu, Lemoa, Mallabia, Mungia, Oñati, Orozko, Urdiain, Zeanuri eta Zornotza.  |
| B   | B1           | Aiherra, Azkaine, Donapauale, Ezpeleta, Garazi, Izura eta Ziburu.  |
|   | B2           | Baigorri, Barkoxe, Hazparne, Maule eta Urdiñarbe.  |
| <b>5. taula:</b> Multzokatze azterketa hierarkikoaren arabeko sailkapenaren multzo eta azpi-multzo bakoitzaren informatzaileak. |              |  |

Dendograman agertzen den lehen banaketa argia Hegoaldea eta Iparraldea bereizten duena da; Hegoaldean, beste azpimultzo bi nahiko argi agertzen dira, bata mendebaldean eta bestea erdialdean, eta Iparraldean ere, zonaldeak hain argiak ez diren arren, ekialdeko herriak (Hazparnerekin eta Baigorriarekin batera) eta mendebaldekoak bereiz agertzen dira (8. irudiko mapa).



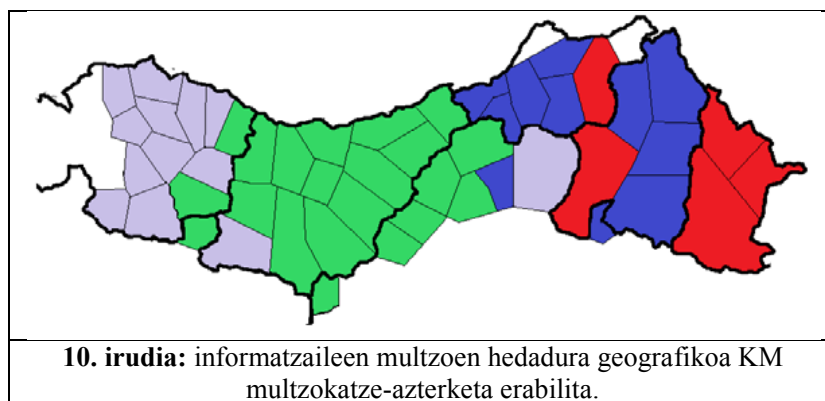
K-batezbestekoak izeneko multzokatze-azterketa lau multzoren arabera egin dugunean, 9. irudiko dispersio grafikoan agertzen diren multzoak lortu dira. Horretan ikusten den bezala, gehien aldentzen den multzoa C da; multzo honetan, Zuberoko herriak biltzen dira, Hazparnerekin eta Baigorriarekin batera. B multzoan, Iparraldeko gainerako herriak biltzen dira, Nafarroako Luzaiderekin eta Sunbillarekin batera. D multzoan, mendebaldeko herri gehienak biltzen dira, eta, azkenik, A multzoan, erdialdeko herri gehienak biltzen dira; 6. taulan, multzo bakoitzeko herri guztiak ikus daitezke.



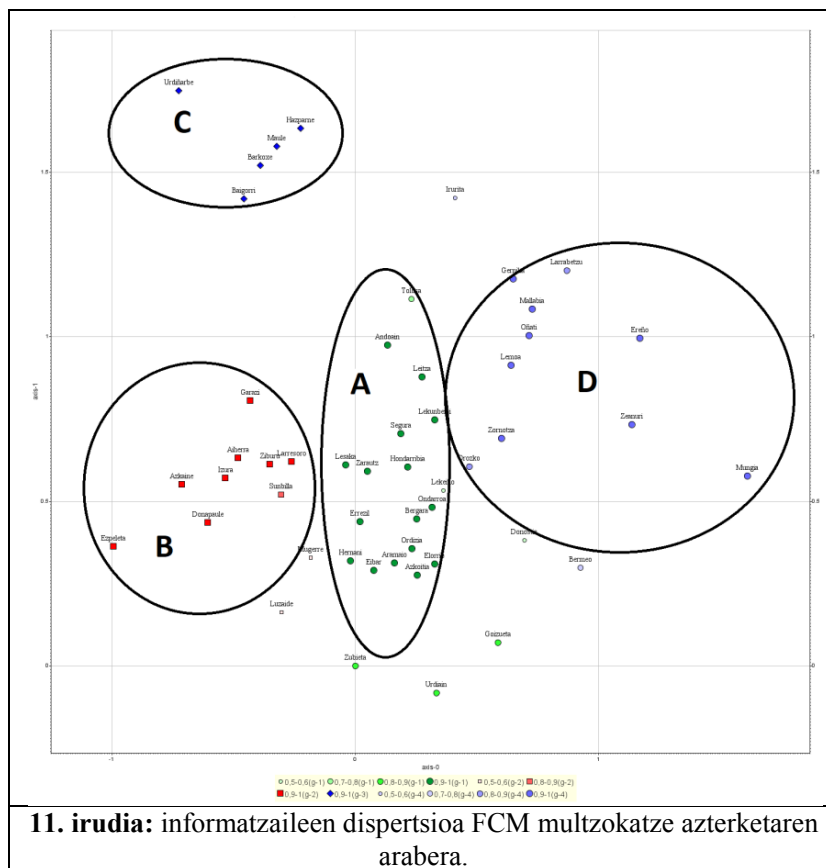
| Multzoa  | Informatzaileak  |
|----------|--|
| <b>A</b> | Andoain, Aramaio, Azkoitia, Bergara, Donostia, Eibar, Elorrio, Errezil, Goizueta, Hernani, Hondarribia, Leitza, Lekunberri, Lesaka, Ondarroa, Ordizia, Segura, Tolosa, Urdiain, Zarautz eta Zubieta. |
| <b>B</b> | Aiherra, Azkaine, Donapauze, Ezpeleta, Garazi, Izura, Larresoro, Luzaide, Mugerre, Sunbilla eta Ziburu.  |
| <b>C</b> | Baigorri, Barkoxe, Hazparne, Maule eta Urdiñarbe.  |
| <b>D</b> | Bermeo, Ereño, Gernika, Irurita, Larrabetzu, Lekeitio, Lemoa, Mallabia, Mungia, Oñati, Orozko, Zeanuri eta Zornotza.   |

**6. taula:** informatzaileen sailkapena KM multzokatzearen arabera.

10. irudiko mapan multzoen hedadura erakusten da; A multzoari kolore berdea dagokio, B multzoari urdin iluna, C multzoari gorria, eta D multzoari more argia.



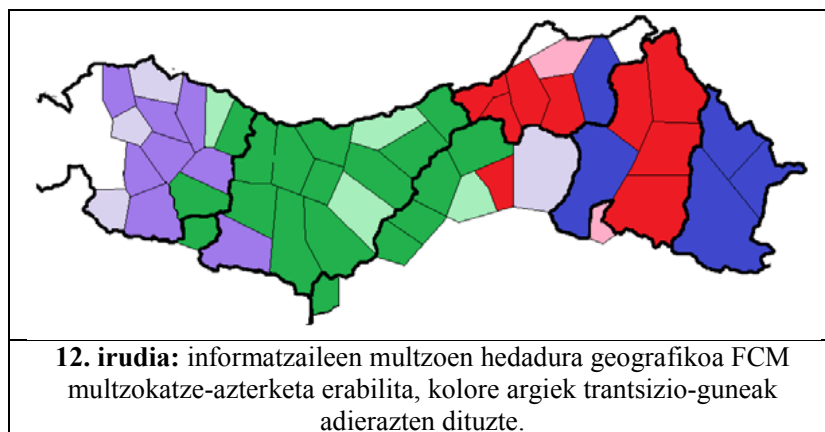
FCM izeneko algoritmoa erabilia multzokatze-azterketa lau multzoren arabera egitean, 11. irudiko dispersio grafikoan agertzen diren multzoak lortu dira. Multzo hauek bat datoz oro har aurreko teknikaren bidez lortu ditugunekin. Aldea trantsizio-gunei dagokie, eta barnekotasun-funtzioaren arabera kalkulatu dugu; horrela, bada, multzoaren barnekotasun osoa adierazten da funtzioa 0,85en eta 1en artean dagoenean. Multzo bakoitzaren herrien kopuruak eta barne-funtzioen tartekak 7. taulan ikus daitezke.



| Multzoa | Azpi-multzoa   | Informatzaileak   |
|---------|----------------|---|
| A       | A1 (0,87-1)    | Andoain, Aramaio, Azkoitia, Bergara, Eibar, Elorrio, Errezil, Goizueta, Hernani, Hondarribia, Leitza, Lekunberri, Lesaka, Ondarroa, Ordizia, Segura, Urdiain eta Zarautz. |
|         | A2 (0,52-0,81) | Donostia, Lekeitio, Tolosa eta Zubieta  |
| B       | B1 (0,89-1)    | Aiherra, Azkaine, Donapauale, Ezpeleta, Garazi, Izura, Larresoro, Sunbilla eta Ziburu.  |
|         | B2 (0,53-0,57) | Luzaide eta Mugerre.  |
| C       | (0,99-1)       | Baigorri, Barkoxe, Hazparne, Maule eta Urdiñarbe  |
| D       | D1 (0,96-0,99) | Ereño, Gernika, Lemoa, Mallabia, Mungia, Oñati, Zeanuri eta Zornotza.   |
|         | D2 (0,51-0,81) | Bermeo, Irurita, Larrabetzu eta Orozko  |

**7. taula:** informatzaileen sailkapena FCM multzokatze-azterketaren arabera, azpimultzoetan parentesien artean barnekotasun-funtzioen tartekak adierazten dira.

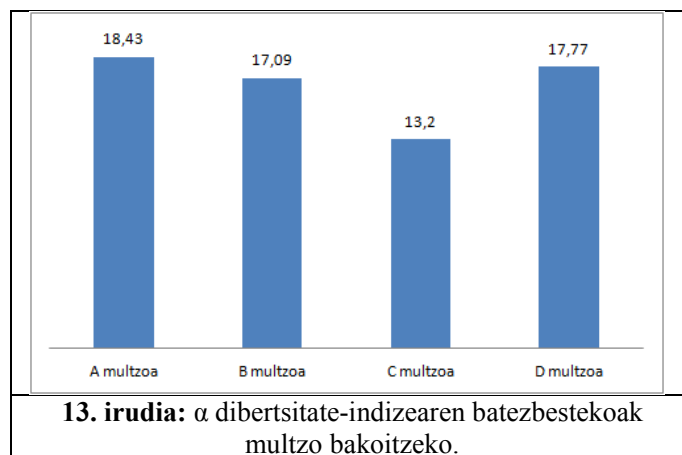
12. irudiko mapan multzo bakoitzaren hedadura geografikoa erakusten da, arteragoko multzokatzean erabilitako kolore berberak erabili ditugu, baina oraingo honetan trantsizio-guneak kolore argiekin adierazten dira.



Egin ditugun lau multzoen barneko koherentzia aztertzeko, metodologia-atalean esaneko moduan, hiru neurri erabili ditugu; 8. taulan,  $\alpha$  dibertsitate-indizearen (De Cáceres et al., 2003) batezbestekoak, desbideratzeak eta aldakuntza koefizienteak (Ak.) ematen ditugu multzo bakoitzeko eta orokorrean; horretan ikus daitekeen bezala, indizerik altuena A multzoak dauka, eta baxuena C multzoak (13. irudiko grafikoa batezbestekoak erkatu daitezke). Aldea Kruskal-Wallis probaren arabera estatistikoki esanguratsua da ( $\chi^2 =$  (a.m.: 3) 8,998;  $p = 0,029$ ).

|           | A multzoa | B multzoa | C multzoa | D multzoa | Orotara |
|-----------|-----------|-----------|-----------|-----------|---------|
| $\bar{x}$ | 18,43     | 17,09     | 13,20     | 17,77     | 17,44   |
| sd        | 2,98      | 3,53      | 2,28      | 4,28      | 3,66    |
| Ak.       | 0,16      | 0,21      | 0,17      | 0,24      | 0,21    |

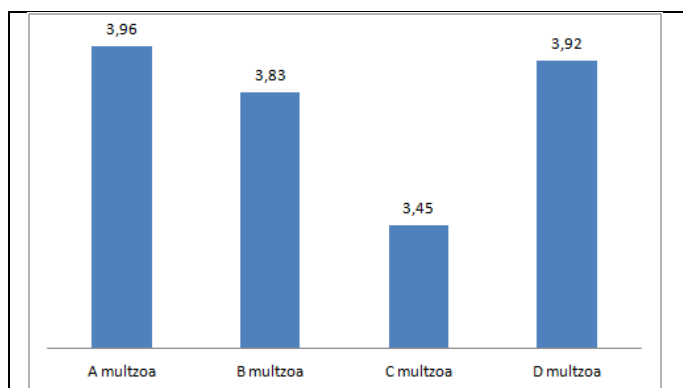
**8. taula:**  $\alpha$  dibertsitate indizearen emaitzak multzo bakoitzeko.



9. taulan, multzo bakoitzaren Shannon-en entropiaren balioaren (Roberto et al., 2012; Wojciechowski eta Gorzynski, 2016) batezbestekoa, desbideratzea eta aldakuntza koefizientea erakusten dira; aurrekoan bezala, entropiarik handiena A multzoak dauka, eta baxuena C multzoak (14. irudiko grafikoa batezbestekoak erkatu daitezke). Aldea Kruskal-Wallis probaren arabera estatistikoki esanguratsua da ( $\chi^2 =$  (a.m.: 3) 11,136;  $p = 0,011$ ).

|           | A multzoa | B multzoa | C multzoa | D multzoa | Orotara |
|-----------|-----------|-----------|-----------|-----------|---------|
| $\bar{x}$ | 3,96      | 3,83      | 3,45      | 3,92      | 3,87    |
| sd        | 0,24      | 0,27      | 0,21      | 0,34      | 0,31    |
| Ak.       | 0,06      | 0,07      | 0,06      | 0,09      | 0,08    |

9. taula: entropiaren emaitzak multzo bakoitzeko.

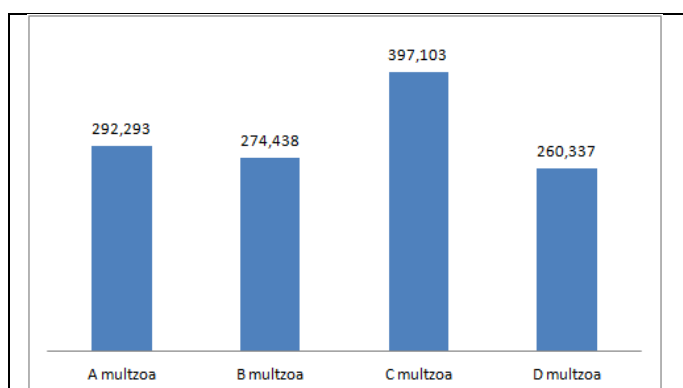


14. irudia: entropiaren batezbestekoak multzo bakoitzeko.

Azkenik, 10. taulan multzo bakoitzaren Yuleren Kren balioaren (Roberto et al., 2012; Wojciechowski eta Gorzynski, 2016) batezbestekoa, desbideratzea eta aldakuntza koefizientea ikus daitezke; horretan ikusten den moduan, altuena C multzoak dauka, hots, ekialderen dagoenak, eta baxuenak, berriz, D multzoak, mendebalderen dagoenak, alegia (15. irudiko grafikoan batezbestekoak erkatu daitezke).

|           | A multzoa | B multzoa | C multzoa | D multzoa | Orotara |
|-----------|-----------|-----------|-----------|-----------|---------|
| $\bar{x}$ | 292,293   | 274,438   | 397,103   | 260,337   | 290,536 |
| sd        | 100,237   | 77,683    | 111,425   | 90,522    | 99,06   |
| Ak.       | 34,293    | 28,306    | 28,059    | 34,771    | 34,096  |

10. taula: Yuleren K indizearen emaitzak multzo bakoitzeko.



15. irudia: Yuleren K indizearen batezbestekoak multzo bakoitzeko.

Aldea Kruskal-Wallis probaren arabera estatistikoki esanguratsua ez den arren, C multzoak gainerako multzoekin dauzkan aldeak Mann-Whitneyren probaren arabera estatistikoki esanguratsuak dira kasu guztietan: C/A ( $p = 0,034$ ), C/B ( $p = 0,036$ ) eta C/D ( $p = 0,020$ ).

Erabili ditugun hiru neurrien araberako multzoen barneko koherentzia handia dela esan dezakegu; izan ere, kasu guztietan C multzoa (ekialdeko informatzaileak biltzen dituena) da gehien aldentzen dena besteetatik; bestalde, Iparraldeko informatzaileak dira lehen bi neurrien arabera gehien aldentzen direnak.

## 5. Ondorioak eta hurrengo lanak

Azken atal honetan datuen azterketatik atera daitezkeen ondorio nagusiak laburbilduko ditugu. Lehenengo eta behin esan behar dugu bat-bateko testuen lexikoaren erabilera metodo kuantitatiboan bidez lantzeak balio diezagukeela azpian dautzan egitura geolinguistikoak azaleratzeko.

Euskara estandarren eraginagatik ere, lau zonalde nagusi eta argi nabarmentzen dira azterketa egindakoan: (a) Mendebaldea (Bizkaiko informatzaile gehienak eta Oñati), (b) Erdialdea (Gipuzkoa eta Nafarroako informatzaileak), (c) Lapurdi eta Nafarroa Behereko informatzaile gehienak, eta (d) Zuberoko informatzaileak, Baigorriekin eta Hazparnerekin batera.

Ikusi dugun bezala,  $\alpha$  dibertsitatearen, Shannonen entropiaren eta Yule-ren Kren emaitzak ondo erlazionatzen dira egindako sailkapen eta banaketa geografikoarekin.

Era honetako azterketa egitean funtsezko baldintza da egin beharreko metodologia koherentea izatea eta ondo aplikatzea. Jarraian, azterketa osoan segitu beharreko ibilbidearen urratsak zehazten dira:

- (1) Informatzaileek testua ekoizteko atazaren diseinua.
- (2) Grabazioak.
- (3) Transkripzioak.
- (4) Lematizazioa:  
etiketen gaineko erabakiak hemen hartu behar dira.
- (5) Lemen maiztasunaren kalkulua: orokorra zein informatzaileko.
- (6) Lemen aukeraketa matrizea osatzeko.
- (7) Lemen matrizeen osaketa: kopurua eta maiztasunaren arabera.
- (8) Multzokatze-azterketaren egitea:  
distantzia aukeratzea,  
multzokatze-metodoa aukeratzea,  
multzo kopurua aukeratzea KM eta FCM teknikak aplikatzeko.
- (9) Multzoen barne koherentzia aztertzekeo neurriak kalkulatzeko ( $\alpha$ , entropia, K).
- (10) Emaitzen interpretazioa egitea.

Hemen saio moduan aurkeztu dugun hau nahi genuke aurrerantzean kilikagarri izan ledin beste ikertzaile batzuk bide hauek urratzen hasteko. Antzeko metodoak erabil daitezke aldagai sozialen arabera informatzaileak sailkatzeko (aldagaiak ama-hizkuntza, generoa, adina, ikasketa-maila, eta abar izan daitezke); metodo kuantitatibo hauen bidez euskara



ikasleek lexikoan egiten dituzten aurrerapenak neur daitezke, era berean hizkuntzaren jabeakuntza-prozesua kuantifikatzeko ere erabil daitezke eta beste hamaika esparrutan.

## 6. Bibliografia

- Aurrekoetxea, G. (1992) “Nafarroako euskara: azterketa dialektometrikoa”, *Uztaro* 5, 59-109.
- Aurrekoetxea, G. (1995) *Bizkaieraren egituraketa geolinguistikoa*, UPV/EHU-ren Argitalpen Zerbitzua, Leioa.
- Aurrekoetxea, G. (2003) “Euskalkiak estandarren uholdepean, Arratiako kasua” in Mendebalde Kultura Alkartea: *Ahozkotatuna aztergai*. Bilbo.
- Aurrekoetxea, G. (2004) “Estandar eta dialektoen arteko bateratze-joerak, ikuspuntu teorikotik begirada bat”, *Uztaro* 50.
- Aurrekoetxea, G. (2005) “Nafarroako Euskararen Sailkapenaz” in Etxebarria, P. eta Knör, H. (2005) *Nerekin yaio nun. Txillardegiri omenaldia*. Euskaltzaindia eta EHU, Iker 17, Bilbo. (109-124)
- Aurrekoetxea, G. (2008) “Bariazio soziolinguistikoa Dimako euskararen”, *Euskalingua* 12.
- Aurrekoetxea, G. (2009) “Bariazio sinkronikoa aztertzeko metodologia(k)”. *Lapurdum XIII*, 43-59.
- Aurrekoetxea, G. (2016) “Distantzia geografikoaren eta hizkuntza distantziaren arteko korrelazioa” in Aurrekoetxea, G., Makazaga, J. M. eta Salaberri, P. (arg.). *Hire Bortatxoan: Txipi Ormaetxea Omenduz*. UPV/EHU. Bilbo
- Aurrekoetxea, G. eta J. L. Ormaetxea (eds.) (2010) *Tools for Linguistic Variation*, Bilbao: UPV-EHU, Supplements of the Anuario de Filología Vasca “Julio de Urquijo”, LIII.
- Aurrekoetxea, G. eta Ormaetxea, J.L. (2006) “Euskararen Atlas Sozio-Geolinguistikoa” ikerketa proiektua”, *Euskalingua* 9.
- Aurrekoetxea, G.; Gandarias, L., Gaminde, I. eta Iglesias, A. (2014a) “Variación prosódica en vasco: áreas acentuales” in Yolanda Congosto Martín, María Luisa Montero Curiel y Antonio Salvador Plans (eds.) *Fonética Experimental, Educación Superior e Investigación*, Arco Libros, Madrid.
- Aurrekoetxea, G.; Gaminde, I.; Gandarias, L. eta Iglesias, A. (2014b) “Prosodic variation in the Basque language: intonational areas”. in Díaz-Negrillo, A. & Díaz Pérez Francisco Javier (eds) *Specialisation and Variation in Language Corpora*, Peter Lang AG, International Academic Publishers, Berna.
- Barrutia, E. (1996) *Bermeo eta Mundakako Arrantzaleen Hiztegia*. UEU: Bilbo.
- Barrutia, E. (2000) *Bizkaiko arrantzaleen hiztegia. Léxico marineru vizcaino*. Labayru Ikastegia eta BBK: Bilbo.
- Basauri, S. eta Sarasua, A. (2003) *Eibarko Hiztegi Etnografikoa: Natura, Baserria, Lanbideak, Ohiturak*. Eibarko Udala, Donostia
- Beckner, C.; Blythe, R.; Bybee, J.; Christiansen, M. H.; Croft, W.; Ellis, N. C.; Holland, J.; Ke, J.; Larsen-Freeman, D. eta Schoenemann, T. (2011). “La lengua es un sistema adaptativo complejo” *Linguística en la Red*. [http://www.linred.es/articulos\\_pdf/LR\\_articulo\\_04092011.pdf](http://www.linred.es/articulos_pdf/LR_articulo_04092011.pdf)
- Bonvin, A. eta Lambelet, A. (2017) “Algorithmic and subjective measures of lexical diversity in bilingual written corpora: a discussion” *Corela*, HS-21; 1-16

- De Cáceres, M.; Oliva, F. eta Font, X. (2003) "Ginkgo, un programa de análisis multivariante orientado a la clasificación basada en distancias" 27 Congreso Nacional de Estadística e Investigación Operativa Lleida, 8-11 de abril de 2003. <https://dialnet.unirioja.es/servlet/articulo?codigo=1153561>
- Eguskiza, N.; De Pablo, I. eta Gaminde, I (2017) "Hizkuntz hizkeren sailkapena metodo kuantitatiboen bidez" agertzeko Psikodidaktika kongresuan
- Ellis, N. C. (2011). "The Emergence of Language as a Complex Adaptive System" In James Simpson (Ed.), *Routledge Handbook of Applied Linguistics*. Routledge/Taylor Francis.
- Ensunza, A. (2015) *Busturialdeko euskararen hizkuntza-aldakortasuna denboran eta espazioan*. Doktorego tesia; UPV/EHU
- Etxebarria, M. (1996) "Disponibilidad léxica en escolares del País Vasco" *Revista española de lingüística*-26, 301-326
- Euskaltzaindia (2008tik gora) *Euskararen Herri Hizkeren Atlas*. Euskadiko Kutxa, Euskaltzaindia, Bilbo
- Gaminde, Iñaki (1997): *Gatikako Euskaraz*. Gogoz Euskara Taldea. Gatika.
- Gaminde, Iñaki (1999): *Bakio Berbarik Berba*. Labayru, Bakioko Udala.
- Gaminde, Iñaki (2000): *Zamudio Berbarik Berba*, Labayru, Bilbo
- Gaminde, Iñaki (2001): *Meñaka Berbarik Berba*, Bizkaiko Diputazioa.
- Gaminde, I. (2007). *Bizkaian Zehar: Euskararen Ikuspegi Orokorra*. Mendebalde Kultura Alkartea eta Bizkaiko Foru Aldundia: Bilbo
- Gaminde, I.; Romero, A.; Legarra, H. (2012) *Gramatika eta hizkuntza bariazioa Bermeon*, Bermeoko Udala: Bermeo.
- Gaminde, I.; Romero, A.; Etxebarria, A. eta Eguskiza, N. (2016a) "Bizkaiko aditz laguntzaileen bilakaeraren azterketaz" in Iglesias, A.; Romero, A. eta Ensunza, A. *Linguistic variation in the Basque language and education – II / Euskararen bariazioa eta bariazioaren irakaskuntza – II*. UPV/EHU, Bilbo.
- Gaminde, I., Etxebarria, A., Eguskiza, N., Romero, A. eta Unamuno, L. (2016b) "Lexikoaren bariazioa eta multzokatze-azterketa" *Euskalingua*-28, 19-3
- Gaminde, I., Olalde, A., Gaminde, U. & Etxebarria, A (2016c). "Populazioaren sailkapena lexikoaren arabera" *Ikastorratza*. e-Revista de Didáctica, 16, 68-84. Retrieved from [http://www.ehu.es/ikastorratza/16\\_alea/4.pdf](http://www.ehu.es/ikastorratza/16_alea/4.pdf)
- Gaminde, I.; Olalde, A.; Etxebarria, A.; Eguskiza, N. eta Gaminde, U. (2017a): *Hizkuntza Aldakortasuna Larrabetzun*. Larrabetzuko Udala. Bilbo.
- Gaminde, I.; Eguskiza, N.; Romero, A. eta Etxebarria, A. (2017b) "Informatzaileen sailkapenerako arazo metodologiko batzuen gainean" in Unamuno, L.; Romero, A., Etxebarria, A. eta Iglesias, A.: *Linguistic variation in the basque language and education - III [Recurso electrónico] = Euskararen bariazioa eta bariazioaren irakaskuntza – III*. UPV/EHU. Bilbo. (29-42 or.)
- Goebel, H. (1992) "Problèmes et méthodes de la dialectométrie actuelle (avec application à l' AIS)", in Euskaltzaindia/Académie de la Langue Basque éd., *Nazioarteko dialektologia biltzarra. Agiriak/Actes du Congrès International de Dialectologie* (Bilbo/Bilbao 1991), Bilbo/Bilbao, 429-475.
- Goebel, H. (2010) Introducción a los problemas y métodos según los principios de la Escuela Dialectométrica de Salzburgo. in Aurrekoetxea, G. eta Ormaetxea, J. L. (2010) *Tools for Linguistic Variation*. UPV/EHU, Bilbo.

- Gregori-Signes, C eta Clavel-Arroitia, B. (2015) "Analysing lexical density and lexical diversity in university students' written discourse" *Procedia-Social and Behavioral Sciences* 198, 546-556.
- Hout, R., eta Vermeer, A. (2007). Comparing measures of lexical richness. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and Assessing Vocabulary Knowledge* (Cambridge Applied Linguistics, pp. 93-115). Cambridge: Cambridge University Press.
- Iglesias, A. (2014) *Igorreko hizkeraren azterketa dialektologikoa*, Doktorego tesia, UPV/EHU
- Kim, J. (2014) "Predicting L2 Writing Proficiency Using Linguistic Complexity Measures: A Corpus-Based Study" *English Teaching*, 69, 4; 27-51. (Neurriak, L2, idatzia)
- Léonard, J. L., Heinsalu, E., Patriarca, M. eta Darlu, P. (2015) "Modeling regional variation from eas: complexity and communal aggregates" in Aurrekoetxea, G.; Romero, A. eta Etxebarria, A. *Linguistic Variation in the Basque an Education-I/Euskararen Bariazioa eta Bariatioaren Irakaskuntza-I* (50-58or.) Bilbo. UPV/EHU
- López Morales, H. (2011) "Los índices de riqueza léxica y la enseñanza de lenguas" in De Santiago, J.; Bongaerts, H.; Sánchez, J. J. eta Seseña, M. (coord.) *Del texto a la lengua La aplicación de los textos a la enseñanza-aprendizaje del español L2-LE*, 15-28
- López Rivera, J. J. (2013) "Aplicación del marco de los sistemas complejos adaptativos a un modelo de variación lingüística" *Moenia* 19, 5-24.
- Lujanbio, O. (2016). *Hizkuntzaa-aldakortasuna euskarán. Nafarroa ipar-mendebaldeko bi udalerritan egindako azterketa*. [Doktoretza-tesia]. UPV/EHU.
- Moreno Cabrera, J. C. (2008) *El nacionalismo lingüístico: Una ideología destructiva*. Ediciones Península, Bartzelona.
- Ormaetxea, J.L. (2008) "Otxandioko hizkera: adinaren araberako bariazioa", *FLV* 108.
- Pardo, A. eta Ruiz, M. A., (2002) *SPSS 11 Guía para el análisis de datos*. McGraw-Hill, Madril.
- Pérez López, C. (2005) *Métodos estadísticos avanzados con SPSS*. Thomson, Madril.
- Roberto, J. A.; Martí, M. A. eta Salamó, M. (2012) Análisis de la riqueza léxica en el contexto de la clasificación de atributos demográficos latentes" *Procesamiento de Lenguaje Natural*, 48; 97-104.
- Sisková, Z. (2012) "Lexical Richness in EFL Students' Narratives" *Language Studies Working Papers*, 4; 26-36.
- Van Gijssel, S., Speelman, D., & Geeraerts, D. (2005). A Variationist, Corpus Linguistic Analysis of Lexical Richness. [http://www.ling.arts.kuleuven.be/qlvl/prints/VanGijssel\\_Speelman\\_Geeraerts\\_2005final\\_Variationist\\_Lexical\\_Richness.pdf](http://www.ling.arts.kuleuven.be/qlvl/prints/VanGijssel_Speelman_Geeraerts_2005final_Variationist_Lexical_Richness.pdf)
- Vermeer, A. (2000) "Coming to grips with lexical richness in spontaneous speech data" *Language Testing* 17 (1), 65-83.
- Wojciechowski, A. eta Gorzynski, K. (2016) "A Method for Measuring Similarity of Books" in Vetulani, Z.; Uszkoreit, H. eta Kubis, M. (2016) *Human Language Technology. Challenges for Computer Science and Linguistics: 6th Language and Technology Conference, LTC 2013, Poznań, Poland, December 7-9, 2013*. (161-176).
- Xiaofei, L. (2012) "The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives", *The Modern Language Journal*, 96; 190-208.

## GESTURES SYNCHRONIZED WITH SPEECH IN INFANT LANGUAGE

Alfonso Igualada  
Universitat Oberta de Catalunya  
[aigualada@uoc.edu](mailto:aigualada@uoc.edu)

### 1. Gestures and speech are an integrated system

In the last few decades, gesture studies have started to investigate the interplay between gestures and speech as being part of the same linguistic system. Within this field, gesture and speech have been argued to be closely linked, forming an integrated system (e.g., Kendon, 1980; McNeill, 1992; McNeill, 2005; Kelly, Ozyürek, & Maris, 2010). According to McNeill (1992), gestures are defined as communicative acts which are closely related with the speech stream at different levels of linguistic analysis. The author suggested that gestures are co-expressive with speech at the phonological, semantic and pragmatic level and proposed the following three rules of synchronization between gesture and speech.

- The phonological synchrony rule predicts that prominences of gesture and speech are temporally aligned.
- The semantic synchrony rule predicts that gestures and speech are related to the same idea or concept.
- The pragmatic synchrony rule predicts that both modalities share the intentional function.

Gestures have traditionally been differentiated by their form and function. Following McNeill (1992) I will describe the four types (or dimensions) of gestures: representational, deictic, conventional, and last but not least, beat gestures.

Representational gestures (including iconic and metaphoric gestures) represent objects, actions, or relations by coding an aspect of their referent's shape or movement. Iconic gestures can express tangible information (e.g., representing the shape of an object with the hands, like producing a cup hand gesture to represent a glass). Metaphoric gestures are related to more abstract concepts (e.g., touching the head with two fingers to represent the verb "think", or moving the arm forward while producing circular movements with the hand to represent the concept "future").

Deictic gestures, such as pointing gestures, are hand and arm gestures which serve to direct attention toward a specific object or event of reference in the surrounding environment. These gestures include requesting (extending the arm toward an object, location or person, sometimes with a repeated opening and closing of the hand), showing (holding up an object in the adult's line of sight), giving (transferring an object to another person), and pointing (index finger or full hand extended towards an object, location, person or event).

Thirdly, conventional gestures are culturally shared symbols, with an arbitrary form and meaning within a given community. For example, a palm up hand gesture shaking with repetitive movements from right to left can serve to warn someone that he/she did something wrong according to the gesturer's opinion, with the potential consequence that there will be some kind of punishment. Another example would be a common “hi” hand gesture.

Finally, beat gestures have been defined as rhythmical movements of the hands which are timed together with speech, specifically with prosodic prominence in speech (e.g., Loehr, 2012; Shattuck-Hufnagel, Ren, Mathew, Yuen, & Demuth, 2016). Beats have been typically described as not having a clear semantic meaning (i.e., non-representational gestures). In contrast with representational gestures, beat gestures do not add propositional content to a given utterance (McNeill, 1992, 2005; Kendon, 1995) but are rather used to mark “the word or phrase they accompany as being significant (...) for its discourse pragmatic content” (McNeill, 1992:15). Beat gestures have been associated with functions typically linked to prosody, such as, focus marking (i.e., prominence) and discourse structure marking (e.g., Loehr, 2012; Wagner, Malisz, & Kopp, 2014; Shattuck-Hufnagel et al., 2016). Research has shown that beat gestures are temporally synchronized with prosodic markers of prominence (i.e., pitch accents) (Yasinnik, Renwick, & Shattuck-Hufnagel, 2004; Jannedy, & Mendoza-Denton, 2005; Loehr, 2012; Shattuck-Hufnagel, et al., 2016). In addition, an increase in prominence perception has been reported when words are produced together with hand gestures (Krahmer & Swerts, 2007) and head/ facial beat gestures (Moubayed, Beskow, & Granström, 2010). Typically, the movements of hand gestures occur together with head and eyebrow movements, which together signal the privileged status of a given piece of discourse in a multimodal fashion.

McNeill's classification has been extensively used in gesture research (for example, see Gullberg, DeBot, & Volterra, 2008, or Cartmill, Demir, & Goldin-Meadow, 2012, for a review). Even though gesture research has typically used this classification as a close categorization of gestures, examples are frequently found where more than one dimension can be observed at the same time. Regarding this conceptual issue on the use of this classification as dimensions that interplay rather than closed clusters, McNeill (1992) provided the following example: a pointing gesture may be temporally aligned with prosodic patterns and also act as a beat gesture. Thus, according to this view any type of gesture associated with a prosodic prominence might be functioning as a multimodal marker of prominence (i.e., a beat gesture).

## **2. Multimodal markers of prominence in language**

Recent research exploring the relationship between prosodic and gestural prominence has shown that prosodic prominence in speech (i.e., expressed through prosodic pitch accents) is typically produced simultaneously with more prominent gestural and articulatory features (such as head nods, eyebrow movements, beat gestures, exaggerated articulation, etc. (see, for example, Dohen, 2009; Shattuck-Hufnagel et al., 2016; Swerts & Krahmer, 2008).

Prosody signals different functions that have been extensively studied across different language types. Among others, prosody functions to structure speech stream into different levels of organization (e.g., word segmentation, syntactic phrase boundaries, and discourse structure), as well as speaker-specific information (e.g., speakers' identity, speech style). Prosody can also signal pragmatic meanings related to the speakers' intentions, such as marking a phrase as a statement or a question or expressing degrees of commitment about the information expressed (e.g., Cole, 2015; Ladd, 2008; Pierrehumbert & Hirschberg, 1990). Prosody may also perform an important semantic function, such as signaling informational focus (or novelty marking) or contrastive focus. Across languages, focus marking has been typically related to the use of pitch accentuation associated with the relevant words (Cole, 2015; Ladd, 2008; Pierrehumbert & Hirschberg, 1990).

Evidence increases on the need of studying beat gestures (or the beat gesture dimension) from a perspective that integrates those gestural markers with prominence functions related to prosody. Interestingly, beat gestures in natural speech have been reported to encode discourse structure functions, such as prosody. In an example, Guellaï, Langus, and Nespors (2014) assessed whether gestures accompanying speech can also signal phrase boundaries, a typical prosodic function. The linguistic stimuli in this study were ambiguous sentences which could differ in their meaning depending on where the prosodic boundary was placed. Video recordings were manipulated to create mismatches between the alignment of prosody and gestures used in a discourse structuring function (chunking sentence). The results showed that adults tended to more often choose the meaning signaled by gestures during mismatches of the ambiguous sentences. The authors conclude that prosody seems not to be the exclusive marker of the prosodic phrasing function. Similarly, beat gestures have been associated with functions typically linked to prosody, such as focus marking (i.e., prominence) and discourse structure marking in both adult populations and children (e.g., Loehr, 2012; Wagner et al., 2014; Shattuck-Hufnagel et al., 2016) (See section 1.1 for the definition of beat gestures).

Regarding the prominence function, perception studies on this field have revealed that that misalignments between visual beats and prosodic pitch accents have a detrimental effect on the perception of prominence (e.g., Leonard & Cummins, 2011; Treffner, Peter, & Kleidon, 2008). Krahmer and Swerts (2007) found that the presence of beat gestures has a significant effect on the perceived prominence independently from prosodic prominence. In their experiment, participants were exposed to the sentence “*Amanda goes to Malta*” in which the two target words were associated with combinations of the presence or absence of prosodic and gestural prominence. When participants saw a manual beat gesture on the focused word, and regardless of the presence or absence of prosodic prominence, this increased its perceived prominence and decreased the prominence of another target word. However, the strongest effect was produced by the multimodal combination of prosodic prominence and visual prominence. Thus, there is a strong perceptual connection between the presence of prosodic prominence (or pitch accentuation) and manual, facial and head beat gestures (Krahmer & Swerts, 2007; Moubayed, et al., 2010).

One of the linguistic functions of beat gestures is thus related to the signaling of prosodic focus, which has been compared to a yellow highlighter, that is, to emphasize

information in the speech stream. McNeill (1992:15) stated that “the semiotic value of a beat lies in the fact that it indexes the word or phrase it accompanies as being significant (. . .) for its discourse pragmatic content”. As in the case of prosodic prominences, beat gestures typically align new information of the discourse context and highlight it. For example, when enumerating the features of a newly introduced character in the story: “his GIRLfriend, ALice, Alice WHItE” (example from McNeill, 2012).

Beat gestures might be studied as the gestural realization of prosodic accentuation as a marker of prominence. That is, to make particular items stand out from their surrounding non-prominent elements (e.g., Terken, 1991; Wagner, et al. 2015). Thus, multimodal marking of prominence might be a focus signaling strategy related to prosody. Based on previous evidence we expect that prominence provided by temporal gesture-speech synchronicities will positively impact on the development of language processing.

### **3. Developmental benefits of temporally synchronized non-representational gestures.**

Temporal integration has been related to perception of prominence when they are naturally aligned. Studies on multimodal integration have shown how the coupling of visual input with oral speech sound provides redundancy of the event in a unitary instance, leading to an increase in intelligibility (Lewkowicz & Hansen-Tift, 2012; van Wassenhove, Grant, & Poeppel, 2007) and gesture-speech sound integration (Esteve-Gibert, et al., 2015; Kraemer & Swerts, 2007). The ability to integrate unitary cues in a redundant way is proposed to be an adaptive strategy related to language comprehension. In fact, research with children with specific language impairments has suggested that children’s difficulty in processing linguistic information might be related to difficulties in integrating multimodal cues of speech perception (Pons, Andreu, Sanz-Torren, Buil-Legaz, & Lewkowicz, 2013). In line with these results, according to the Dual Coding Theory, memory traces are enhanced by the integration of multimodal channels, in this case redundancy between information conveyed by representational gestures and verbal information (Clark and Paivio, 1991).

Research on gestures has extensively reported the beneficial results of representational gestures for various linguistic abilities. While research has shown that adults can benefit from the presence of beat gestures in word recall tasks, very few studies have explored this with preschool children (Austin & Sweller, 2014). Igalada, Esteve-Gibert & Prieto (2017) investigated whether the presence of beat gestures helps children to recall information when these gestures have the function of singling out a linguistic element in its discourse context. One hundred and six 3-to-5-year-old children were asked to recall a list of words within a pragmatically relevant context (i.e., a story-telling activity) in which the target word was accompanied or not by a beat gesture. Results showed that children recalled the target word significantly better when it was accompanied by a beat gesture than when it was not, indicating a local recall effect. Moreover, the recall of adjacent non-target words did not differ depending on the condition, revealing that beat gestures seem to have a strictly local highlighting function (i.e., no global recall effect). Interestingly, even though the effect was weak, the analyses show that the children’s ability to recall the target word improved with age only in the beat condition. These results demonstrate that preschoolers benefit from the

pragmatic contribution offered by beat gestures when they function as multimodal markers of prominence.

Pointing gestures' temporal synchronization to speech has been shown to develop around the age of 12 months. Moreover, children have been shown to modify their communicative responses depending on adult's joint attention to the children's object of interest during natural interaction (Miller & Lossia, 2013; Miller & Gros-Louis, 2013; Gros-Louis & Wu, 2012), and during controlled experimental settings (Liszkowski, Albrecht, Carpenter & Tomasello, 2008). However, less is known about whether children produced synchronized gesture-speech productions during experimentally controlled pragmatic context (i.e., varying on the adult's joint attention), and whether this ability is related to later language abilities (Murillo & Belinchón, 2012; Wu & Gros-Louis, 2014). Igualada, Bosch & Prieto (2015). investigates the degree to which an infants' use of synchronous gesture-speech combinations during controlled social interactions predicts later language development. Nineteen infants participated in a declarative pointing task involving three different social conditions: two experimental conditions (a) available, when the adult was visually attending to the infant but did not attend to the object of reference jointly with the child, and (b) unavailable, when the adult was not visually attending to either the infant or the object; and (c) a baseline condition, when the adult jointly engaged with the infant's object of reference. At 12 months of age measures related to infants' speech-only productions, pointing-only gestures, and synchronous pointing-speech combinations were obtained in each of the three social conditions. Each child's lexical and grammatical output was assessed at 18 months of age through parental report. Results revealed a significant interaction between social condition and type of communicative production. Specifically, only synchronous pointing-speech combinations increased in frequency during the available condition compared to baseline, while no differences were found for speech-only and pointing-only productions. Moreover, synchronous pointing-speech combinations in the available condition at 12 months positively correlated with lexical and grammatical development at 18 months of age. The ability to selectively use this multimodal communicative strategy to engage the adult in joint attention by drawing his attention towards an unseen event or object reveals twelve-month-olds' clear understanding of referential cues that are relevant for language development. This strategy to successfully initiate and maintain joint attention is related to language development, as it increases learning opportunities from social interactions.

#### 4. References

- Austin, E. E. & Sweller, N. (2014). Presentation and production: The role of gesture in spatial communication. *Journal of Experimental Child Psychology*, 122, 92-103.
- Cartmill, E. A., Demir, O. E., & Goldin-Meadow, S. (2012). Studying gesture. In Hoff, E. (Ed.), *Research methods in child language: a practical guide*. Blackwell Publishing.
- Clark, J. M. & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review*, 3(3), 149-210.
- Cole, J. (2015). Prosody in context : a review Prosody in context : a review. *Language, Cognition and Neuroscience*, 30, 1-30.



- Dohen, M. (2009). Speech through the ear, the eye, the mouth and the hand. In Esposito, A., Hussain, A., & Marinaro, M. (Eds.), *Multimodal signals: Cognitive and algorithmic issues* (pp. 24-39). Berlin/Heidelberg: Springer.
- Esteve-Gibert, N., Prieto, P., & Pons, F. (2015). Nine-month-old infants are sensitive to the temporal alignment of prosodic and gesture prominences. *Infant Behavior & Development, 38*, 126–9.
- Guellaï, B., Langus, A., & Nespors, M. (2014). Prosody in the hands of the speaker. In *Language by mouth and by hand* (Vol. 5, p. 700). Frontiers in Psychology.
- Gullberg, M., deBot, K., & Volterra, V. (2008). Gestures and some key issues in the study of language development. *Gesture, 8*(2), 149–179.
- Igalada, A., Esteve-Gibert, N., & Prieto, P. (2017). Beat gestures improve word recall in 3- to 5-year-old children. *Journal of Experimental Child Psychology, 156*, 99-112.
- Igalada, A., Bosch, L., & Prieto, P. (2015). Language development at 18 months is related to multimodal communicative strategies at 12 months. *Infant Behavior and Development, 39*, 42-52.
- Jannedy, S. & Mendoza-Denton, N. (2005). Structuring information through gesture and intonation. *Interdisciplinary Studies on Information Structure, 3*, 199–244.
- Kelly, S. D., Ozyürek, A., & Maris, E. (2010). Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. *Psychological Science, 21*(2), 260–7.
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–227). The Hague, The Netherlands: Walter de Gruyter.
- Kendon, A. (1995). Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of pragmatics, 23*(3), 247-279.
- Krahmer, E. & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language, 57*(3), 396–414.
- Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge and New York, NY: Cambridge University Press.
- Leonard, T. & Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes, 26*, 1295-1309.
- Lewkowicz, D. J. & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences, 109*(5), 1431-1436.
- Liszkowski, U., Albrecht, K., Carpenter, K., & Tomasello, M. (2008). Infants' visual and auditory communication when a partner is or is not visually attending. *Infant, Behavior & Development, 31* (2), 157-167.
- Loehr, D. P. (2012). Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology, 3*, 71–89.
- McNeill, D. (2012). *How language began: Gesture and speech in human evolution*. Cambridge University Press.
- McNeill, D. (2005). *Gesture and Thought*. Chicago: University of Chicago Press.
- McNeill, D. (1992). *Hand and Mind*. Chicago. The Chicago University Press.
- Miller, J.L. & Gros-Louis, J. (2013) Socially guided attention influences infants' communicative behavior. *Infant Behavior & Development, 36* (4), 627-634.

- Miller, J. L. & Lossia, A. K. (2013). Prelinguistic infants' communicative system: Role of caregiver social feedback. *First Language*, 33 (5), 433–448.
- Moubayed, S.A., Beskow, J., & Granström, B. (2010). Auditory visual prominence. *Journal of Multimodal User Interfaces*, 3, 299-309.
- Murillo, E. & Belinchón, M. (2012). Gestural-vocal coordination. Longitudinal changes and predictive value on early lexical development. *Gesture*, 12 (1), 1, 16-39.
- Pierrehumbert, J. & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press.
- Pons, F., Andreu, L., Sanz-Torrent, M., Buil-Legaz, L., & Lewkowicz, D. J. (2013). Perception of audio-visual speech synchrony in Spanish-speaking children with and without specific language impairment. *Journal of child language*, 40(03), 687-700.
- Shattuck-Hufnagel, S., Ren, A., Mathew, M., Yen, I., & Demuth, K. (2016). Non-referential gestures in adult and child speech: Are they prosodic? *Proceedings of Speech Prosody*. Boston, USA, May 31-June 3.
- Terken, J. (1991). Fundamental frequency and perceived prominence. *Journal of the Acoustical Society of America* 89(4), 1768–1776.
- Treffner, P., Peter, M., & Kleidon, M. (2008). Gestures and phases: The dynamics of speech-hand communication. *Ecological Psychology*, 20(1), 32-64.
- Yasinnik, Y., Renwick, M., & Shattuck-Hufnagel, S. (2004). The timing of speech-accompanying gestures with respect to prosody. *Proceedings of From Sound to Sense* (pp. 97–102). Cambridge, MA: Massachusetts Institute of Technology.
- Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, 45(3), 598-607.
- Wagner, P., Malisz, Z., & Kopp, S. (2014). Gesture and speech in interaction: An overview. *Speech Communication*, 57, 209–232.
- Wu, Z. & Gros-Louis, J. (2014). Infants' prelinguistic communicative acts and maternal responses: Relations to linguistic development. *First Language*, 34 (1), 72-90.

## ADITZ JOKATUAN NON DITU ZAILTASUNAK EUSKARA H2 IKASTUN BERANTIARRAK?

Juan Abasolo  
UPV/EHU, EUDIA ikerketa-taldea  
juan.abasolo@ehu.eus

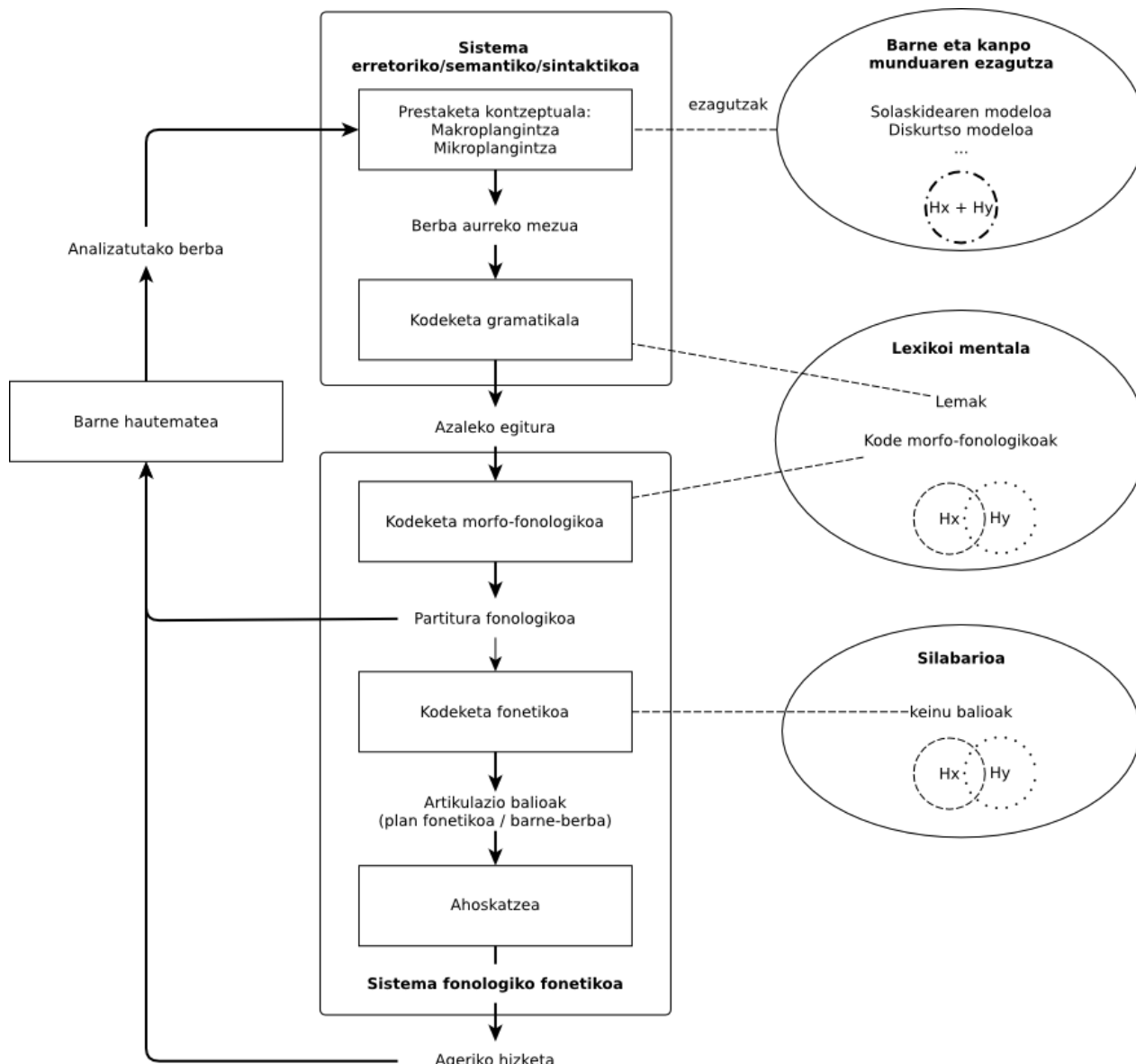
**Laburpena:** Artikulu honetan aurkezten da euskara-ikastun baten berbazko ekoizpenaren analisi bat; fokua adizkiaren inguruan neurtutako zailtasunetan kokatzen da, ea aditz nagusien, adizki laguntzaileen ala zeren inguruan multzokatzen diren hauteman nahirik. Horretarako, zailtasunen banaketan hondarrak  $\chi^2$  banaketarekin alderatu dira. Ikastun honen kasuan, zailtasunak adizki trinko jokatu edo adizki laguntzaileen inguruan multzokatzen direla antzeman dugu.

### 1. Sarrera

Lan honetan, adizki jokatuaren elementuen sailkapen bat erakusten da, zeinetan adizki nagusi perifrastikoak elementu erraztat hautematen diren, eta adizki trinkoek zein laguntzaileek, dirudienez, beste talde bat osatzen duten: adizkiaren alderdi zaila.

Aurkezten ditugun datuak ikerketa handiago batean ulertu behar dira, euskara-ikastun batek ahozko ekoizpenean erakutsitako zailtasunen analisisian (Abasolo, 2017), hain zuzen. Lan horretan, eta beste batzuetan (Abasolo, 2015a, 2016), erakutsia da antzerako portaera erakusten dutela zailtasuntzat har litezkeen errore, fonazio luzatu, errepikapen eta birformulazio zenbaitek.

Segidan, Segalowitz-ek bigarren hizkuntzan (H2) egiten diren etorri-etenen azterketarako aurkeztu zuen hizketaren modeloa aurkeztuko dugu (Segalowitz, 2010); berez, *Speaking* liburuan Levelt-ek aurkeztutako bigarren modeloa (Levelt, 1989), eta de Bot-ek Levelt-en lehenengo modeloari iradoki zizkion H2rako egokitzapenekin eraikia (De Bot, 1992).



1. irudia : Levelten hizketa-eredua H2 ere kontuan hartuta.

Modelo horren arabera, etorri-etenen eragileek asko eta toki askotan lukete eragina karratuetan aipatzen diren azpiprozesuetako bakoitzean zein erdiko biribilaz irudikatzen den lexikoi mentaleko elementuak berreskuratzeko unean ere. Denera, zazpi une identifikatu zituen Segalowitzek etorri-etenak edo desjariakortasunak eragiteko. Gure interes berezikoak kodeketa gramatikalaren unea eta lexikoi mentaletik berreskuratzekoa dira, une horietan kokatzen direlako aztergai dugun Abasolo 2017ko corpuseko zenbait zalantza eta bestelako zailtasun-adierazleak, adizki jokatuaren inguruan multzokatzen direnak, alegia.

Zalantzazkoa da ea aditza jokatzearen ondorio den ala adizki jokatu zehatz bat berreskuratzeko eragiten dituen zalantza-adierazleak eta bestelako zailtasun-adierazleak diren. Kormosek (Kormos, 2006) dioenez, bigarren hizkuntza batean ari den ikastunak

kode gramatikalik ez du eraikitzen, baizik eta elementuen formaren arabera aukeratzen ditu lexikoi mentalean (1. irudian erdiko biribilean irudikatzen da).

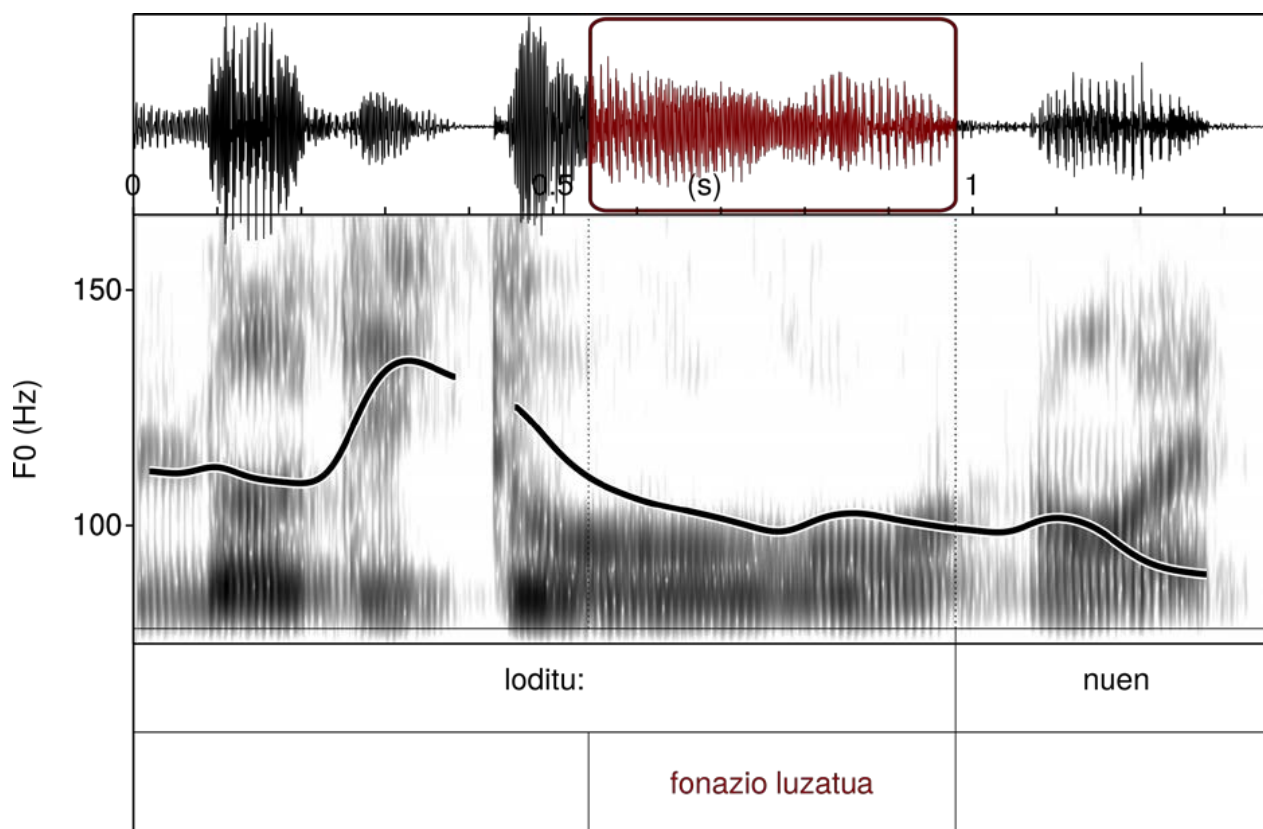
Hizketa ekoizteko uneko ezaugarri dira zalantzak eta autozuzenketak; H1 zein H2 izan. Errakuntzak H1ean eta H2an gertatzen badira ere, H2ko gaitasun baxuko ikasleen ezaugarria dela ere esan ohi da. Ikerketa honetan darabilgun corpuseko datuetan, zalantzak, autozuzenketak eta erroreak aditz jokatuaren inguruan kokatzen dira (Abasolo, 2017).

Zailtasun-adierazleen inguruko iritzi-batasunik ez dagoela esan liteke. Guk, lan honetan, zalantza adierazleetatik gramatika erroreetaraino dirauen kontinuuma, (Gaëtanelle & De Cock, 2011) lanean iradokitzen eta (Abasolo, 2017) lanean sailkatu den moduan:

Zailtasunen kontinuuma:

- Ahoskatu gabeak
  - Zalantza isilak
- Ahoskatuak
  - Zalantza ahostunak
    - Fonazio luzatuak
    - Silabaz silabako ahoskera
    - Errepikapenak
  - Erroreak
    - Hautemandakoak
      - Zuzendutakoak
      - Errepikapenak
      - Berba zatiak
      - Birformulazioak
    - Zuzendu gabekoak
  - Hauteman gabekoak

2. irudian ikusten den zalantza-adierazlea fonazio luzatua da, lan honetan zehatzen ikertu dugun adierazlea; aurreko zerrendako hasierako elementuetako bat da fonazio luzatua. Zalantza-adierazle hori adizki jokatuak kokatzen da, aditz nagusiaren ahoskatze-unean, hain zuzen; adizki laguntzailearen aurre-aurrean.



2. irudia: fonazio luzatuaren irudi bat PRAAT erabilia

Euskarazko aditz jokatuak zenbait ikertzaile liluratu du; ez dago ingelesezko<sup>1</sup> edo gaztelaniazko<sup>2</sup> Wikipediako artikuluen hasierak irakurtzea baino. Argiro azaltzen du Moreno Cabrerak falazia batez ari garela hizkuntzaren konplexutasun gramatikalaz ari garenean (Moreno Cabrera, 2007), oposizioak markatzen baitu aldearen tamaina — oposizioa egoteak edo ez egoteak (Newmeyer & Preston, 2014)— eta ez forma kopuruak.

Beste alde batetik, euskarazko adizki jokatuak hainbat informazio barnebil dezake, H2an trebatzen ari den batentzat zailtasun eragile izan daitekeena.

- (1) Etor zitezkeen
- (2) Badago

Aurreko adibideetan, konfigurazio bana ikus dezakegu. Lehenengo kasuan, adizki perifrastikoaren adizki nagusiak badu informazio semantikoa, eta bigarren elementuak, berriz, hots, adizki laguntzaileak, ia erabat erantzia du informazio semantikoa, aditz nagusirik gabe ulertu nahi bada; baina, era berean, bigarren elementu hori ekoizteko unean

1 [https://en.wikipedia.org/wiki/Basque\\_verbs](https://en.wikipedia.org/wiki/Basque_verbs)

2 [https://es.wikipedia.org/wiki/Verbo\\_vasco](https://es.wikipedia.org/wiki/Verbo_vasco)

informazio asko kodetu edo askoren artean aukeratu behar da, baldin eta ikastunak horixe argi adierazi behar badu euskaraz.

Bigarren adibidean, berriz, batera ikusten dira adizkiaren informazio semantikoa, *egotea*, eta horri dagozkion modu, pertsona eta zenbaki informazioa.

EGA azterketako idatziaren analisisian Pello Esnalek oso errakuntza gutxi aurkitu zituen 1988ko bere tesinan, eta honela azaltzen zuen:

aditzaren arloa dugu euskal hizkuntz formen artean egituratuenerikoa (Esnal, 1988)

Esnalek egindako ikerketa-lana idatzitakoaren analisisian oinarritu zen, euskara H1 ala H2 zen sailkatu ere egin barik. Idatzitakoaren analisiak, zailtasunen continuumean, errakuntzen berri baino ezin lezake eman, zalantza eta birformulatze guztiak amaituak baitira idatzia amaitutakoan. Horra ahozko jardunaren analisiak duen azalpen ahalmen handiagoa.

### 1.1. Corpusaren ezaugarriak

Gorago esan bezala, lan honetarako, Abasolo 2017 laneko datuak erabili dira. Datuak urtebetean zehar batu ziren, 2014-2015 ikasturtean. Horretarako, ikastunari grabatu zitzaizkion zenbait ahozko ariketa, aurrez prestatutako monologia egitean zeutzanak.

Ikerketa horretarako, propio garatutako transkribatze- eta etiketatze-sistema erabili zen (Abasolo, 2015b). Kodeketa horretan grafema bakarreko edo biko etiketak erabili ziren zailtasun-adierazle bakoitza identifikatzeko, baita zenbait ezaugarri morfosintaktiko nabarmentzeko ere. Bigarren pauso batean, etiketen multzokatze motak nola banatzen diren ikusteko bidea ematen du kodetutako testuaren azterketak.

|          |                          |               |                     |
|----------|--------------------------|---------------|---------------------|
| esaldia: | @æ jakiteko betebeharrak | \$ betetze:n  | ¡ duten *2          |
| etiketak |                          | luzatua0      | luzatua1            |
|          |                          |               | errorea (eza)       |
|          | aditz argumentua ABS     | aditz nagusia | adizki laguntzailea |

**1. taula:** testu eta etiketen adibidea

Corpusaren ezaugarri kuantitatiboak aztertzeko, eurak laburtuta aurkezten dira hurrengo taulan.

|   |                              |
|---|------------------------------|
| 3057 berba  | 1325 zailtasun               |
| 14 monologo   | 564 fonazio luzatu           |
| 3057 berba  | 38 berba silabatze markatuaz |
| 467 adizki  | 275 osatu gabeko berba       |
| 304 adizki nagusi perifrastiko / 278 adizki laguntzaile | 66 errepikapen               |
| 189 adizki trinko                                       | 144 birformulazio            |
| 407 aditz argumentu                                     | 238 errore                   |
| 64 Ezezko partikula                                     |                              |

**2. taula:** aztertutako corpusaren ezaugarriak.

Datuen ezaugarri morfosintaktikoen arabera aztertu dira zailtasunen banaketak; gehienetan,  $\chi^2$ -aren banaketarekin alderatuta. Horrela, antzeman dugu zenbait ezaugarrik badutela aditzen inguruko banaketa bereizia.

| Zailtasunen menpeko banaketak ( $\chi^2$ )        | ez  | aditz nagusia | trinkoa | laguntzailea |
|---|-----|---------------|---------|--------------|
| Fonazio luzatua                                   | **  | ***           | ***     | ***          |
| Fonazio luzatua aurrekoan                         |     |               | ***     | ***          |
| Fonazio luzatua azken-aurrekoan                   |     |               |         | *            |
| Berba zatiak ahoskatu dira aurretik               | **  |               | ***     | *            |
| Aurrekoaren errepikapena                          | *** | ***           | **      | ***          |
| Errepikapena hasi eta geroko lehenengo berba da   |     |               | *       | **           |
| Birformulazioa bertan hasten da                   | **  |               |         |              |
| Birformulazioa hasi eta geroko lehenengo berba da |     |               | **      | ***          |
| Zerbait falta du                                  |     |               |         | ***          |

**3. taula:** zailtasunen eta ezaugarri morfosintaktikoen banaketaren adierazgarritasun mailak,  $\chi^2$  banaketaren arabera.

Aurreko taulan ikusten denez, corpus honetan adizki nagusiaren zein adizki laguntzaileen inguruan agertzen dira multzokatuta zailtasun-adierazleak. Corpus horretako datuen beste analisi bat proposatuko dugu lan honetan.

## 2. Hipotesiak

Gure hipotesi nagusia da aditz nagusiaren inguruan aurkitu diren zailtasunak ez dagozkiola aditz nagusiari berari, baizik eta adizki laguntzaileei. Eta hori hizketaren ekoizpenaren ezaugarri inkrementalak azal lezake.



H<sub>1</sub>: Zalantza-adierazleen multzokatzeak adizki nagusietatik independenteak izango dira ezezko esaldietan.

H<sub>2</sub>: Bat datoz zalantza-adierazleen banaketak adizki laguntzaile zein adizki trinkoen kasuetan.

H<sub>3</sub>: Aditz nagusien aspektuetatik independentea da zalantza adierazleen banaketa

### 3. Metodologia

Zailtasunen agerrera euskarazko adizkien ezaugarriren batek ala zerk egiten duen zuzenean ezin azter daiteke corpusaren araketa hutsez; bai, ordea, hobeto ezagutu horien elkarreraginak. Horixe aztertzea helburu, hurrengo lerroetan aurkezten dugun lanerako ikusizko azterketa-teknikak erabili ditugu, *Visualizing Categorical Data* (VCD), banaketen ezaugarrien ulerkuntzan sakontzeko (Friendly, 1994). R estatistika-lengoiarako garatutako teknikak eta baliabideak erabili ditugu, vcd paketearekin (Meyer, Zeileis, & Hornik, 2006).

VCD teknikak tamainak banaketen dimentsioak irudikatzeko erabiltzen ditu; koloreak, konparaketa adierazgarriak erakusteko, eta tonalitateak, aldearen indarra irudikatzeko. Datuen banaketa mosaiko-grafiko batez irudikatzen du proportzioen berri emateko. Gure kasuan,  $\chi^2$ -aren banaketarekiko alderaketa kolorez irudikatzen du. Kolore gorritz, banaketa idealean litzatekeen baino maiztasun gutxiago duten kategoriak erakusten ditu, eta kolore urdinez, alderatutakoaren gainerako maiztasuna dutenak. Adierazgarritasunik erakusten ez duten multzokatzeetan, zuria erabili dugu lanean, eta grisa, banaketaren

osotasunean  $\chi^2$ -aren banaketarekiko alderik erakusten ez denean. Koloreon intentsitatea erabiltzen da aldea handiagoa ala txikiagoa den erakusteko.

Aukera-arrazoia erabili ohi da 2x2 dimentsioko alderaketak egiteko; guk, lan honetarako, horren erabilera ez baizik  $\chi^2$ -arena lehenetsi dugu; arrazoi bigatik: batetik, aukera-arrazoien indargunea banaketaren aurretiko ezagutzan oinarritzen delako eta corpuseko datuak ausaz hartu direlako, eta, betetik, aurrez egindako alderaketekiko koherentzia gordetzeko, 2x2 taulok beste batzuen laburpenak direlako.

### 4. Emaitzak

Hiru lan-hipotesiei dagozkien irudiak eta horien interpretazioak erakusten dira segidan, hurrenez hurren.

#### 4.1. Zalantza-adierazleak eta adizki nagusiak independenteak dira ezezko esaldietan

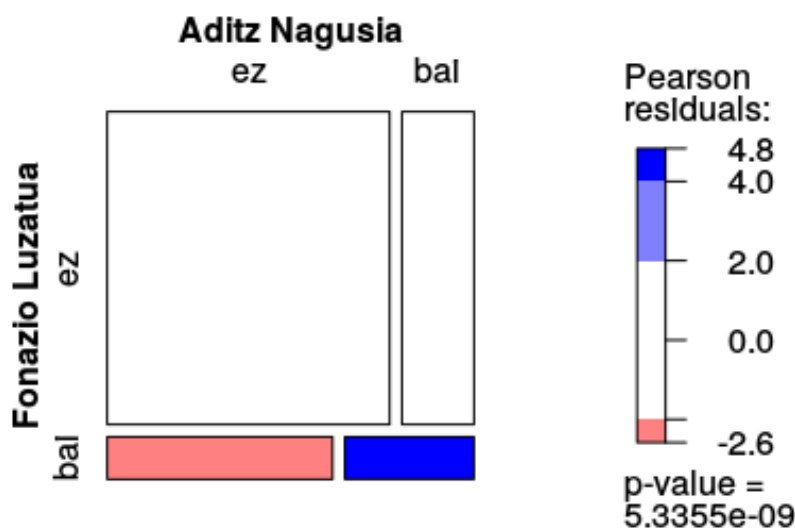
Alderaketa hau egiteko, fonazioak luzatzeko unek eta berben balio sintaktikoa alderatu dira; zehazki, ea adizki nagusia ala beste zerbait den fonazioa luzatzeko unean ahoskatua. Kasu honetan, konfigurazio bitan alderatu behar dira elementuak: baiezko eta ezezko esaldietan.

Lehenengo konfigurazioan, hau da, baiezko esaldietan, banaketak  $\chi^2(1, N=2157) = 33,138$ ,  $p = 0$  balioa du. Balio horrek zelako datu-banaketa duen azpian, vcl teknikak erabiliz aztertu dugu.

3. irudiko ezkerreko eskeman ikusten da aditz nagusian fonazioa luzatzeko joera handia dela baiezko esaldietan; izan ere, konfigurazio horretan etiketatutako fonazio luzatuen herenak aditz nagusian neurtu dira. Kolore gorri argiak adierazten du fonazio luzatuak aukera baxuagoa duela aditz nagusia ez diren elementuetan agertzeko.

Ezezko esaldietako adizki nagusietan, berriz, ez da aurkitu banaketa baldintzaturik khi karratuaren probaren bitartez:  $\chi^2(1, N=413) = 0$ ,  $p = 1$ .

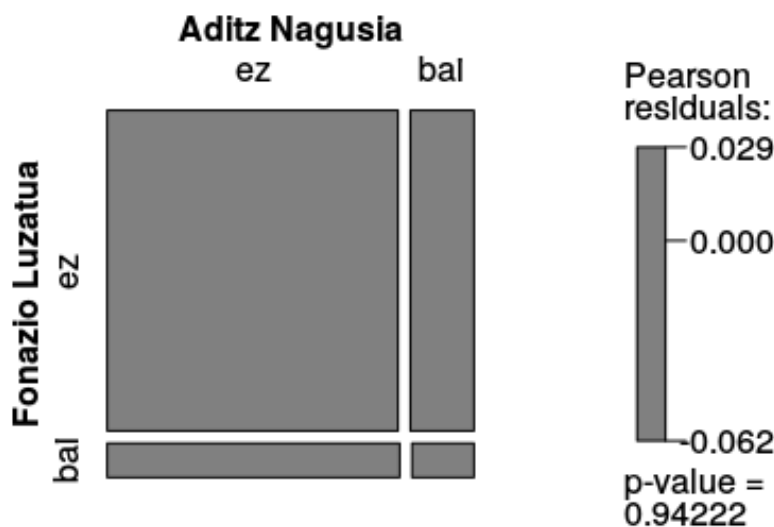
Hala ere, irudi-azterketa batek lagun lezake datu-banaketen moduak aztertzen, adierazgarritasunaren aurretik ere.



### 3. irudia: fonazio luzatuak eta aditz nagusien banaketa baiezko esaldietan

Aurreko ikerketekin bat etorrira, 3. irudian, baiezko esaldietako joera ikusten dugu: fonazioa luzatzeko joera handiagoa da adizki perifrastikoetan aditz nagusia ahoskatzeko

orduan beste kasu guztietan baino; era berean, kolore gorri argiak adierazten digu neurri adierazgarrian gutxiago ikusten ditugula fonazio luzaturik gabeko aditz nagusiak fonazio luzatudunak baino.



**Irudia 4:** fonazio luzatuak eta aditz nagusien banaketa ezezko esaldietan

Goiko 4. irudian ikusten den banaketa ez-esanguratsuak, nabarmentzekoa da, baiezko esaldietakoaren kontrakoa da. Hau da, zenbakietan ez da aurkitu baiezko esaldietako banaketaren antzik, adizki nagusian neurtutako fonazio luzatuei dagokienez.

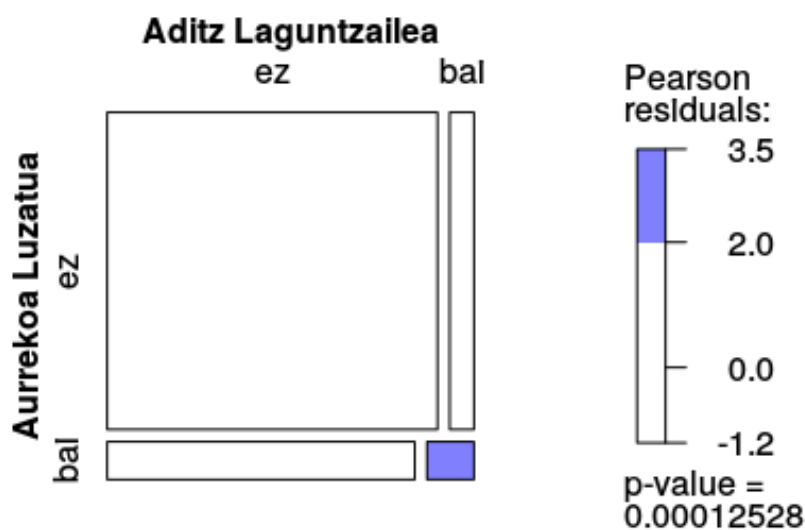
#### 4.2. Fonazio luzatuaren banaketak adizki laguntzaile eta trinkoen kasuetan.

Atal honetan ez da aztertzen ea adizkia ahoskatzeko unean fonazioa luzatu den ala ez, baizik eta ahoskatu aurretiko berban fonaziorik luzatu denetz, 2. irudian bezala.

**Adizki laguntzaileen** kasuan, fonazio luzatuaren banaketa ez da independentea, espero zitekeen bezala. Honako balioa hartzen du:  $\chi^2(1, N=2615) = 13,805$ ,  $p = 2,0275026 \times 10^{-4}$ . Honetan  $\chi^2$ -ak hartzen duen balioaren norabidea eta tamainak aztertzeko, mosaiko-grafikoa eraiki da; 5. irudiak erakusten duenez, koadrante bakar batean aurkitu da emaitza adierazgarria; horrek erakusten du banaketa ez dela independentea. Hain zuzen, aditz

laguntzaileen kasuan, aurreko hitzean fonazio luzatua aurkitzeko aukerak handiagoak dira aztergai darabilgun banaketa estatistikoaren arabera espero litekeen baino.

**Adizki trinkoen** inguruko azterketan ere mendekotasuna aurkitu dugu, honelako balioarekin:  $\chi^2 (1, N=2615)=14,399$ ,  $p=1,5 \times 10^{-4}$ . Horrek adierazten du adizki trinkoen aurreko hitzean fonazio luzatua aurkitzeko aukerak eta adizki jokatu trinkoak harremanetan daudela. Norabideak eta tamainak aztertzeko mosaiko-grafikoa erabili da.



5. irudia: fonazio luzatua aurrekoan eta adizki laguntzaileak

Goiko 5. irudian eta beheko 6.ak pareko banaketa erakusten dute, portaera bera edo bertsua.



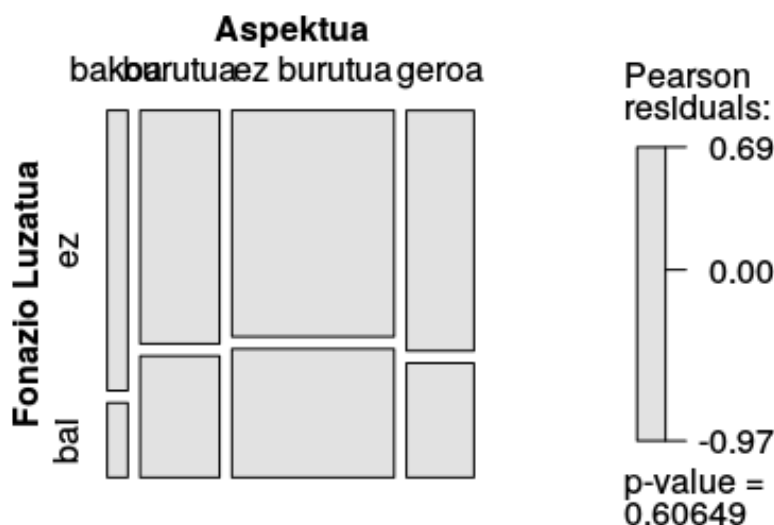
**6. irudia:** fonazio luzatua aurrekoan eta adizki trinkoak

Azterketak erakusten du fonazio luzatua aurreko hitzean aurkitzeko aukerak handiagoak direla adizki trinkoen kasuan bestelako kasuetan baino, adizki laguntzaileekin gertatzen den antzera.

*4.3. Aditz nagusiaren aspektua eta fonazio luzatua*

Aditz nagusiari eragiten dion gramatika-marka dugu aspektua. Aditz-aspektua zein den, baliteke fonazio luzatuaren agerrera menpeko banaketa izatea. Banaketa horren emaitzen arabera ( $\chi^2(3, N=303)=1,839, p=0,60649$ ), ez dago horrelakorik esaterik.

Era grafikoan ere azter daiteke banaketa hori, hipotesi berri edo bestelako informazioaren bila.



7. irudia: fonazio luzatua aditz nagusietan, aspektuen arabera.

Begizko azterketa batean ikus dezakegu fonazio luzatuaren gertaldiak gutxiago direla aspektu bako aditz nagusietan (maiztasunik baxuena dutenak) aspektu ez-burutuan baino (maizen ikusi den aspektua). Aspektu burutu eta geroan, berriz, pareko banaketa du ezaugarriak.

## 5. Ondorioak

Aurkeztu diren emaitzen argitan planteatu diren hiru lan-hipotesiak bete dira. Aditz nagusian aurkitu diren fonazio luzatuen mendeko banaketa baiezko esaldietan aditz nagusiak aditz laguntzailearekiko hartu ohi duen kokapenak baldintzatua dela ondoriozta daiteke. Ezezko esaldietan ikusten den portaera erabat bestelakoak argi uzten du funtzio sintaktiko horretatik independentea dela ahoskatze-ezaugarri hori.

Antzerako jokaera ikusi dugu fonazio luzatuaren agerrera adizki laguntzaile edo trinkoen presentziaren aurrean. Horrek pentsarazten du funtzio desberdin bi horiek osotasun bat dutela corpuseko informatzailearen prozesamenduan.

Aditz nagusi perifrastikoaren funtzio sintaktikoan gramatika elementu bakarra agertzen da aztergai: aspektu-marka edo markarik eza. Horren arabera analisiak ez digu erakutsi aspektu-markaren arabera oposizioak ezelako loturarik duen zailtasunen agerpenarekin.

Ez dago argi, corpusaren analisi huts honetan aztertuta, ea zerk eragiten duen zailtasuna, aditza jokatzek ala aditz jokatu formak berreskuratzeak. Lan honetan zailtasun-eragilea identifikatzera hurreratu gara, eta euskara H2ko ikastun berantiarraren honen prozesamendua ulertzera ere hurreratu gara. Zer zehaztu ere definitu ahal dugu: ea aditza jokatzek ala aditz-formak berreskuratzea den zailtasun-eragilea, ala bien arteko zelanbaiteko elkarreragina.

Hurrengo lan batzuetan formen maiztasuna eta zailtasunak alderatu daitezke, ea formen zailtasuna maiztasunarekin lotuta dagoen aztertzeko ala ea aditz inflexioaren presentziak; beraz, funtzio sintaktiko horrek azaltzen duen ikusteko. Corpora zabaltzek ere emaitza orokorgarriagoak ekar litzake; beraz, corpora beste hiztun batzuetara zabaltzea ere gomendatzen dugu.

## 6. Erreferentziak

- Abasolo, J. (2015a). Adizki jokatuaren sailkapen hirukoitza: Zailtasunak, etenak eta erroreak. In *Nuevos retos en la investigación en Psicodidáctica - Psicodidaktikako ikerkuntzaren erronka berriak* (6–20. pp. ). Leioa: Argitarapen Zerbitzua, UPV/EHU.
- Abasolo, J. (2015b). Transkribatze arauak eta transkripzio helburuak. In *Investigar en psicodidáctica: Una realidad en auge - Psicodidaktikako ikerketa gorabidean* (316–325. pp. ). Leioa: Argitarapen Zerbitzua, UPV/EHU.
- Abasolo, J. (2016). Desjariakortasuna eta gramatika errorea adizki jokatuaren B2-C1 ibilbidean: Kasu baten analisia. In *XXIII Jornadas de Investigación en psicodidáctica - XIII Psicodidaktikako ikerkuntza jardunaldiak* (345–358. pp. ). Leioa: Argitarapen Zerbitzua, UPV/EHU.
- Abasolo, J. (2017). *Zailtasunak aditz jokatuaren inguruan B2-C1 ibilbidean* (Doktorego-tesia). UPV/EHU, Leioa. Berreskuratua <https://addi.ehu.es/handle/10810/25306>-(e)tik
- De Bot, K. (1992). A bilingual production model: Levelt's 'Speaking' model adapted. *Applied Linguistics*, 13(1), 1–24.
- Esnal, P. (1988). Erroreak helduen euskalduntzearen azken urratsetan. *Zutabe: Euskalduntze Alfabetatzearen Aldizkaria*, 17. zenb, 9–66 or.
- Friendly, M. (1994). Mosaic Displays for Multi-Way Contingency Tables. *Journal of the American Statistical Association*, 89(425), 190.
- Gaëtanelle, G., & De Cock, S. (2011). Errors and disfluencies in spoken corpora: Setting the scene. *International Journal of Corpus Linguistics*, 16(2), 141–172.
- Kormos, J. (2006). *Speech Production and Second Language Acquisition*. London: Lawrence Erlbaum Associates.

- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: The MIT Press.
- Meyer, D., Zeileis, A., & Hornik, K. (2006). The Strucplot Framework: Visualizing Multi-way Contingency Tables with vcd. *Journal of Statistical Software*, 17(3).
- Moreno Cabrera, J. C. (2007). Sobre la complejidad y dificultad de las lenguas: El caso del euskera. *Revista de Lenguas Y Literaturas Catalana, Gallega Y Vasca*, (13), 199–216.
- Newmeyer, F. J., & Preston, L. B. (2014). *Measuring Grammatical Complexity*. Oxford University Press.
- Segalowitz, N. (2010). *Cognitive Bases of Second Language Fluency*. New York: Routledge.



## EUSKALKIEN SAILKAPEN ZIENTIFIKOA: LEXIKOA (2)

Aurrekoetxea, G., Ormaetxea, J.L. & Videgain, X.

Eudia ikerketa-taldea

[gotzonaurre@gamil.com](mailto:gotzonaurre@gamil.com), [txipirrin@gmail.com](mailto:txipirrin@gmail.com), [charles.videgain@univ-pau.fr](mailto:charles.videgain@univ-pau.fr)

**Laburpena:** Ekarpene honen ohiko euskalkien sailkapen zientifikoa egiteko bidean bigarren ekarria aurkeztea du helburu. Horretarako, EHHA egitasmoan bildutako datuak erabili dira, zehazki lehen bi liburukietan biltzen den lexikoa. Horrek esan nahi du gaur egun 100 urte inguru dituen edo izango lituzkeen jendearen lexikoa duela langai. Guztira, 183 galdera aukeratu dira, eta datuak datu-basean ezarri, transkribaketa ortografikoa erabiliz. Ondoren, erantzunetan jasotako datuak lematizatu egin dira: hots, hitzetan agertzen diren ezaugarri fonetikoak ezabatu eta hitzek izan ditzaketen aldaeren artean bat hautatu da. Hau da berrikuntza ekarpen honetan: datuen analisia ez dela erantzunetan (datu gordinetan) oinarrituz egin, datuak lematizatu ostean sortu diren lehen gainean baizik.

Sortu den datu-basea *Diatech* programara igo da analisi estatistikoa eta mapagintza-lanak egiteko. Analisisian bi mota erabili dira: lehenik, lemetan oinarritutako emaitzak aurkeztu dira eta, ondoren, erantzunetan oinarritutako ikerketetan lortutako emaitzekin (Aurrekoetxea, Ormaetxea, Videgain 2018) alderatu dira. Erabili diren tekniken artean, similaritate-mapa esanguratsuenak aztertu dira, batetik; hots, herri batetik abiatuz gainerako herrietara dagoen desberdintasun lexikala nolakoa den zehaztu da, mapak eta hizkuntza-distantzien taula ere aztertuz. Bestetik, cluster analisiaren bidez egin den hizkeren sailkapen hierarkikoa erakutsiko da. Sailkapen hau bi grafikoz erakusten da: batetik, dendrogramaren edo zuhaitz hierarkikoaren bidez, zeinetan aztertzen diren 145 hizkerak hizkuntza-hurbiltasunaren hierarkia baten barnean agertzen diren, eta, bestetik, emaitzak mapa batean ezarri.

**Gako-hitzak:** bariazio geolinguistikoa, dialektometria, EHHA, lexikoa

### 1. Sarrera

Berriketan, Aurrekoetxea, Ormaetxea eta Videgainek (2018), lexikoan oinarrituz, euskalkien sailkapenaz ekarpen berria plazaratu dute, datu-baseak eta prozedura kuantitatiboak erabiliz. Argitalpen hau aurretik zetozen Aurrekoetxea (2016), Videgain & Aurrekoetxea (2016) eta Iglesias & Aurrekoetxea (2017) lanekin abiatu zen euskalkiei buruzko sailkapen berriaren ikerketa-lerroari dagokio. Ekarpene hau euskalkien banaketan eta sailkapenean beste urrats bat ematera dator.

Azken urteotan, poliki-poliki baina etengabe, euskararen kasuan ere jada gaintzen ari gara dialektologia tradizionalaren ekarriak (Bonaparte 1868, Zuazo 1998), bai metodologian, bai datu kopuru erraldoien erabileran eta bai sailkapena egiteko prozeduretan.

Izenburutik beretik abiatuz, “sailkapen zientifikoa” terminoa erabiltzeko zilegitasuna aldarrikatzen dugu. Horren zergatia gure aurreko ekarpenean azaldu bagenuen ere, labur-labur errepikatuko dugu hemen: sailkapena zientifikoa dela diogu, batetik, ikerketa

zientifikoetan, objektibotasun eta sinesgarritasunari begira, datu kopuru handiak erabili ohi direlako; azterketa honetan 26.535 erregistro erabili dira. Ez da, egia esan, datu-kopuru erraldoia, baina bai orain artean euskararen sailkapen geolinguistikoak egiteko erabili diren datu-kopuruak baino askoz zabalagoa. Bestetik, zientifikoa da ikerketa honetan ere zeinahi zientzian ezer sailkatu behar denean erabiltzen diren teknikak erabili direlako; hots, sailkapen hierarkikoa edo, bestela esanez, cluster analisia. Eta azkenik, horretarako erabili diren datu guztiak argitaratuak eta egiaztagarriak dira. Hori dela eta, aldarrikatzen dugu sailkapen honen zientifikotasuna, ikerketa hau ez delako oinarritzen elukubrazioetan, ez hizkuntza-ezaugarrien hautu subjektiboetan, ez ezaugarri gutxi batzuetan.

Hain zuzen ere, horrexek, datu-kopuru handiak eta sailkapen-teknikak erabiltzeak, ematen dio ekarpen honi gure ustez duen ezohiko garrantzia. Horrek ez du esan nahi dena eginda dagoenik, ezta hurrik eman ere, baina metodologia eta prozedura berekin burututako laugarren euskararen sailkapen honek bide egokian jartzen gaituela eta gure oinatzak norabide egokian ipini ditugula erakusten du. Zientziak ere unean uneko teknika eta teknologia berrietara egokitu behar izaten du, guk ere halaxe egiten dugu euskararen sailkapenari dagokion ekarpena, gogoan harturik J. Nerbonneek aspaldi zioena (2008: 7): “Good dialectological analysis need to adopt techniques of aggregation more extensively.”

Ondorioz, esan dezakegu hemen lan honetan ez dagoela ez elukubrazioerik, ez suposizioerik, ezta usteen gainean ezer eraikitakorik ere.

## **2. Metodologia**

EHHako datu-bilketaren metodologiaz ez dugu hemen jardungo; jada hainbat lekutan argitaratu dira ezaugarri horiek eta irakurleak eskura du, hala nahi izanik, horren berri zehatza (Aurrekoetxea 1986, 2002; Aurrekoetxea & Videgain 1987). EHHaren datu eta argitalpenetik haratagoko bidea erakutsiko da hemen.

### *2.1. EHHatik datu-base berrira*

EHHak bi formatutan argitaratu ditu datuak: datu gordinak alfabetikoki emanak zerrendatan, lurraldeka eta herrika antolaturik aurkezten ditu lehenik. Eta ondoren, datuon interpretazio geolinguistikoa eginez, mapen bidez erakusten zaio irakurleari. Ikerketa honetan abaiapuntua zerrendetan agertzen diren erantzun fonetikoak izan dira. EHHak erantzunen zerrendetan argitaratu dituen datu fonetikoak idazkera ortografikora bihurtu dira datu-base berria sortzean; zuzenean eraldatu dira datuak. Galderaz galdera eta herriz herri datuak banan-banan idatzi dira sortu den datu-base berrian.

Ikerketa honetan EHHako lehen eta bigarren liburukietako informazioa erabili da ikerketa honetan; horretarako, paperezko bertsioak izan dira iturri. Paperezko bertsioa esku artean, aurrez zehaztutako bi baldintza hauek betetzen zituzten galderak aukeratu dira:

a) Erantzunik gabeko herri gutxi edo bat ere ez zuten galderak hartu dira kontuan. Galderaren erantzunen zerrendan bost herri edo gehiagotan erantzunik ez denean izan, galdera hori ez da kontuan hartu.

b) Lexiko orokorra bakarrik aukeratu da, edozein gai dela, kontzeptu orokorrak biltzen dituzten galderak bakarrik onartu dira. Hori dela eta, adibidez, baztertu egin dira intsektuei, arrainei eta abarrei dagozkien galderak.

Bi baldintza hauek kontuan izanik eratu da datu-base berria. EHHako lehen bi liburukietan 509 galdera biltzen diren arren, ikerketa honetako datu-baserako 183 galdera bakarrik aukeratu dira; gainerako galdera guztiak, edo erantzun hutsak direlako edo galderetan lexiko berezitua lantzen delako, bazter utzi dira. Eraiki den datu-base berriak ondoko ezaugarriak biltzen ditu:

- Formatua: .csv formatuan eraiki da (*Openoffice*-eko *Calc* formatua).
- Unicode UTF-8an kodetu da datuak gordetzeko.
- 183 galdera ditu datu-baseak.
- 145 herri biltzen dira.
- Guztira 26.535 erregistro biltzen dira.

Lan hau guztia eskuz egin da, ekarpen hau izenpetzen dugun hiru kideon artean, egonaldi baten, bakoitzak galdera desberdinak landuz. Horretarako, aste osoa behar izan da, bost eguneko jarduera latz eta gogorra, zeinetan aldian-aldian zalantzak eta elkarrizketaren bidez argituak izan diren. Datuen eraldatze-lanean eta datu-base berrian, EHHAn agertzen diren erantzunak (lekukoek euren kasa emandako erantzunak) eta proposamen onartuak (inkestazaileek proposatu eta lekukoek herriko forma edo berbatzat onartuak) maila berean ezarri dira. Zilegizkoa da herri bat ordezkatzeko duen hiztun batek herrikotzat jotzen duen hitz bat herrikotzat hartzea, nahiz ez zaion inkestagunean bat-batean atera eta inkestagileak proposatua izan.

|      |    |    |                             |
|------|----|----|-----------------------------|
| 6236 | 44 | 1  | kopo                        |
| 6237 | 44 | 2  | alpar,matasa                |
| 6238 | 44 | 3  | matasa,txingor              |
| 6239 | 44 | 4  | oja,txotor                  |
| 6240 | 44 | 5  | metxa                       |
| 6241 | 44 | 6  | edurtxotor                  |
| 6242 | 44 | 7  | edurmata,txingor            |
| 6243 | 44 | 8  | edururtantakada,edurtxingor |
| 6244 | 44 | 9  | tantakada,matasa            |
| 6245 | 44 | 10 | edurtxistil,edurmako        |
| 6246 | 44 | 11 | matasa,edurlapatz           |
| 6247 | 44 | 12 | edurgarrafa,edurmatasa      |
| 6248 | 44 | 13 | edursati,edurlume           |
| 6249 | 44 | 14 | edurmatasa,edurtxingor      |

**1. taula:** datu-base berrian 44. galderari (EHHako 09090) dagokion datu-base zatia.

## 2.2. Datuen lematizazioa

Datu-base hau lematizatu egin da ondoren. Zeregin honetan jasotako hitzen aldaeren artean bakarra onartu da, herri guztietarako bera. Hiztegi sarrerak erabakitzeko egiten den lanaren antzekoa izan da lematizazio hau; hitz beraren aldaera desberdinak izanik, aldaera bakarra (gainerakoen ezaugarriak ondoen gordetzen dituen) aukeratu da. Baliteke sortu den lema berria forma asmatua izatea ere; hots, jaso diren aldaeren artean ez izatea, berreraiketa izatea.

Datu-base honen egitura honako hau da (1. taula): lehen zutabea erregistro zenbakia da; bigarrena galdera-zenbakia (kasu honetan 44. galderaren zenbakia da). Hirugarren zutabea, herriaren kodea agertzen da (“1” Getxo da, eta horrela jarraitzen du mendebaldetik ekialderanzko norabidean). Beraz, taula honetan agertzen diren erantzunak Bizkaiko mendebaldeko herrietan jaso direnak dira.

## 2.3. Hizkeren arteko desberdintasunak neurtzeko sistema

Eraikitako datu-basea *Diatech* tresnara igo da. Hizkeren arteko analisi geolinguistiko kuantifikatuak (hizkeren arteko hizkuntza-distantzien neurketa eta sailkapenak) egiteko, oso egokia da online tresna hau: <http://eudia.ehu.eus/diatech/index/>. *Diatech* EUDIA ikerketa-taldeak bariazio geolinguistikoa ikertzeko sortutako tresna da. Tresna honek hizkeren arteko desberdintasunak estatistikoki aztertuz emaitzak mapetan ezartzen ditu.

Datuak igo ondoren, hizkeren arteko desberdintasunak neurtzeko erabiliko den izari edo neurria erabaki behar da ezer baino lehen. Dialektometrian neurri bat baino gehiago erabiltzen da horretarako: RIV neurria (*Relative Index Value*), Salzburgo-ko eskolak proposatua eta eskola horren jarraitzaileek erabiltzen dutena (ikus Goebel 1978 nola eraiki den jakiteko). Beste izari bat *Levenshtein algoritmoa* da, Groningen-go eskolakoek eta haren jarraitzaileek erabiltzen dutena (ikus besteak beste, Nerbonne & Heeringa 2010). Bi neurriak erabilgarri ditu *Diatech* programak analisi estatistikoak egin nahi direnerako. Neurriok desberdinak dira oso. Lehena, RIV delakoa, neurri kategorikoa edo nominala da, eta konparatzen diren hitzak bere osotasunean hartzen ditu: nahikoa da zeinu edo hizki bat desberdin izatea hitzak desberdintzat hartzeko (neurri horren arabera eta adibide bat jartzeko, *ixar* eta *izar* % 100ean desberdinak izango lirateke). Neurri hori egokia da lehen arteko edo tipologian oinarritutako hizkuntza distantzia neurtzeko, adibidez. Bigarren neurria, *Levenshtein algoritmoa*, ez da distantzia kategorikoa, eta hizkien galtze, aldatze edo sortzean oinarritzen da: hitzen arteko desberdintasuna bi hitzen arteko soinu-desberdintasunean datza (neurri honen arabera, *ixar* eta *izar* aldaeren arteko desberdintasuna % 25 da, lau hizki edo soinutatik bat dute desberdin); neurri hau oso da egokia distantzia fonetikoak edo lematizatu gabeko “datu brutuak” neurtzeko. Bata ala bestea erabiltzea ikertzaileari dagokio eta ikerketaren helburuen arabera izaten da (ikus zehatzago Aurrekoetxea 2016: 28-30). Honako ekarpen hau lemetan oinarritzen denez (ikus Aurrekoetxea, Ormaetxea & Videgain 2018 erantzun ortografikoetan oinarritutako azterketarako), distantziak neurtzeko RIV izaria erabiliko da. Aipatu neurriez gain, badira beste izari batzuk ere hizkuntza-datuak neurtzeko, baina ez dira hemen kontuan izango.

#### 2.4. Erabiliko diren estatistika-teknikak

Eremuen arabera bariazio geolinguistikoa ikertzeko garatu diren teknika kuantitatiboen artean, lehenik, hizkeren arteko hizkuntza-distantzien taula eta similaritate-mapa bat aurkeztuko dira. Izan ere, hizkeren arteko hizkuntza-distantziak neurtu eta hizkeren sailkapena egiteko, datu linguistikoaren arteko desberdintasunak zenbatu behar dira. Zenbatze honen ondorioz, hizkera batetik gainerakoetara dagoen desberdintasuna zenbakiz eman liteke; hori da, hain zuzen ere, hizkuntza-distantzien taula. Dialektologian ere jada, hizkeren arteko desberdintasunei buruz zehaztasunez eta zientifikoki hitz egitean, desberdintasun hori kuantifikatu behar da, ez da nahiko “desberdintasun handia edo txikia” dagoela esatea.

Ondoren, hizkeren sailkapenari begira, sailkapen hierarkikoa (cluster analisia) egingo da. Teknika hau, zientzietan datuak sailkatu behar direnean, ohikoa bihurtu da; teknika unibertsala da, zientzia guztietan erabiltzen dena. Azterketa honetan sailkapen honen emaitza bi modutan aurkeztuko dira: dendrogramaren (edo zuhaitz hierarkikoaren) bidez,

lehenik, eta mapen bidez, ondoren. Mapetan, gainera, balizko euskalkien eta azpieuskalkien mapak aurkezteko ahaleginak egingo dira.

### 3. Analisia

Aurrekoetxea, Ormaetxea & Videgain (2018) lanean hizkuntza-izaritzat *Levenshtein algoritmoa* (*Lv*) erabili zen, hizkuntza-datu gordinak (lematizatu gabeak) erabili zirelako. Oraingoan, aldiz, RIV erabili da; hain zuzen ere, datu lematizatuak erabili direlako. Atal honetan, lehenik, hizkuntza-distantzien taula aurkeztuko da, hizkera batetik bestera dagoen hizkuntza-desberdintasuna erakusteko; ondoren, taula hau eta *Lv* izariarekin lortutakoa erkatuko dira. Bigarrenik, similaritate-mapa bi: bata Getxori dagokio, ezker muturrean kokatuta dagoen herriari, eta bestea Eskiulari, maparen eskuin muturreko herriari. Hirugarrenik, eta azkenik, sailkapen hierarkikoari dagozkion dendrograma eta mapak azalduko dira; hots, lexikoari dagozkion sailkapen zientifikoa. Atal honetan, RIV izariarekin lortutako mapa *Lv* izariarekin lortutakoarekin erkatuko da, lematizatu gabeko datuekin (eta *Lv* izariarekin) eta lematizatutako datuekin (eta RIV erabiliz) lortutako mapak erkatuz.

#### 3.1. Hizkuntza-distantziak

Dialektometriak egindako aurrerakadaren ondorioz, egun hizkera batetik besterako hizkuntza-desberdintasunak neur litezke eta hizkuntza-distantzia zertan den zehaztu. Ikerketa honetan egindako eragiketa matematiko eta estatistikoak *Diatech* programaren bidez egin dira. Programa honek lortutako datu eta emaitza guztiak deskargatzeko aukera ematen du, eta hori da oraingoan egin den eragiketa (ikus 2. taula).

|              | Abaurreg | Ahetze | Aia   | Aldude | Alkotz | Altzai | Altzuruk | Amezka | Andoain | Aniz  |
|--------------|----------|--------|-------|--------|--------|--------|----------|--------|---------|-------|
| Abaurregaina | 0        | 57,69  | 61,2  | 54,95  | 50,28  | 69,4   | 69,95    | 61,75  | 64,48   | 46,99 |
| Ahetze       | 57,69    | 0      | 63,74 | 48,07  | 58,89  | 67,58  | 65,38    | 65,38  | 66,48   | 52,75 |
| Aia          | 61,2     | 63,74  | 0     | 63,74  | 53,59  | 74,86  | 74,32    | 42,08  | 47,54   | 53,55 |
| Aldude       | 54,95    | 48,07  | 63,74 | 0      | 55     | 65,38  | 65,38    | 65,38  | 68,68   | 49,45 |
| Alkotz       | 50,28    | 58,89  | 53,59 | 55     | 0      | 72,93  | 72,38    | 57,46  | 55,8    | 38,12 |
| Altzai       | 69,4     | 67,58  | 74,86 | 65,38  | 72,93  | 0      | 43,17    | 75,96  | 77,6    | 70,49 |
| Altzuruku    | 69,95    | 65,38  | 74,32 | 65,38  | 72,38  | 43,17  | 0        | 76,5   | 77,6    | 69,4  |
| Amezketak    | 61,75    | 65,38  | 42,08 | 65,38  | 57,46  | 75,96  | 76,5     | 0      | 40,98   | 55,74 |
| Andoain      | 64,48    | 66,48  | 47,54 | 68,68  | 55,8   | 77,6   | 77,6     | 40,98  | 0       | 56,28 |
| Aniz         | 46,99    | 52,75  | 53,55 | 49,45  | 38,12  | 70,49  | 69,4     | 55,74  | 56,28   | 0     |
| Aramaio      | 68,85    | 71,98  | 66,12 | 70,33  | 64,09  | 80,33  | 79,23    | 64,48  | 67,21   | 63,39 |

**2. taula:** hizkeren arteko distantzien taula (zatia) RIV izaria eta lema erabiliz

2. taulan ikus daitekeenez, Abaurregaineko hizkeratik Ahetzeko hizkerara dagoen berdintasuna % 57,69 da (eta beraz, hizkera bien arteko desberdintasuna % 42,31koa da). Abaurregainak hurbilen duen herria (taularen zati honetan) Aniz da (% 46,99) eta urrunen Altzürükü (% 69,95). Bestalde, taularen zati honek erakusten dituen zenbakietatik distantzia handienak Aramaio eta Altzaireren artean biltzen dira (% 80,33).

Taula osoan arakaturaz gero, hizkuntza-distantzia handienak Etxebarri eta Eskiularen artean jaso dira (% 83,52). Ondoren Etxebarri eta Altzai, eta Etxebarri eta Altzürüküren artean agertzen dira (% 83,06). Beti uste eta esan izan da euskararen dialektalizazioa handia dela; orain, datu hauekin, baieztatu baino ez da egin uste hori: euskararen mendebaldetik ekialderako hizkeren artean, eta lexikora mugatuta, % 83,52 desberdin izatea oso kopuru handia da. Hots, 100 hitzetatik 17 hitz baino ez dira herri bietan erabiltzen. Aldiz, distantzia txikienak, Baztanen agertzen dira: distantzia txikiena Aniz eta Lekarozen artean jaso da (% 26,92); ondoren, Erratzu eta Lekarozen artean (29,67), eta, ondoren, Aniz eta Erratzuren artean. Bizkaian, hizkuntza-distantzia txikiena Lemoiz eta Bakioren artean jaso da (31,15); Gipuzkoan, Asteasu eta Elduainen artean; Lapurdin, Ahetze eta Senpereren artean (34,62); Nafarroa Beherean, Landibarre eta Behorlegiren artean (35,36), eta Zuberoan, Sohüta eta Barkoxeren artean (38,80)<sup>3</sup>. Honek esan nahi du lexikoan hizkeren sare honetan ez dela ehunetik 26,92 baino hitz berdin gehiago dituen herririk (kontuan izan behar da datu-basean 183 kontzeptu erabili direla; beraz, Aniz eta Lekarozen artean 49 hitz desberdin jaso direla).

Erka ditzagun, orain, Lv eta RIV izariekin lortutako emaitzak (2. eta 3. taulak): batetik, eta gaingiroki begiraturaz, Lv izariarekin lortutako distantzia-emaitzak txikiagoak dira RIV-arekin lortutakoak baino (kasu bat baino ez aipatzearen, Abaurregaineko hizkeratik Ahetzeko hizkerara RIV izariarekin % 57,69 da eta Lv-arekin % 47,52).

|              | Abaurregaina | Ahetze | Aia   | Aldude | Alkotz | Altzai | Altzürükü | Amezketeta | Andoain | Aniz  |
|--------------|--------------|--------|-------|--------|--------|--------|-----------|------------|---------|-------|
| Abaurregaina | 0,00         | 47,52  | 51,73 | 44,87  | 46,73  | 61,12  | 62,93     | 50,70      | 53,56   | 41,09 |
| Ahetze       |              | 0,00   | 51,09 | 36,41  | 50,98  | 57,02  | 57,65     | 51,71      | 51,39   | 42,66 |
| Aia          |              |        | 0,00  | 53,93  | 50,13  | 66,02  | 66,18     | 37,06      | 39,46   | 47,84 |
| Aldude       |              |        |       | 0,00   | 51,57  | 56,14  | 58,13     | 52,63      | 54,33   | 42,38 |
| Alkotz       |              |        |       |        | 0,00   | 66,23  | 67,34     | 49,69      | 49,15   | 36,90 |
| Altzai       |              |        |       |        |        | 0,00   | 31,26     | 65,92      | 66,11   | 61,65 |
| Altzürükü    |              |        |       |        |        |        | 0,00      | 67,62      | 67,70   | 63,49 |
| Amezketeta   |              |        |       |        |        |        |           | 0,00       | 33,26   | 46,47 |
| Andoain      |              |        |       |        |        |        |           |            | 0,00    | 47,61 |
| Aniz         |              |        |       |        |        |        |           |            |         | 0,00  |

**3. taula:** hizkeren arteko erantzunen distantzien taula (zattia) Lv izaria erabiliz

Hizkeren arteko distantzia handienak erkatuz, RIV erabiliz, Etxebarri eta Eskiularen artean jaso da desberdintasun handiena (% 83,52); Lv izaria erabiliz, ordea, Altzürükü eta Getxoko hizkeren artean (% 73,26)<sup>4</sup>. Desberdintasun txikienari erreparatzen bazaio, RIV izariarekin, Aniz eta Lekarozen artean jaso da distantzia txikiena (% 26,92), eta, kasu

<sup>3</sup> Nahiz oso interesgarria den jakitea zein diren herri hauek bereizten dituzten ezaugarri lexikalak beste lan baterako utziko da, artikulu honen helburu eta luzeratik kanpo daudelako.

<sup>4</sup> Hemen ere irakurketa zehatzagoaren beharra antzematen dugu, itaunduz ea zergatik diren batean hizkera batzuk eta bestean beste batzuk distantzia handienak eta txikienak eragiten dituztenak.

honetan, *Lv* izariarekin ere bi hizkera horien artean jaso da distantzia txikiena (% 22,87). Aurreikusten zenez, *Lv* izariarekin hizkuntza-distantzia txikiagoak lortzen dira RIV distantziarekin baino.

Datuok eskuetan izanik, euskal dialektologiak ikaragarrizko jauzia egin duela iruditzen zaigu: metodologia eta prozedura egokiak erabiliz, zehaztasunik gabeko garaietatik hizkeren arteko distantzia modu zehatzean kuantifikatu da; eta gainera, ez modu bakarrean, distantzia unitate bi erabiliz. Honek euskal dialektologia inoiz aurkitu ez den egoera batean kokatzen du: hizkeren arteko desberdintasunen analisisian eta hizkeren sailkapenean.

### *3.2. Similaritate-mapak*

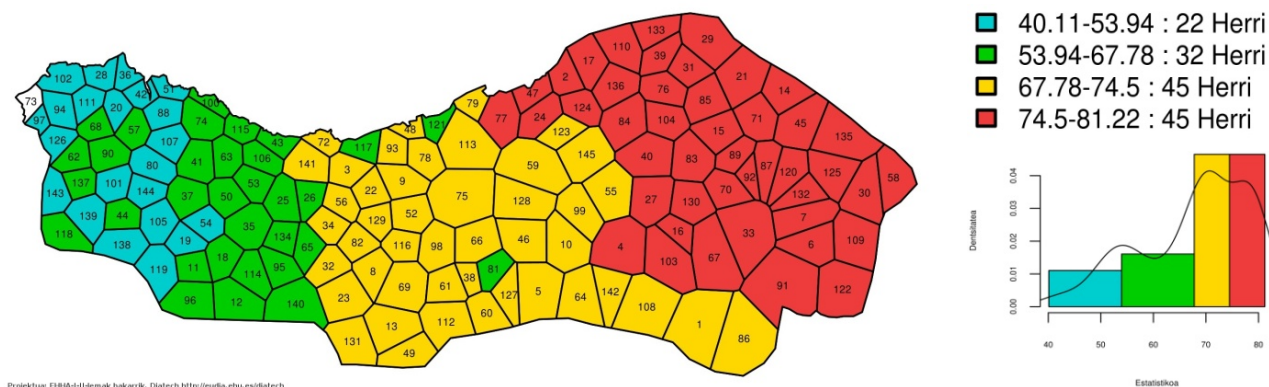
Emaitzak mapetan ezarriz, lehenik similaritate-mapak aztertuko dira. Similaritate-mapak hizkera batetik gainerakoetara dagoen hizkuntza-desberdintasuna neurtzen dute. Hizkera (herri bakoitzeko hizkuntza-ezaugarriak) bakoitza gainerako guztiekin konparatzen dira eta (des)berdintasunak zenbatzen dira. Adibide gisa, bi herri aukeratu dira, biak muturrekoak, bata mendebaldekoa (Getxo) eta bestea ekialdekoa (Eskiula).

#### *3.2.1. Getxoren similaritate-mapa*

Getxo herriari dagokionez (1. mapa), ikertzen diren gainerako herrietako hizkerarekin dituen berdintasunak lau multzotan banandu dira. Lehen multzoan (urdinez), linguistikoki hurbilenak dituen herriak agertzen dira; guztira 22 herri, denak Bizkaian kokatutakoak. Bigarren multzoan (orlegiz), 32 herri aurkitzen dira, bat izan ezik guztiak Bizkaian eta Gipuzkoan kokatuak. Hirugarren multzoan (horiz), 45 herri, Gipuzkoan eta Nafarroan kokatuak. Azkenik, laugarren multzoan (gorriz), 45 herri, denak Iparraldean kokatuak, Iparraldeko guztiak, hain zuzen ere.

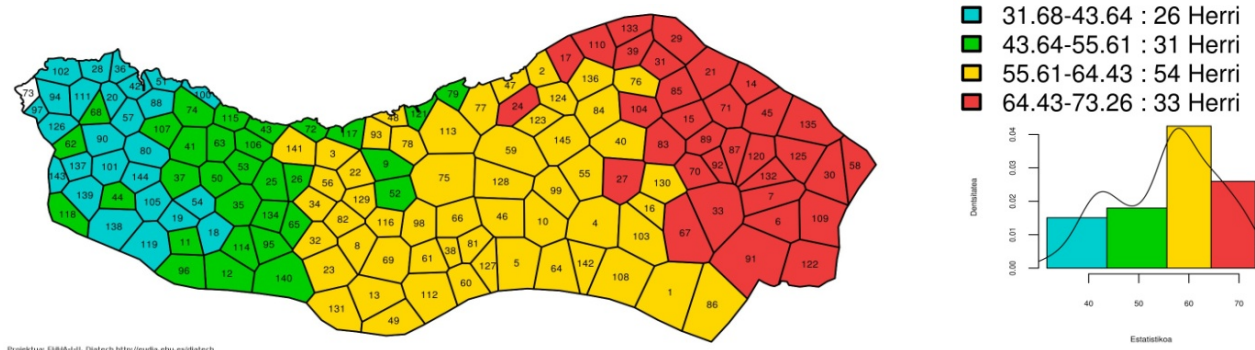
Deigarrienak diren zertzeladak baino ez aipatzearren: batetik, urdinez koloreztatutako eremuan agertzen diren orlegiz koloreztatutako herriak aipatu behar dira. Herri horiek hizkuntza-desberdintasun handiagoak dituzte gorriz koloreztatuta dauden inguruko herriek baino. Batetik, Errigoiti-57, Etxebarri-62, Fika-68, Larrabetzu-90, Zaratamo-137 herriek osatzen duten multzoa dago, eta, bestetik, Dima-44 eta Orozko-118 herriek osatua. Bigarren zertzelada horiz koloreztatutako multzoan agertzen diren orlegiz koloreztatutako hiru herriak dira (Igoa-81, Orío-117 eta Pasaia-121). Inguruko herrien eremuaz kanpo kokatzea zerk eragin duen jakiteko azterketa sakonagoa merezi dute, nahiz ez den hemen egingo.





**1. mapa:** lemak eta RIV izaria erabiliz, Getxo herriari dagokion similaritate-mapa.

Elkarren ondoan jarriz, batetik, lemak RIV izariarekin neurtuz lortutako mapa eta, bestetik, erantzunak  $L_v$  izariarekin neurtutakoa (2. mapa) konparatuz, bai multzoen arteko mugetan eta bai multzoen barnean desberdintasun batzuk badirela egiaztatu ahal dugu. Lehenik eta behin, hizkuntza-distantzien kopuruak azertu behar dira: lemekin eta RIV erabiliz lortutako mapak desberdintasun handiagoa erakusten du, bai distantzia txikienean (40,11) bai distantzia handienean (81,22); erantzunekin eta  $L_v$  izariarekin lortutakoan, ordea, 31,68 da, txikiena, eta 73,26, handiena. Multzo bakoitzean sartzen diren herrien kopurua ere desberdina da; batez ere, gorritz koloreztatutako multzoan (45 vs. 33).

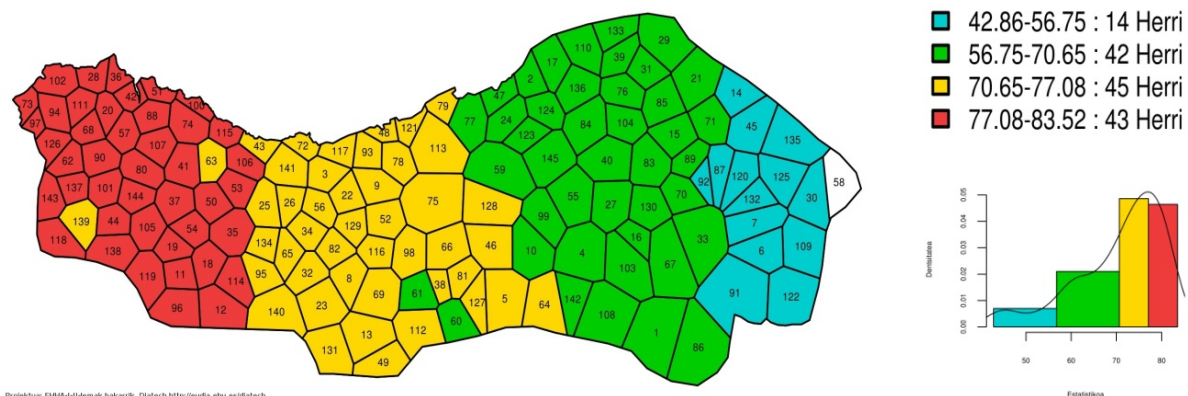


**2. mapa:** erantzunak eta  $L_v$  izaria erabiliz, Getxo herriari dagokion similaritate-mapa.

### 3.2.2. Eskiularen similaritate-mapa

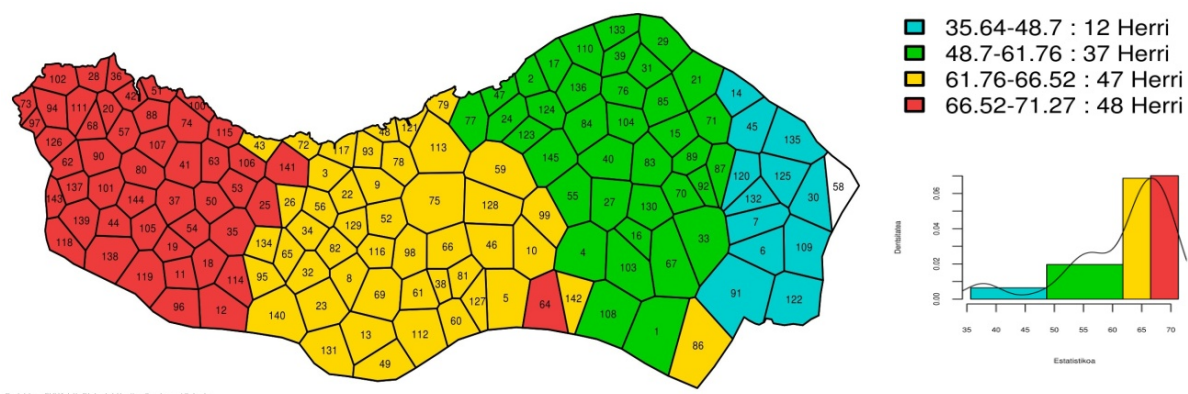
Ekialderen dagoen herria aukeratuz, Eskiulak gainerako herriekin dituen similaritateak agertuko dira (3. mapa). Eskiulak (mapan zuriz dagoen poligonoa) gainerakoekin dituen similaritateak lau multzotan bananduz, lau multzo geografikoki aski koherenteak lortzen dira: lehen multzoan (urdinez), Zuberoako eta Amikuzeko 14 herri agertzen dira; bigarren multzoan (orlegiz), 42 herri kokatzen dira: gainerako Nafarroa Behereko, Lapurdiko guztiak, eta Nafarroan, ekialdeko herriez gain, Etxaleku-60 eta Etxarri-Larraun-61. Hirugarren multzoan (horiz), 45 herri agertzen dira: Nafarroako eta Gipuzkoako herriak eta

Bizkaiko Etxebarria-63 eta Zeberio-139. Eta, azkenik, laugarren multzoan (gorriz), Gipuzkoako ekialdeko eta Bizkaiko herriak, gorago aipatutako Etxebarria eta Zeberio izan ezik; etorkizunean aztertu beharreko herriak ditugu hauek.



**3. mapa:** lemak eta RIV izaria erabiliz, Eskiulari dagokion similaritate-mapa.

Mapa hau erantzunekin eta *Lv* izariarekin lortutakoa (4. mapa) elkarren artean konparatuz, batetik, nabarmentzen da hizkuntza-distantziak ez direla berak: distantzia txikiena 42,86, batean, eta 35,64, bestean; handienak ere berdintsu: batean, 83,52 eta, bestean, 71,27. Bietan, erantzunekin *Lv* izaria erabiliz lortu da txikiena. Multzoetan ezarritakoen gure begiak, aldaketa handirik ez bada ere, badira xehetasunak: lehen mapan, Bizkaian, Etxebarria eta Zeberio dira agertzen gorrien eremuan horiz koloreztatuak. Gorriz koloreztatutakoaren eta horizkoaren arteko muga ere ez da guztiz bera. Hirugarrenik, Nafarroan ere badira desberdintasunak bai iparraldean eta bai hego-ekialdean. Eta azkenik, eremu orlegiaren eta urdinaren arteko muga ere ez da arras berdina.



**4. mapa:** erantzunak eta *Lv* izaria erabiliz, Eskiulari dagokion similaritate-mapa.

### 3.3. Sailkapen hierarkikoa

Hizkeren sailkapen hierarkikoa egin da, *cluster analisia* deitzen dena. Cluster analisiak hizkerak hizkuntza-berdintasunen arabera multzokatzen ditu: hurbilen diren bi hizkera biltzen dira, lehenik, eta multzo bat eratzen dute; ondoren, berriro aztertzen da zein hizkera edo hizkera-multzo diren hurbilen, eta beste multzo bat egiten da, eta horrela hizkera guztiak multzo bakarrean bildu arte.

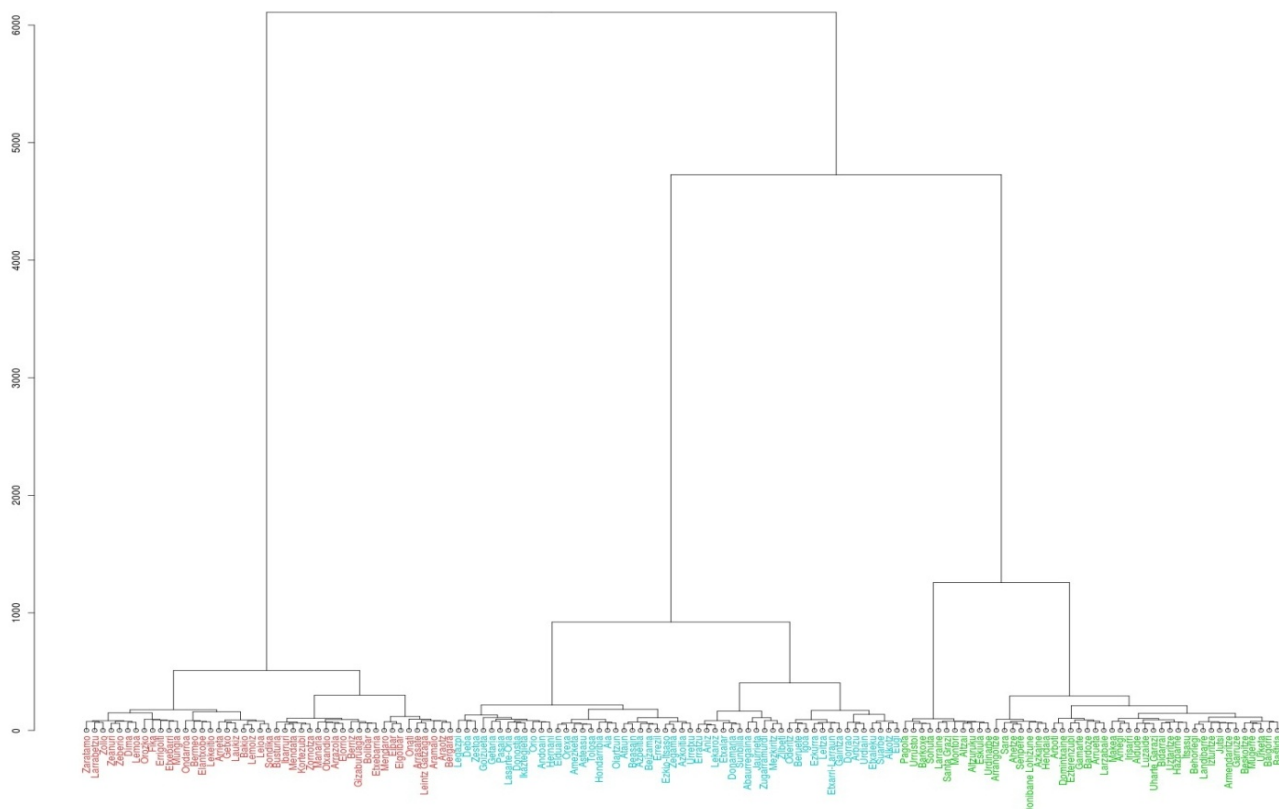
Bariazioa kuantifikatzeko, dialektologian, gaur egun, oso ohikoa da teknika hau. Emaidza estatistiko bi ematen ditu ohiz: dendrograma, edo zuhaitz hierarkikoa deitua, eta honetatik ondorioztatzen diren mapak. Ekarpak honetan biak aurkeztuko dira.

#### 3.3.1. Dendrograma

Dendrogramak hizkerak, hurbilenetik hasiz, hierarkia batean ezartzen ditu, eta buruz behera ezarri den zuhaitz baten itxura hartzen du. Dendrograma eraikitzeke metodo edo algoritmo asko dira. Ikerketa honetan, *Ward metodoa* erabili da, metodo hau delako aseptikoena eta distortsio gutxien sortzen duena. Gainerako zehaztasunak aurreko analisian erabilitako berak dira.

Dendrogramaren mozketarako, bi irizpide erabiliko dira: Guter-en irizpidea eta dendrogramaren besoen luzera. Guterrek (1973) hizkeren arteko desberdintasunak mailakatzea proposatu zuen; haren arabera, % 0-25 bitarteko desberdintasunek hizkerak bereizteke balio lezake; % 25-50 bitarteko desberdintasunek, azpidialektoak bereizteke; % 50-75 bitartekoek, dialektoak bereizteke, eta hortik aurrerako ehunekoek, hizkuntzak bereizteke. Aurrekoetxeak (2012) hizkeren arteko hizkuntza-mugen beharrezko mailakatzeaz jardutean, ideia berrartu zuen, eta hiru maila proposatu zituen. Ikerketa honetan, sortutako dendrogramaren eskalak orotara 6.000 maila ditu (1. irudia): Guterren arabera jokatu, eskalako % 25erainoko tartean, barietateak edo hizkerak bereizten dira; % 50erainokoan, azpidialektoak, eta % 75erainokoan, dialektoak. Gorago aipatu dendrogramara itzuliz eta % 25ean moztuz, eskalako 1.500 mailan, hiru multzo genituzke; % 50ean eginik, eskalako 3.000 mailan ere aurreko hiru multzo berak izango genituzke; eta azkenik, % 75ean eginik, eskalako 4.500 mailan ere aurreko hiru multzoak lortuko lirateke. Gure kasuan, eta aurreko arlo honetako literatura kontuan izanik, ez dirudi sistema egokiena denik.

Bigarren irizpidea estatistikan cluster analisia egiten dutenen artean oso ezaguna da: dendrograma bat izanik, hizkeretatik (irudiaren behealdean koloretan) sortzen diren adarren luzera hartzen da kontuan.



1. irudia: cluster analisiaren bidez lortutako dendrograma

Adar horien luzetasunak hizkuntza-desberdintasuna adierazten du. Hizkuntza-desberdintasunarekin hizkuntza-distantzia esan nahi da. Adarra zenbat eta luzeagoa izan, orduan eta hizkuntza-distantzia handiagoa da adar-muturrean biltzen diren bi multzoen artean. Dendrogramak argi erakusten du gorritz idatzita dauden hizkerak osatzen duten multzoa dela, batetik, trinkoena (beste bi multzoak baino eskalan beherako biltzen da) eta, bestetik, hiruetarik bereziena (gainerako bi multzoekin biltzen azkena delako). Aztertzen diren hizkera guztiak bi multzotan bildu behar balira, gorritz idatzitzako hizkerak batetik eta urdin-orlegiz idatzitakoek osatzen dutena bestetik lortuko lirateke. Beraz, erdiguneko hizkeren eta ekialdeko guneko hizkeren artean hurbiltasun handiagoa dago gorritz idatzitako multzoa osatzen dutenen artean baino.

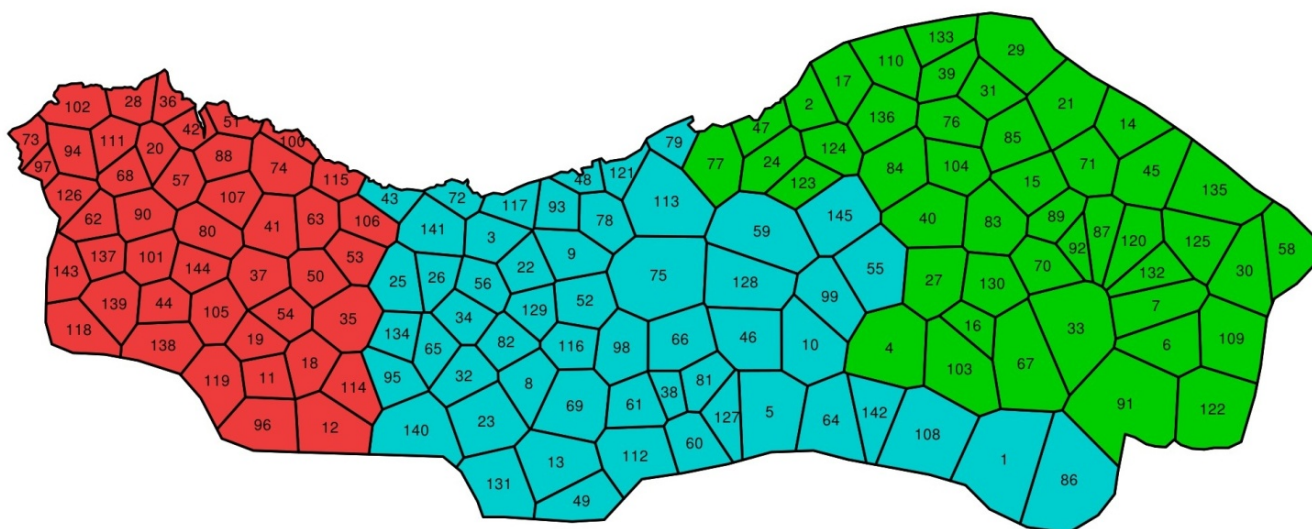
Dendrograman agertzen den sailkapena mapa batean islatu behar denean, dendrograma hori nonbait moztu behar izaten da; mozketa horretan sortzen diren multzoen kopurua mapan kolorez bereizten diren eremuetan islatuko da. Beraz, garrantzi handia du dendrograma non moztu behar den zehazteak, edo zein den mozteko lekurik egokiena zehazteak, banaketa dialektal egokia lortzeko.

Funtsean, honek zera esan nahi du: muga dialektal guztiek ez dutela garrantzi bera, muga batzuk garrantzitsuagoak direla, hizkuntza-desberdintasun handiagoa erakusten dutelako, eta, alderantziz. Gauzak horrela, dendrograma moztu nahi denean, garrantzi

handia dauka mozketa-leku edo maila ondo aukeratzeak. Mota honetako estatistika ikertzen duten tresnen artean badira tresna automatizatuak dendrogramaren mozketa-maila egokiena egiten laguntzen dutenak. Denek bat egiten dute mozketa adarraren luzetasunarekin lotzean. Gure kasura etorrira, eskalako 1.200 maila ingurutik gora, hiru adar esanguratsu agertzen dira: lehenak gorritz idatzita dauden herriak biltzen ditu (askoz beherago bildutako multzoa); bigarrenak, urdinez idatzitakoak, eta hirugarrenak, orlegiz idatzitakoak, hain zuzen maila horretan biltzen direnak.

### 3.3.2. Lexikoari dagokion euskalkien mapa zientifikoa

Dendrograma 1.250 mailan moztuz (hizkuntza-desberdintasunen % 20 gutxi gorabehera) hiru multzotan biltzen dira aztertzen diren 145 hizkerak. Multzo horiek, mapan ezarrita, hiru eremu dialektal marraztuko dute (5. mapa).



Proiektua: EHHA-II-Hemak bakarrik. Diotech <http://eudia.ehu.es/diotech>

**5. mapa:** lemak RIV izariarekin neurtuz cluster analisiaren bidez lortutako mapa (ward, 3 multzo).

Mapan hiru multzo azaltzen dira: mendebaldekoa, gorritz margotutakoa; erdialdekoa, urdinez margotua, eta ipar-ekialdekoa, orlegiz margotua. Bigarren aldia da datu lexikalekin euskararen eremu osoko datuetan oinarrituz eta prozedura kuantitatiboak erabiliz mapa bat egiten dena eta lehen aldia datu lematizatuekin egiten dena. Garrantzi handiko emaitza da, zalantzarik gabe; orain arte inoiz ez delako lexiko hutsean oinarritutako euskalkien maparik egin. Laburbilduz, mapa hau lortzeko, Euskal Herri osoan barreiatutako 145 herritan lortutako lexiko orokorreko 183 galderaren erantzun lematizatuak erabili dira, gorago esan bezala. Guztira 26.535 datu bildu dira sortu den datu-basean. Datu horiek *Diotech* programarekin aztertu dira mapa hau egiteko.

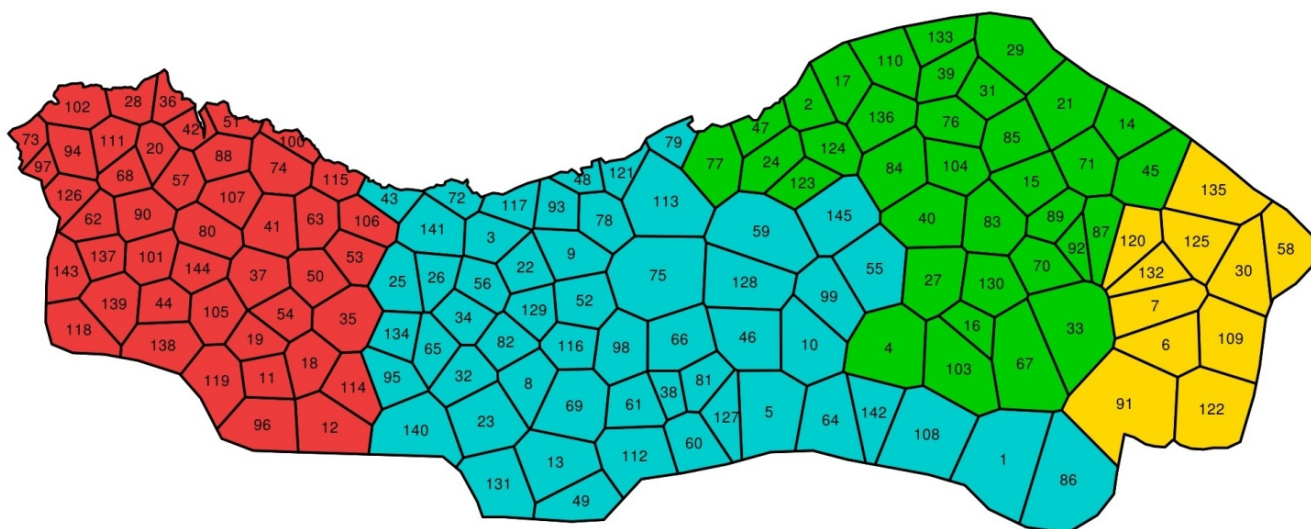
Agian, gorriz koloreztatutako mendebaldeko eremua arrunt egokitzen jo lezake edonork, orain arteko banaketa dialektaletan ohiko bihurtu baita eremu hori; gune honen eskuinaldeko hizkeren lerroa goitik behera hauxe da: Ondarroa-115, Mendaro-108, Elgoibar-53, Bergara-35, Oñati-114 eta Araotz (Oñati)-12. Guztira, 45 hizkera biltzen ditu, zeinek orain arteko dialektologia tradizionalaren metodologiaz lortutako mapek marrazten duten eremuaren eskuineko muga osatzen baitute, nahiz eta Elgoibar eta Mendaro, banaketa batean edo bestean, zalantzazkotzat edo zubi-hizkeratzat hartuak izan diren.

Erdiguneko eremu urdina aztertzerakoan, zalantza gehiago sor lezake Gipuzkoako eta Nafarroako lurralde eremu bakarrean agertzeak. Baina dendrograman begiratu gero, oso eskalako maila apalean biltzen dira (900 mailan gutxi gorabehera): desberdintasunen % 15en mailan eta Iparraldeko hizkerak biltzen diren maila baino apalagoan. Eremu honetan agertzen dira (mendebaldeko eremuan sartzen ez diren) Gipuzkoako hizkerak eta Nafarroako denak, Luzaide-103 izan ezik. Eremu trinkoa da oso, nahiz alboetako bi eremuen tartean izan. Guztira, 54 hizkera biltzen ditu,

Orlegiz koloreztatutako eremua dela eta, aldiz, inor gutxi egon liteke ados, aditu izan zein ez izan, ezen, Lapurdiko eta Behe Nafarroako hizkeren arteko mugak zalantza sortzen badu ere, zubereraren eta behe-nafarreraren arteko muga oso gardena dela, oso barneratuta dago-eta. Dena den, irakurleak kontuan izan behar du orain arteko banaketa eta sailkapen guztietan nagusiki gramatika ezaugarriak hartu direla kontuan, eta ez lexikoa. Dendrogramako eskalan begiratu gero, multzo hau 1250 mailan biltzen dela egiaztatuko da. Hegoalde eta Iparraldeko “muga” ere garbi-garbi da, Luzaide-103 izan ezik, hau Iparraldeko herriekin lerratzen da eta. Guztira 46 hizkera biltzen ditu eta desberdintasunen 20,8 biltzen ditu (eskalaren 1.250 mailan biltzen da).

Mapa, bestalde, oso koherentea da, eremuak oso argiak dira eta ez da hizkerarik edo herririk eremuz kanpo kokatzen denik. Koherentziaz gain, eta gure orain arteko ezagutzaz gain, funtsezko arazo baten, arazo teoriko baten, aurrean kokatzen gaitu honek: zenbat euskalki daude edo zer behar da izan euskalki izateko? Eztabaida hau modu kuantitatiboan bideratzeko —eta ez dakit nola bidera daitekeen modu kuantitatiboan ez bada— hizkuntza-distantziak neurtu behar badira mugak ezartzeko. Horretarako, dendrograma aztertzea da aukera zehatzena: dendrograman agertzen diren lerro nagusiak hiru eremu hauei dagozkie. Kontuan hartzen bada lerroaren luzerak desberdintasuna adierazten duela (hemen hizkuntza-distantzia), dendrogramak adierazten du aztertzen diren hizkerak sailkapena hiru eremutan egitea dela.

Baina dendrograma hiru eremutan banandu ordez lautan banatzea bada aukera? Dendrograma, lau multzotan banatzeko, 1.200 inguruko mailara jaitsi beharko genuke. Maila horretatik behera banatzen dira Zuberoako hizkerak (Domintxaine izan ezik) Nafarroa Behereko eta Lapurdiko hizkeretatik.



Proiektua: EHHA-I-II-Hemak bakarrik. Diatex <http://eudia.ehu.es/diatex>

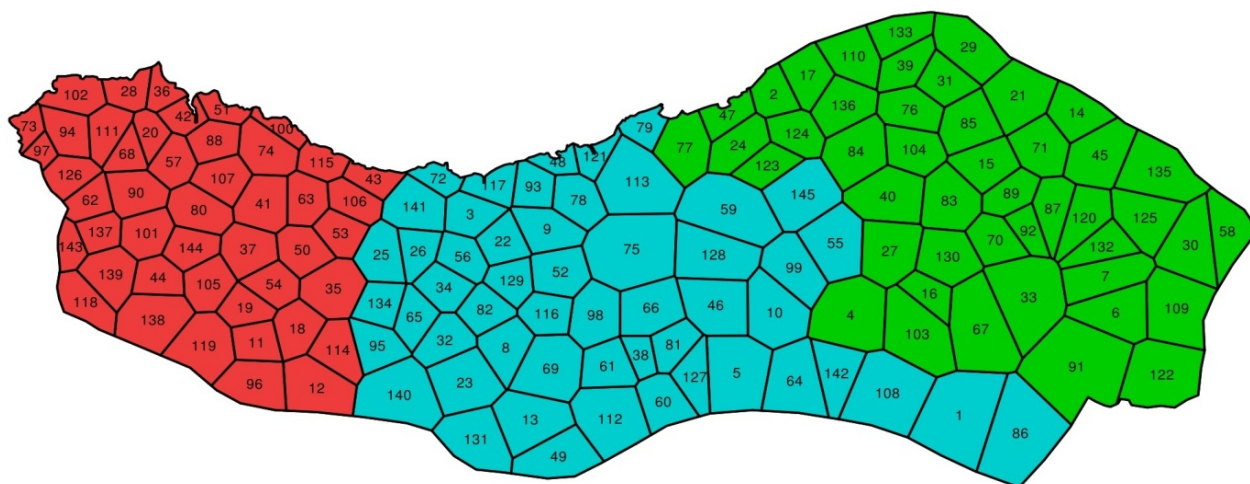
#### 6. mapa: cluster analisisian dendrograma 4 multzotan banandu osteko mapa (Lemak, RIV, Ward-4)

Baina dendrogramari begiratzuz gero, Zuberoako eremu honen eta Iparraldeko gainerako herrien muga oso desberdina da, beste bi mugekin erkatzen bada: muga hau oso maila apalean biltzen da dendrograman, eta gainerako biak (eremu gorria eta urdina bereizten dituen, eta urdinaren eta orlegiaren artekoa) oso gora dagoen mailan biltzen dira. Hots, Zuberoako herri hauen eta Iparraldeko gainerakoen artean dagoen hizkuntza-desberdintasuna oso txikia dela gainerako eremuen artean dagoenarekin konparatuz. Mugen pisua, balioa, garrantzia, desberdintasuna handia da maila berean ipintzeko.

Hori dela eta, uste dugu lexikoari dagokion euskalkien mapa hiru eremutan banatzea dela egokiena: mendebaldeko eremua, erdialdean oso gune handia osatzen duen eremua eta ipar-ekialdeko eremua. Mendebaldeko hizkerak 500 mailan biltzen dira multzo batean; oso maila apala da, eta eremu horren kohesioa eta trinkotasuna erakusten du. Gipuzkoako eta Nafarroako hizkerak (Goizueta izan ezik, hau Gipuzkoakoekin agertzen da eta) 900-1.000 mailan biltzen dira.

#### 4. Mapen erkaketa

Lemen mapa erantzunen maparekin erkatuz gero, harrigarriki bi mapak puntu guztietan bat datoz, hizkera batean izan ezik: Deba-43. Gainerako herri guztiak eremu edo multzo beretan kokatuta agertzen dira. Beraz, kasu honetan, erantzunak eta *Lv* izaria erabiliz sortutako hiru eremuko sailkapen hierarkikoa eta lemak eta RIV izaria erabiliz sortutako hiru eremuko sailkapen hierarkikoaren ondorio den mapa berak dira.



Proiektua: EHH4-I-II. Diatex <http://eudia.ehu.es/diatex>

7. mapa: erantzunak eta *Lv* izaria erabiliz 3 eremuko mapa (Ward).

## 5. Euskalkien eta azpieskalkien mapa

Emaitzak aztertu ondoren, oso garbi adierazi behar da, erabili diren datuak kontuan izanik, hiru eremuko banaketa egiten dela antzeko mailan biltzen diren multzoak mapetan irauliz. Eta hiru eremuon arteko hizkuntza-distantzia handia dela. Eremuok bereizteko edo banatzeko, eskalako maila asko jaitsi behar da.

### 5.1. Euskalkien mapa

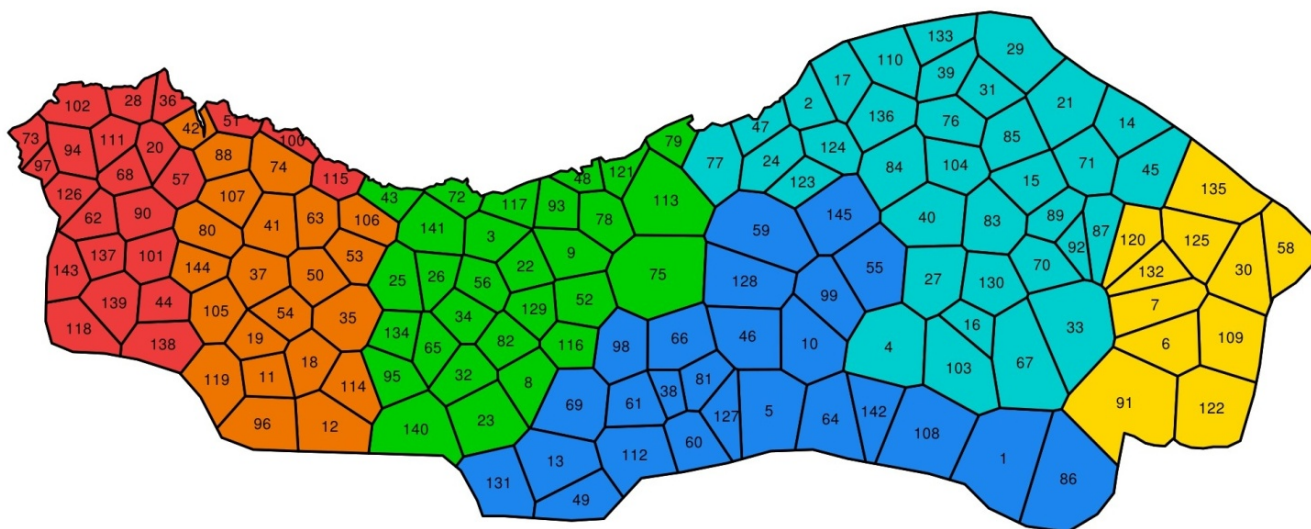
Gorago esan da hiru eremuko mapak argi erakusten duela banaketa egokiena bera dela. Hiru eremu horiek nola izendatu da, orain, eskuarteko arazoa. Dialektologia tradizionalak ez luke arazorik izango eremu hauei “euskalki” izendapena emateko. Baina gaur egungo dialektologian ez da garbi izendapen hori emateko behar diren baldintzak zein izan behar diren, eta hori aztertu eta horretaz adostasun bat lortu bitartean, eremu dialektalez hitz egiten da, izendapenean gehiago zehaztu gabe.

Gure kasuan, eta erabaki bat hartu behar balitz, eztabaidarako ez bada ere, eremu hauek “euskalki” izenez izendatzea ez litzateke guztiz desbideratua izango (ikus 5. mapa). Honek, oster, gaiari dagokion orain arteko literatura guztia hankaz gora jartzen du. Sor litekeen eztabaidarako hau ere, kontuan hartu behar da ezaugarri lexikalak baino ez direla erabili sailkapen honetan eta gramatikari buruzkoak ez direla erabili. Bistan da lexikoari dagozkion ezaugarriek eta ezaugarri gramatikalek oso sailkapen eta banaketa desberdina sor lezaketela; ez dute nahitanahiez sailkapen bera edo berdintsua eman beharrik.



## 5.2. Azpieuskalkien mapa

Dendrogramaren mozketan egokienak hiru multzotan banatzea eskatzen duela aipatu da gorago. Horiek dira eremu honetako euskalkiak. Baina euskalki bakoitza bitan banandu nahi izanez gero, eskalako 500 maila ingurura jaitsi beharko ginateke. Maila horretan, hiru eremuak bikoizten dira, eta honako mapa lortzen da:



Proiektua: EHHA-I-II-lemak bakarrik. Diatex <http://eudia.ehu.es/diatex>

8. mapa: azpieuskalkien mapa

Mendebaldeko eremuan bi azpierrezu ikusten dira, goitik beherako marra bertikal batek bereziak. Baina itsasaldeko hizkerak mendebalde-muturreko eremuan kokatzen dira, Busturia (42) hizkera izan ezik. Erdialdeko euskalkia bitan banatzen da: Nafarroakoak (Goizueta salbu), alde batetik, eta Gipuzkoakoak, bestetik. Ekialdeko euskalkia ere bitan banatzen da: Lapurdi eta Nafarroa Behereko hizkerak (eta Nafarroako Luzaide-103), batetik, eta Zuberoakoak, bestetik, (Domintxaine-45 izan ezik).

## 6. Ondorioak

Lexikoa erabiliz, euskalkien sailkapen hierarkiko automatizatuaren bigarren emaitzak dira lan honetan aurkeztu direnak; lehenak datuen hizkuntza-analisirik gabe egindakoak izan ziren (Aurrekoetxea, Ormaetxea & Videgain 2018). Horretarako, erantzunen lematizazioa egin da, lehenik, eta hizkuntza-distantzia gisa RIV izaria erabili da. Euskal dialektologian datu lexikalak erabiliz burutzen den bigarren euskalkien sailkapena da. Eta ez hori bakarrik: erabili diren metodologia eta sailkapen-teknikak ere aurreko sailkapenetan erabilitakoetatik oso desberdinak dira. Dena den, uste dugu lexikoaren analisia egitean datu lematizatuak erabili behar direla sailkapena ezaugarri lexikal hutsen gainean egiteko, hitzetan ager litezkeen ahoskera-ezaugarriek ez dezaten eraginik izan; beraz, artikulua

honetan aurkeztu diren mapak dira, gure ustez, egokienak. Euskararen sailkapen geolinguistikoari begira, lexikoari dagokion lana biribildu dela deritzogu.

Lexikoan egin den bezala, gainerako gramatika-ataletan ere egin dira beste sailkapen batzuk, etorkizunean bildu eta elkarrekin konparatu beharko direnak.

## 7. Erreferentzia bibliografikoak

- Aurrekoetxea, G., 1986, “Euskal Herriko Hizkuntz Atlasa (EHHA): inkesta metodologia eta ezezko datuak”, *Euskera* XXXI, 413-424.
- Aurrekoetxea, G., 2002, “Algunas consideraciones sobre la contrapregunta en las encuestas lingüísticas”, in L. Rabassa (ed.), *Mélanges offerts à Jean-Louis Fossat*, Université de Toulouse II-Le Mirail, CerCLid 11/2, 57-65. [[http://artxiker.ccsd.cnrs.fr/docs/00/07/14/00/PDF/Contrapregunta\\_encuestas\\_ling.pdf](http://artxiker.ccsd.cnrs.fr/docs/00/07/14/00/PDF/Contrapregunta_encuestas_ling.pdf)]
- Aurrekoetxea, G., 2012, “Towards a scientific measurement of linguistics boundaries”, in Álvarez Pérez, Xosé Afonso, Ernestina Carrilho & Catarina Magro (eds.), *Proceedings of the International Symposium on Limits and Areas in Dialectology (LimiAr). Lisbon, 2011*. Lisboa: Centro de Linguística da Universidade de Lisboa, 23-34. ISBN: 978-972-96404-9-0 [http://limiar.clul.ul.pt/proceedings\\_en.html](http://limiar.clul.ul.pt/proceedings_en.html)
- Aurrekoetxea, G., 2016, “Analysis of the morphological variation of Basque”, *DiG* 24, 21-41 [DOI 10.1515/dialect-2016-0002]
- Aurrekoetxea, G., Ormaetxea, Tx., Videgain, X., (2018), “Euskalkien sailkapen zientifikoa: lexikoa”. In L. Unamuno, A. Romero, A. Etxebarria & A. Iglesias (arg.), *Linguistic Variation in the Basque and Education-III / Euskararen bariazioa eta bariazioaren irakaskuntza-III*, Bilbo: UPV/EHUko Argitalpen Zerbitzua, 126-140.
- Aurrekoetxea, G. eta Videgain, X., 1987, “Euskal Herriko Hizkuntz Atlasaren ezaugarri nagusiak”. In askoren artean, *II. Euskal Mundu Biltzarra*, 2.II, 57-462.
- EHHA: Euskaltzaindia, 2010-2018, *Euskararen Herri Hizkeren Atlasa*, I-IX, Bilbo: Euskaltzaindia.
- DIATECH: <http://eudia.ehu.es/diatech/index/>
- Guiter, H., 1973, “Atlas et frontières linguistiques”. In G. Straka & P. Gardette (eds.) *Les dialectes romans de France à la lumière des atlas régionaux*, (col. de Strasbourg, 1971). Paris. 61-109.
- Goebel, H., 1978, “Analyse dialectométrique de quelques points de l'AIS (italien standard valdotain provençal alpin turinois milanais)”. In G.E. Clivio & G. Gasca Queiraza (arg.), *Lingue e dialetti nell'arco alpino occidentale. Atti del Convegno Internazionale di Torino (1976)*. Turin, 282-294.
- Iglesias, A., Aurrekoetxea, G., 2017, “Intonazioaren egituraketa geolinguistikoa”. In A. Etxebarria & N. Eguskiza (arg.), *Bariazioa esaldien intonazioan*. Bilbo: UPV/EHUko Argitalpen Zerbitzua, 277-295.

- Nerbonne, J., 2008, “Variation in the Aggregate: An Alternative Perspective for Variationist Linguistics”. In K. Dekker, A. MacDonald, & H. Niebaum (Eds.), *Northern Voices: Essays on Old Germanic and Related Topics offered to Professor Tette Hofstra*. Leuven: Peeters, 365-382.
- Nerbonne, J., & Heeringa, W., 2010, “Measuring dialect differences”. In Jürgen Erich Schmidt & Peter Auer (arg.) *Language and Space: An International Handbook of Linguistic Variation, Volumen I*. Berlin: Mouton De Gruyter, 550-565.
- Séguy, J., 1971, “La relation entre la distance spatiale et la distance lexicale”, *Revue de Linguistique Romane* 35, 335-357.
- Videgain, X., Aurrekoetxea, G., 2016, “Euskararen bariazio geo-morfologikoaren azterketa”. In Aitor Iglesias, Asier Romero, Ariane Ensunza (arg.), *Linguistic variation in the basque language and education - II / Euskararen bariazioa eta bariazioaren irakaskuntza -II*, Bilbao: Universidad del País Vasco / Euskal Herriko Unibertsitatea, Argitalpen Zerbitzua / Servicio Editorial, 25-36.
- Zuazo, K., 1998, “Euskalkiak, gaur”, *FLV* 78, 191-233.

## **BIDEO-PODCASTA: ERRONKETAN OINARRITUTAKO IKASKUNTZA<sup>5</sup>**

Iglesias, A., Etxebarria, A., Pagadigorria, A. & Unamuno, L.

UPV/EHU

[aitor.iglesias@ehu.eus](mailto:aitor.iglesias@ehu.eus), [aintzane.etxebarria@ehu.eus](mailto:aintzane.etxebarria@ehu.eus), [aintzane.oagadigorria@ehu.eus](mailto:aintzane.oagadigorria@ehu.eus),  
[lorea.unamuno@ehu.eus](mailto:lorea.unamuno@ehu.eus)

**Laburpena:** Lan honetan hezkuntza berrikuntzarako proiektu baten lehen fasea aurkezten da. Erronketan oinarritutako ikaskuntza erabiliz, titulazio ezberdinetako ikasleek modu kolaboratiboan bideo-podcasta sortu behar dute Lehen Hezkuntzako geletara eroateko material gehigarri gisa. Lehen fase honetan, proiektuan parte hartu duten ikasleen motibazioaren balorazioa aztertzen da datu estatistikoaren bidez. Datuok motibazio estrinseko handia eta amotibazio txikia erakusten dute.

**Abstract:** In this work the first phase of an educational innovation project is presented. Through learning-based on challenges, students from different degrees have created a video-podcast that will be taken to primary education classrooms as complementary material. The assessment of the motivation of the participants has been analyzed through statistical data. These data show high extrinsic motivation and low amotivation.

### **1. Sarrera**

Gaurko irakaskuntza-sistemaren helburu nagusietakoa da gaitasunen garapena IKTen erabilpenaren laguntzarekin. Testuinguru honetan oso garrantzitsua da erronka errealak planteatzea ikasleei, eurek lan kooperatiboan bitartez egin ditzaten (Evans & More, 2013). Artikulu honetan aurkezten duguna hau da: ikasle unibertsitario batzuen motibazioa erronka erreal batean zehar. Erronka honetan hiru diziplinatarako kideekin lan egin behar dute, Lehen Hezkuntzako umeentzat izango den bideo-podcast bat sortzeko.

### **2. Marko teorikoa**

Autodeterminazioaren teoriari (Deci & Ryan, 1985) egoki deritzogu hezkuntzako motibazio motak ulertzeko; beraz, horretan oinarrituko dugu proiektu honen marko teorikoa.

Self-Determination Theory is an approach to human motivation and personality that uses traditional empirical methods while employing an organismic metatheory that highlights the

---

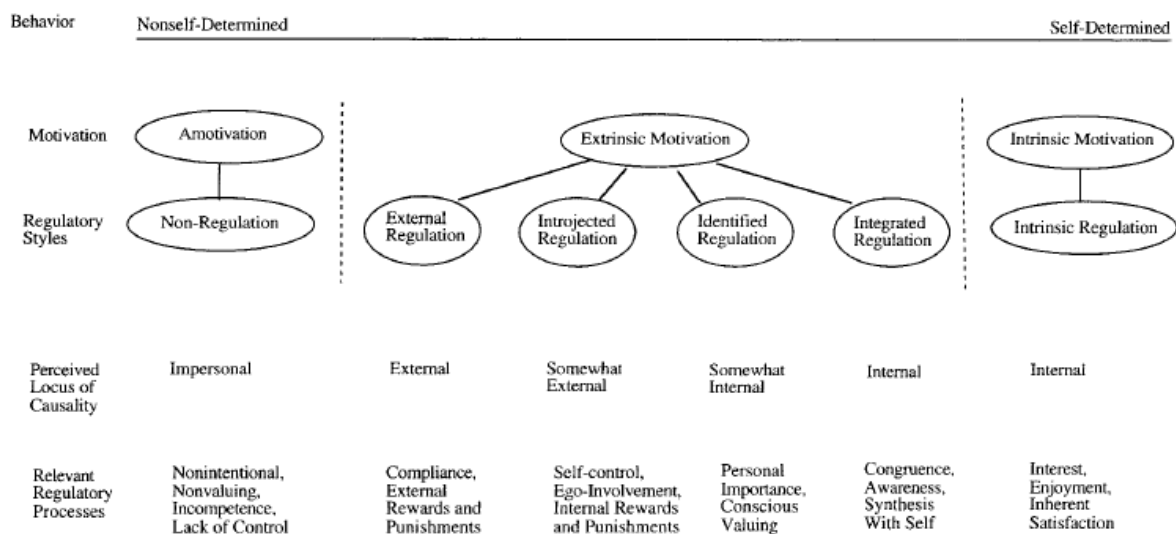
<sup>5</sup> Videocastinga eta podcastinga: diziplina ezberdinetako ikasleen euskarazko ahozko komunikazio gaitasuna lantzeko bideak izeneko Hezkuntza Berrikuntzarako Proiektuak (HBP) UPV/EHUko Berrikuntzaren, Gizarte Konpromisoaren eta Kulturagintzaren arloko errektoreordetzaren laguntza jaso du.

importance of humans' evolved inner resources for personality development and behavioral self-regulation. (Ryan & Deci, 2000: 68)

Teoria honek hiru motibazio mota bereizten ditu autodeterminazio-mailaren arabera: motibazio eza, motibazio estrinsekua eta motibazio intrinsekua (Martín-Albo et al., 2009). Motibazio eza, izenak adierazten duen moduan, motibaziorik ez dagoeneko da eta ez du erregulaziorik.

Motibazio estrinsekua pizgarrien bidez egiten den jarduerari deritzo, eta duen erregulazioaren arabera, lau motibazio estrinseko mota bereizten dira; hala, batetik, kanpoko erregulazioa, egiten den hori saria lortzeko edo zigorrik ez pairatzeko egiten denekoa eta autodeterminazio maila baxuenekoa. Bestetik, erregulazio introiektatua, norbaitek errudun ez sentitzeko egiten duenekoa. Bestetik, identifikatutako erregulazioa, eginkizuna egilearen helburuetarako garrantzitsua denekoa; honek erregulazio-maila autodeterminatuagoa du. Azkenik, integratutako erregulazioa, eginkizunaren ondorioak norberaren balio eta beharrek bat datozeneko, motibazio estrinsekoetan autodeterminatuena (Martín-Albo et al., 2009: 800).

Motibazio intrinsekua norberak du; ez du kanpoko faktoreen premiarik, eta autodeterminazio-maila handiena eta erregulazio intrinsekua ditu.



**1. irudia:** The Self-Determination Continuum Showing Types of Motivation With Their Regulatory Styles (Ryan & Deci, 2000: 72)

Erronketan oinarritutako ikaskuntzak trebetasun batzuk garatzen laguntzen du, hala nola bai lan kolaboratiboa eta bai diziplina anitzekoa egiteko trebetasunak, erabakiak hartzekoak, komunikatzekoak, etika arlokoak edota lidergoari dagozkionak (Malmqvist, Rådberg & Lundqvist, 2015). Zenbait urrats edo pausori (2. irudia) jarraituz eta ikasleen rol aktiboa bultzatuz ikaskuntza eraikitzea eta ikaskuntza bera lortzea da helburua. (Schnackenberg et al., 2009).



**2. irudia:** ikasleen urratsak (Cordray, Harris & Klein, 2009: 336).

Lehenik eta behin, planteatzen den erronkak pizgarria izan behar du ikaslearentzat. Ikasleek aukera bat ikusi behar dute euren ezaguera praktikan jartzeko erronketan oinarritutako ariketa batean. Planteatzen den erronka aztertzeko, aurretiaz, liburuak, aldizkarien artikuluak, bideoak, etab. landu behar dituzte, askotariko ikuspegia lortzeko eta ondoren ikertzeko. Ikertzeko materiala sortu behar da eta ostean probatu, zeren probak adierazten du materialak balio duenentz. Materiala ez bada baliokoa, berraztertu egin behar da eta, aldiz, proba gainditu badu, publiko egiten da arazoaren konponbidea (Cordray, Harris & Klein, 2009).

Hainbat egilek jadanik podcastaren erabilerak eragin positiboa duela frogatu dute; hau da, ikasketa-prozesua errazten du (Piñeiro & Costa, 2011), ikaskuntzaren prozesuan laguntzen du (Guiza et al., 2016) eta aberastasuna sortzen du, ikasleen arreta eta gogoeta sustatzen ditu, asetasun-maila handiagoa sorraratzen du (Miller & Piller, 2005) eta antsietatea murrizten du (Piñeiro & Caldevilla, 2011).

### 3. Metodologia

Esan dugun moduan, ikasle hauen erronka bideo-podcasta sortzea izan da, Lehen Hezkuntzako geletara erodateko material gehigarri gisa. 170 ikaslek parte hartu dute lan honetan: Hezkuntza Fakultateko 99k, Zientzia eta Teknologia Fakultateko 29k eta Gizarte eta Komunikazio Zientzien Fakultateko 42k. Partaideen generoari dagokionez, % 70 neskak dira, eta % 30, mutilak; ama-hizkuntzari dagokionez, % 58,8k gaztelania du, eta % 41,2k euskara.

Une honetan proiektu honen lehen fasea aztertzen ari gara. Bertan ikasleak lan kolaboratiboan ditugu beharrean eta aztertu nahi dugu euren motibazioa zein den. Beraz, lan honetan aurkezten diren emaitzak hezkuntza-berrikuntzarako proiektu honek sortu dien motibazioarekin zerikusi duten faktoreak dira. Proiektuaren lehen urtean lortutako emaitzak dira: 2016ko lehen lauhilabetekoan hasi, eta 2017ko maiatzean amaitu zen, guztira 15 aste. Ikasleek, bost-sei pertsonako taldeetan, beste gradu batzuetako ikasleekin lan egin dute.

Hauek dira gradu bakoitzaren ezaugarriak eta landutako oinarritzko eta zeharkako gaitasunak.

Zientzia eta Teknologiako Fakultatean, Matematika, Fisika eta Ingeniaritza Elektronikoko graduetak ikasleak izan ditugu *Komunikazioa Euskaraz* irakasgai. Hauek landutako oinarritzko gaitasunak honako hauek dira: bilatzea, ulertzea, sintetizatzea eta zientzia era kritikoan aztertzea; eta beste hauek zeharkako gaitasunak: talde-lana eta berritzapena.

Hezkuntza Fakultatean, Lehen Hezkuntzako Graduak ikasleak izan ditugu *Komunikazio Gaitasunaren Garapena I: euskara eta gaztelania* irakasgai. Hauek landutako oinarritzko gaitasunak honako hauek dira: bilatzea, ulertzea, sintetizatzea eta hizkuntzari eta literaturari buruzko informazioa era kritikoan aztertzea; eta beste hauek zeharkako gaitasunak: talde-lana eta berritzapena.

Gizarte eta Komunikazio Zientzien Fakultatean, Ikus-entzunezko Komunikazioko Graduak ikasleak izan ditugu *Ikus-entzunezko Sormen Teknikak* irakasgai. Hauek landutako oinarritzko gaitasunak honako hauek dira: ikus-entzunezko proiektuak garatzeko eta diseinatzekeo baliabideak ezagutzea eta behar den teknologia aplikatzea; eta beste hauek zeharkako gaitasunak: talde-lana eta berritzapena.

Hona hemen proiektu honetan eman diren urratsak:

- 1) Erronka: zientzietako eta hezkuntzako ikasleek minutu bateko testua sortzea; ikus-entzunezkoek testua egin eta bideoa sortzea.
- 2) Brainstorming gai posibleekin.

- 3) Ikuspegi anitzak landzea, proiektua egin ahal izateko ikerketa-eremuen arteko loturak aztertu dira.
- 4) Ikertu eta berrikustea: ikerketa- eta berrikuste-lanak egitea, eztabaidatzekoa eskolara egokitzeko.
- 5) Trebetasunak erakustea: bideoak grabatzea.
- 6) Inplementazioa, konponbidea argitaratzea: 17/18 ikasturtean egin da, 4. mailako Lehen Hezkuntzako ikasleek bideoak erakusteko plataforma sortu.

Behin proiektu honen erronka lortuta, Google Drive plataformaren bidez diseinatu eta balidatutako EMSI galdetegia pasatu zitzaizen ikasleei, eta plataforma beraren bidez egin zen datu-bilketa.

Galdetegi honetan *Zergatik hartzen duzu parte jarduera honetan?* galderari erantzuten dioten itemak hiru motibazio motarekin daude erlazionatuta: motibazio eza, motibazio estrinsekoa eta motibazio intrinsekoa.

#### 4. Emaitzak

Galdetegi honekin lortutako emaitzak IBM SPSS Statistics programaren bidez aztertu dira. 1. taulan 16 itemen emaitzak ditugu: batezbestekoa, desbideratze estandarra eta kurtosia.

1., 5., 9. eta 13. itemak motibazio intrinsekoari dagozkio; 2., 6., 10. eta 14. itemak identifikatutako erregulaziodun motibazio estrinsekoari; 3., 7., 11., 15. itemak kanpoko erregulaziodun motibazio estrinsekoari; eta azkenik, 4., 8., 12. eta 16. itemak motibazio ezari.

| <b>¿Por qué participas actualmente en esta actividad?</b>                                    | <b>Batez bestekoa</b> | <b>Desbideratze estandarra</b> | <b>Kurtosia</b> |
|--|-----------------------|--------------------------------|-----------------|
| <b>1. Porque creo que esta actividad es interesante.</b>                                     | 4,51                  | 1,469                          | -0,009          |
| <b>2. Porque se supone que debo hacerlo.</b>   | 5,17                  | 1,519                          | -0,025          |
| <b>3. Por mi propio bien</b>   | 4,76                  | 1,454                          | -0,49           |
| <b>4. Puede que haya buenas razones para realizar esta actividad, pero yo no veo ninguna</b> | 3,19                  | 1,624                          | -0,475          |
| <b>5. Porque disfruto con esta actividad</b>   | 4,06                  | 1,47                           | -0,503          |
| <b>6. Porque creo que esta actividad es buena para mí</b>                                    | 4,61                  | 1,46                           | -0,435          |
| <b>7. Porque es algo que tengo que hacer</b>   | 5,26                  | 1,454                          | 0,27            |



|  |      |       |        |
|--|------|-------|--------|
| <b>8. Realizo esta actividad, pero no estoy seguro de si vale la pena</b>                      | 3,56 | 1,495 | -0,46  |
| <b>9. Porque esta actividad es divertida</b>   | 3,84 | 1,445 | -0,419 |
| <b>10. Por decisión personal</b>   | 3,48 | 2,024 | -1,072 |
| <b>11. Porque no tengo otra alternativa</b>  | 4,55 | 1,862 | -0,997 |
| <b>12. No lo sé; no veo qué me aporta esta actividad</b>                                       | 3,39 | 1,731 | -0,506 |
| <b>13. Porque me siento bien realizando esta actividad</b>                                     | 4,1  | 1,413 | -0,192 |
| <b>14. Porque creo que esta actividad es importante para mí</b>                                | 4,19 | 1,593 | -0,726 |
| <b>15. Porque creo que tengo que hacerlo</b>   | 4,93 | 1,441 | -0,006 |
| <b>16. Hago esta actividad, pero no estoy seguro de que sea conveniente continuar con ella</b> | 3,05 | 1,591 | 0,031  |

1. taula: EMSI galdetegiaren emaitzak

Batezbestekoari erreparatzen badiogu, emaitzarik altuenak 2., 7. eta 15. itemek dituzte; item hauek guztiak motibazio estrinsekoari dagozkio: 2.a, *porque se supone que debo hacerlo*, identifikatutako erregulazioari, eta 7.a, *porque es algo que tengo que hacer*, eta 15.a, *porque creo que tengo que hacerlo*, kanpoko erregulazioari. Emaitzarik baxuenak, berriz, 4. itemak (*puede que haya buenas razones para realizar esta actividad, pero yo no veo ninguna*), 12.ak (*no lo sé; no veo qué me aporta esta actividad*) eta 16.ak (*hago esta actividad, pero no estoy seguro de que sea conveniente continuar con ella*); hirurak motibazio ezari dagozkio.

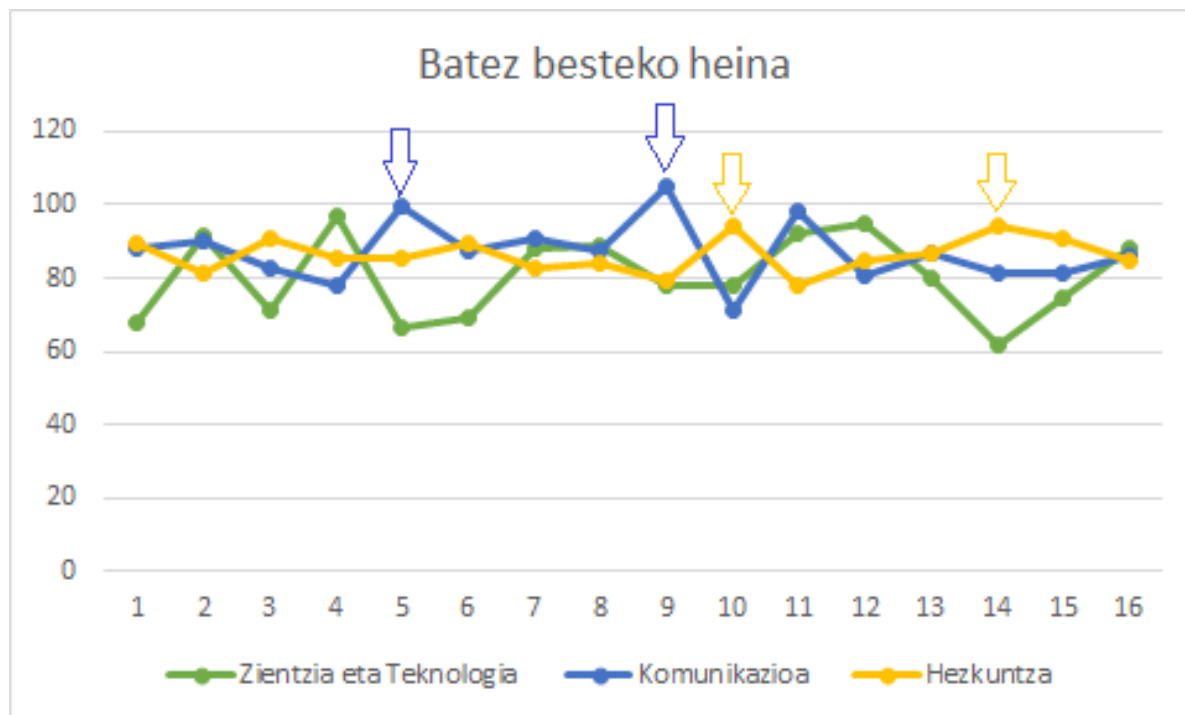
Kontuan hartu ditugun hiru aldagaien (generoa, ama-hizkuntza eta ikasketak) arabera ere aztertu ditugu emaitzak. Azterketa horretan, hasteko, normalitate-probak egin ditugu. Laginaren tamaina 50 datu baino handiagoa izan denez, Kolmogorov-Smirnov proba egin dugu. Kolmogorov-Smirnov probaren emaitzek banaketa normala edo banaketa asimetrikoa erakuts dezakete. Hiru aldagaien kasuan banaketa asimetrikoa eman du proba honek. Hortaz, proba ezberdinak egin dira datu multzoen arabera. Bi datu-multzorekin, Mann-Whitneyren U proba egin da, kasu honetan generoa eta ama-hizkuntza aztertzeko, eta bi datu-multzo baino gehiagorekin, Kruskal-Wallis proba egin da, kasu honetan ikasketak aztertzeko; izan ere, hiru gradutako (Zientzia eta Teknologia, Hezkuntza eta Gizarte eta Komunikazio Zientzietako) ikasleen motibazioa aztertu baita.

*Generoa eta ama-hizkuntza* aldagaiekin egindako Mann-Whitneyren U probaren bidez lortutako emaitzak ez dira esanguratsuak izan. Aldiz, ikasketekin egindako Kruskal-Wallis probaren bidez lortutako emaitzak esanguratsuak izan dira, lau galderatan (2. taula eta 3. irudia): 5., 9., 10. eta 14. galderetan, alegia.

| <b>Heinak</b> |                         |     |                      |
|---------------|-------------------------|-----|----------------------|
|               | Fakultatea              | N   | Batez besteko heinak |
| I1            | Zientzia eta Teknologia | 29  | 67,50                |
|               | Komunikazioa            | 42  | 88,02                |
|               | Hezkuntza               | 99  | 89,70                |
|               | Guztira                 | 170 |                      |
| I2            | Zientzia eta Teknologia | 29  | 91,84                |
|               | Komunikazioa            | 42  | 90,14                |
|               | Hezkuntza               | 99  | 81,67                |
|               | Guztira                 | 170 |                      |
| I3            | Zientzia eta Teknologia | 29  | 71,14                |
|               | Komunikazioa            | 42  | 82,57                |
|               | Hezkuntza               | 99  | 90,95                |
|               | Guztira                 | 170 |                      |
| I4            | Zientzia eta Teknologia | 29  | 96,98                |
|               | Komunikazioa            | 42  | 77,75                |
|               | Hezkuntza               | 99  | 85,42                |
|               | Guztira                 | 170 |                      |
| I5            | Zientzia eta Teknologia | 29  | 66,24                |
|               | Komunikazioa            | 42  | 99,49                |
|               | Hezkuntza               | 99  | 85,21                |
|               | Guztira                 | 170 |                      |
| I6            | Zientzia eta Teknologia | 29  | 69,17                |
|               | Komunikazioa            | 42  | 87,31                |
|               | Hezkuntza               | 99  | 89,52                |
|               | Guztira                 | 170 |                      |
| I7            | Zientzia eta Teknologia | 29  | 88,36                |
|               | Komunikazioa            | 42  | 90,62                |
|               | Hezkuntza               | 99  | 82,49                |
|               | Guztira                 | 170 |                      |
| I8            | Zientzia eta Teknologia | 29  | 88,60                |
|               | Komunikazioa            | 42  | 87,48                |
|               | Hezkuntza               | 99  | 83,75                |
|               | Guztira                 | 170 |                      |
| I9            | Zientzia eta Teknologia | 29  | 78,07                |
|               | Komunikazioa            | 42  | 104,73               |
|               | Hezkuntza               | 99  | 79,52                |
|               | Guztira                 | 170 |                      |

|     |                         |     |       |
|-----|-------------------------|-----|-------|
| I10 | Zientzia eta Teknologia | 29  | 77,64 |
|     | Komunikazioa            | 42  | 71,01 |
|     | Hezkuntza               | 99  | 93,95 |
|     | Guztira                 | 170 |       |
| I11 | Zientzia eta Teknologia | 29  | 92,10 |
|     | Komunikazioa            | 42  | 98,49 |
|     | Hezkuntza               | 99  | 78,06 |
|     | Guztira                 | 170 |       |
| I12 | Zientzia eta Teknologia | 29  | 94,72 |
|     | Komunikazioa            | 42  | 81,00 |
|     | Hezkuntza               | 99  | 84,71 |
|     | Guztira                 | 170 |       |
| I13 | Zientzia eta Teknologia | 29  | 80,17 |
|     | Komunikazioa            | 42  | 86,83 |
|     | Hezkuntza               | 99  | 86,49 |
|     | Guztira                 | 170 |       |
| I14 | Zientzia eta Teknologia | 29  | 61,81 |
|     | Komunikazioa            | 42  | 81,31 |
|     | Hezkuntza               | 99  | 94,22 |
|     | Guztira                 | 170 |       |
| I15 | Zientzia eta Teknologia | 29  | 74,72 |
|     | Komunikazioa            | 42  | 81,14 |
|     | Hezkuntza               | 99  | 90,51 |
|     | Guztira                 | 170 |       |
| I16 | Zientzia eta Teknologia | 29  | 87,81 |
|     | Komunikazioa            | 42  | 86,02 |
|     | Hezkuntza               | 99  | 84,60 |
|     | Guztira                 | 170 |       |

**2. taula:** Kruskal-Wallis probaren batez besteko heinak



**3. irudia:** Kruskal-Wallis proba ikasketen arabera

5. itemean (*porque disfruto con esta actividad*), argi ikusten da emaitzarik altuena Gizarte eta Komunikazioko ikasleek eman diotela; aldiz, Zientzia eta Teknologiakoek baxuena.

9. itemean (*porque esta actividad es divertida*), berriro ikusten da notarik altuena Gizarte eta Komunikazioko ikasleek eman diotela eta, baxuena, berriz, Zientzia eta Teknologiakoek.

10. itemean (*por decisión personal*), Hezkuntzako ikasleek eman diote notarik altuena eta Gizarte eta Komunikaziokoek baxuena.

Amaitzeko, 14. itemean (*porque creo que esta actividad es importante para mí*), Hezkuntzako ikasleek nota altua eman diote eta Zientzia eta Teknologiakoek, ordea, baxua.

## 5. Ondorioak

Motibazioari dagokionez, balidatutako EMSI galdetegiaren batezbestekoen emaitzek ikasleek amotibazio baxua izan dutela erakusten dute; honi dagozkion itemak lau dira, eta horietako hiruk lortu dituzte batezbesteko baxuenak. Hortaz, ondoriozta daiteke motibazioa ez dela izan baxua.

Batetik, motibazio eragilerik handienak zeintzuk izan diren ikusteko, batezbesteko altuenei erreparatuko diegu. 1. taulako datuetan ikusten denez, motibazio estrinsekoa da emaitza altuenak lortu dituenak. Puntuazioa kontuan hartuta, lehenengo bost itemetatik hiru kanpoko erregulazioari dagozkio, eta beste biak, identifikatutako erregulazioari. Beraz, esan dezakegu proiektu honetan parte hartu duten ikasleak motibatuta egon direla eta motibaziorik altuena estrinsekoa izan dela, hau da, ikasleentzat erakargarria izan dela proiektu honetan parte hartu izana.

Beste alde batetik, generoari eta ama-hizkuntzari dagozkien datu estatistikoek emaitzek adierazten dute aldagai hauek ez direla esanguratsuak. *Ikasketa* aldagaiari dagokionez, berriz, batez besteko heinak esanguratsuak izan dira lau itemetan (5.a: *porque disfruto con esta actividad*; 9.a: *porque esta actividad es divertida*; 10.a: *por decisión personal*, eta 14.a: *porque creo que esta actividad es importante para mí*). Hargatik, baieztatu daiteke, batetik, Zientzia eta Teknologiako ikasleentzat erronka ez dela garrantzitsua izan. Bestetik, motibatuenak Gizarte eta Komunikazioko ikasleak izan direla eta, azkenik, erreal, esanguratsu eta beraien etorkizun profesionalarekin erlazionaturik dagoen erronka bezala Hezkuntzako ikasleek hauteman dutela.

## 6. Bibliografia

- Cordray, D.S., Harris, T.R. & Klein, S. (2009). A research synthesis of the effectiveness, replicability, and generality of the VaNTH challenge-based instructional modules in bioengineering. *Journal of Engineering Education* 98, 335-348.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Evans, M. J. & Moore, J. S. (2013). Peer tutoring with the aid of the Internet. *British Journal of Educational Technology*, 44(1), 144-155.
- Guiza, M., Rivera, R., Lamadrid, L. & Aguilar, W. (2016). Podcast educativo como herramienta en el proceso de enseñanza/aprendizaje: un estudio de caso universitario. In Roig-Vila, R. (Eds.) *Tecnología, innovación e investigación en los procesos de enseñanza-aprendizaje*. Barcelona: Octaedro.
- Malmqvist, J., Rådberg, K. K., & Lundqvist, U. (2015). Comparative analysis of challenge-based learning experiences. Proceedings of the 11th International CDIO Conference, Chengdu University of Information Technology, Chengdu, Sichuan, P.R. China.
- Martín-Albo, J., Núñez, J.L. & Navarro, J.G. (2009). Validation of the Spanish version of the Situational Motivation Scale (EMSI) in the educational context. *The Spanish journal psychology* 12(2), 799-807.
- Miller, M., & Piller, M. (2005). Principal factors of an audio reading delivery mechanism: Evaluating educational use of the iPod. Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications, Chesapeake, VA.
- Piñeiro, T. & Caldevilla, D. (2011). Podcasting didáctico: una aproximación a su uso en el ámbito didáctico de la universidad española. *Sapiens: Revista Universitaria de Investigación* 12(2), 14-30.

- Piñeiro, T. & Costa, C. (2011). Potencialidades del podcast como herramienta educativa para la enseñanza universitaria. *Ética net* 15(2).
- Ryan, R. M. & Deci, E. L. (1985). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist* 55(1), 68-78.
- Schnackenberg, H.L., Vega, E.S. & Dustin, H. (2009). Podcasting and vodcasting: legal issues and ethical dilemmas. *Journal of Law, Ethics, and Intellectual Property* 3(1), 361-367.

## GAZTE EUSKALDUNEN BAI/EZ GALDEREN INTONAZIOAZ: EUSKARAZ ETA GAZTELANIAZ

UPV/EHU

Naia Eguskiza, Irati de Pablo eta Iñaki Gaminde

[naia.egusquiza@ehu.eus](mailto:naia.egusquiza@ehu.eus), [irati.depablo@ehu.eus](mailto:irati.depablo@ehu.eus), [inaki.gaminde@gmail.com](mailto:inaki.gaminde@gmail.com)

**Laburpena:** Euskararen ikaskuntza-irakaskuntza prozesuetan intonazioa ez da gehiegi sakontzen den gaia. Hutsune hori asebetetzeko, hemen aurkezten den lan honen helburu nagusia da aztertzea euskaldun gazte elebidun goiztiarrek bai/ez galderen intonazioaren tasun batzuk desberdin egiten dituzten erabiltzen duten hizkuntzaren arabera, dela euskaraz dela gaztelaniaz. Horrez gain, erabileretan ager dakizkigukeen aldeetan informatzaileen ama-hizkuntzaren eragina zenbatekoa den aztertu da eta, bestetik, agerian jarri diren tasunen arabera informatzaileen sailkapena egingo da; horren helburua datuen azpian dagokeen egituraketa soziolinguistikoa azalera da, sistema konplexu dinamiko moldagarrien ildotik, gure kasuan ama- hizkuntzaren arabera balizko azpimultzoak azalera. Hortaz, datuak aztertzean, aldagai bi erabiliko ditugu: esaldiak egiten direneko hizkuntza eta informatzaileen ama-hizkuntza. Horretarako, 160 esaldiko corpusa osatu da 40 andrazko informatzaileen kolaborazioarekin.

**Gako hitzak:** intonazioa, bariazio linguistikoa, fonetika akustikoa, informatzaileen sailkapena

### 1. Sarrera

Beste batzuetan aldarrikatu den bezala (Gaminde et. al., 2014), euskararen ikaskuntza-irakaskuntza prozesuetan intonazioa ez da gehiegi sakontzen den gaia; irakasleok sarri askotan zer landu behar den ez dakigu, are gutxiago zelan landu behar den. Bestalde, gizarte elebiduna omen den gure honetan ez da ikertzaile trebeegia izan behar jabetzeko euskararen egoera gaztelaniarenaren aldean, Hegoaldean, ez dela parekidea. Egoera honetan, euskarari eusteko indar egin behar izaten da esparru guztietan eta euskaldunen borondateari esker eusten zaio, nekez bada ere.

Intonazioa landu ezak bultzatu gaitu lantxo honen burutzerara. Egia da gaztelaniaz ere ez dugula ezagutzen metodologiarik edo ahalegin berezirik intonazioa lantzeko; baina egoera soziolinguistikoa gaztelaniaren aldekoa izanik, ikusi nahi izan dugu hizkuntza bietan elebidunak direnek alde handia daukaten ala ez.

Gure lana mugatzeko, esaldi mota baten intonazioa hartu dugu aztergaitzat; aukeratu duguna bai/ez galdera edo galdera absolutua izan da.

Beraz, gure lan honen helburu nagusia da aztertzea euskaldun gazte elebidun goiztiarrek bai/ez galderen intonazioaren tasun batzuk desberdin egiten dituzten erabiltzen duten hizkuntzaren arabera.

Helburu hau lortzeko, esaldien deskripzio fonetikoetan oinarrituko gara. Gaztelaniaren kasuan, eredu metriko autosegmentala erabil genezake, eta deskripzio fonologikoan oinarritu: hau izan da aspaldion hizkuntza honetan gehienetan erabili den moldea (Prieto, P. eta Roseano, P. (eds), 2010; Estebas, E. eta Prieto, P., 2009; Face, T. eta Prieto, P., 2007; besteak beste); euskararen barietate batzuetan egin diren monografia zenbaitetan eredu metriko autosegmentala erabili da (Gaminde, 2010a, Gaminde et al., 1997, 2012; Elordieta, 1999, 2000, 2003, 2007a, 2007b; Elordieta et al., 1999, Ito 2002a, 2002b, 2003; Jun eta Elordieta, 1997; Legarra, 2011, besteak beste). Hala ere, euskararen azentuaren kokagunea zehaztea oso konplexua, eta nekeza denez gero (Hualde, 1997, 2006, Gaminde 2007, 2011), Gamindek et al.ek (2017) eta Etxebarriak eta Eguskizak (2018) proposatutako ildotik, azterketa fonetikoan oinarritzea erabaki dugu.

Gorago aipatutako helburu orokor horrezaz gain, beste helburu zehatz bi ere hartu ditugu; batetik, erabileretan ager dakizkigukeen aldeetan informatzaileen ama-hizkuntzaren eragina zenbatekoa den aztertuko dugu eta, bestetik, agerian jarriko ditugun tasunen arabera informatzaileen sailkapena egingo dugu; beronen helburua datuen azpian dagokeen egituraketa soziolinguistikoa azaleratzea da, sistema konplexu dinamiko moldagarrien ildotik (Clua et. al., 2017; Eguskiza et. al., 2017), gure kasuan ama-hizkuntzaren arabera balizko azpimultzoak azaleratuz. Hortaz, datuak aztertzea aldagai bi erabiliko ditugu: esaldiak egiten direneko hizkuntza eta informatzaileen ama hizkuntza.

Badago lehendik hizkuntza biak erkatzen dituen lan bat (Gaminde, 2010a), haren emaitzak ezin ekar ditzakegu hona, han aztertzeke erabili zen metodologia ezberdina izan zen eta.

Lana bost ataletan banatuta aurkezten dugu. Sarrera honen osteko bigarren atalean, corpusaren ezaugarriak eta hori jasotzeko eta prozesatzeko metodologia aurkezten dira; hirugarren atalean, intonazioaren azterketa egiten da; laugarrenean, informatzaileen sailkapena egiten da, eta, azkenik, bosgarren atalean, ondorio nagusiak laburbiltzen dira.

## 2. Corpora eta Metodologia

Aurreko atalean zehaztu ditugun helburuak lortzeko, 160 esaldiko corpora osatu dugu 40 informatzailerekin. Informatzaile guztiak andrazkoak izan dira, 1996-1999 urte artean jaiotakoak; euren adinaren batezbestekoa 19,08 urtekoa izan da, D ereduan burutu dituzte euren ikasketa guztiak, eta Bilboko Hezkuntza Fakultateko ikasleak izan dira inkestak egitean. Euretako 20ren ama-hizkuntza euskara izan da (A taldea) eta beste 20ena (B taldea) gaztelania; hemen, beste lan batzuetan egindako sailkapen berberari jarraituko diogu (Gaminde, 2010a, 2010b). Denak elebidun goiztiarrak izan dira: A taldekoek gaztelania ikasi dute hezkuntzaren bidez, eta B taldekoek, berriz, euskara. 1. taulan informatzaileen zerrenda ematen dugu euren ama-hizkuntzaren arabera sailkatuta. Informatzaileen identifikaziorako parentesien arteko kodearekin batera jatorrizko herriaren izena erabiltzen da; herriren baten informatzaile bat baino gehiago izan dugunean, zenbaki bat paratu dugu.



| Taldea | Informatzaileak   |
|--------|---|
| A      | Mungia1 (A01), Mungia2 (A02), Gamiz (A03), Sondika (A04), Zamudio (A05), Larrabetzu (A06), Bermeo (A07), Gernika (A08), Muxika (A09), Lekeitio1 (A10), Lekeitio2 (A11), Aulesti (A12), Markina1 (A13), Markina2 (A14), Elorrio (A15), Abadiño (A16), Zornotza1 (A17), Zornotza2 (A18), Igorre1 (A19) eta Igorre2 (A20)  |
| B      | Abanto (B01), Trapagaran (B02), Santurtzi (B03), Portugalete (B04), Bilbo1 (B05), Bilbo2 (B06) Bilbo (B07), Basauri1 (B08), Basauri2 (B09), Etxebarri (B10), Arrigorriaga (B11), Galdakao (B12), Sondika (B13), Leioa (B14), Getxo (B15), Sopela1 (B16), Sopela2 (B17), MungiaB1 (B18), Mungia2 (B19) eta Gernika (B20) |

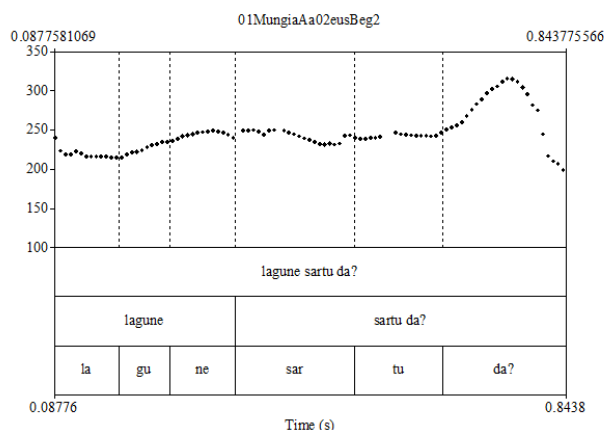
1. taula: informatzaileen zerrenda, euren ama-hizkuntzaren arabera sailkatua.

Informatzaile bakoitzeko, esaldi bat jaso da euskaraz (a) eta beste bat gaztelaniaz (b), eta informatzaileek esaldi bakoitza bi aldiz esan zuten; beraz, lau esaldi jaso dira informatzaileko.

- (a) Laguna sartu da?
- (b) ¿La nena llevaba?

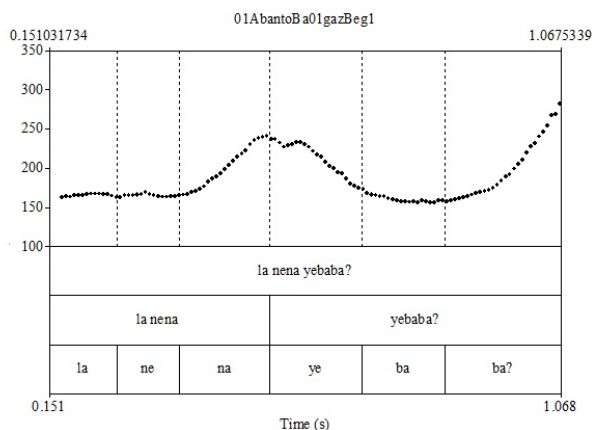
(a) eta (b) esaldiek osagai bi dauzkate, eta osagai bakoitzaren silaba kopurua bera da hizkuntza bietan. Euskararen kasuan, A taldeko informatzaileek euren barietateetako arauak eta aldakiak erabili dituzte; hala ere, honek ez dio eragin silaba kopuruari.

Lan honetako helburuetarako, esaldien bigarren osagaiaren  $f_0$ -aren azterketa hartu dugu oinarritzat; esaldiak transkribatzeko, etiketatzeko eta aztertzeko, “praat” izeneko programa erabili dugu (Boersma eta Weenink, 2016). 1. irudiko kurbaren bigarren zatia irreparatzen badiogu, ikusten dugu aditzaren lehen silaban beheranzko tonua agertzen dela, bigarren silaban goranzkoa eta hirugarrenean goranzko/beheranzkoa.



1. irudia: euskarazko esaldia (A taldeko Mungiako informatzailea)

2. irudiko esaldiaren aditzaren lehen silaban beheranzko tonua ikusten dugu; bigarren silaban ere beheranzkoa dugu, eta hirugarrenean, goranzkoa.



**2. irudia:** gaztelaniazko esaldia (B taldeko Abantoko informatzailea)

Horrela bada, kurba guztietako silaben tonuen gorabeherak aztertu ondoren, ikus dezakegu zeintzuk diren silaba bakoitzeko erabiltzen diren tonu motak eta baita agertzen diren tonu konbinazio nagusiak ere.

Informatzaileak sailkatzeko eta balizko egitura soziala azaleratzeko, lau metodo erabiliko ditugu. Lehenengo eta behin, esaldien bigarren zatiaren silaba guztietako tonuen norabideak aztertuta, matrize bat eraiki dugu hizkuntza bakoitzeko silaba bakoitzean egon daitezkeen tonu motekin. Matrize horretan, informatzaile bakoitzaren datuen arabera emaitzak sartu ditugu. Emaitzak 0 eta 1 bitartean kokatu dira; izan ere, informatzaile bakoitzak esaldi bi egin ditu. Orduan, esaldi bietan tonu bera erabili duenean, 1 jarri da; batean erabili duenean, 0,5, eta tonu motarik erabiltzen ez duenean, 0 (3. irudiko taula).

| Infor | .E1g | .E1b | .E2g | .E2b | .E3gb | .E3g | .E3b | .E3bg | .G1g | .G1b | .G2g |
|-------|------|------|------|------|-------|------|------|-------|------|------|------|
| A01   | 1    | 0    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A02   | 0    | 1    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A03   | 0    | 1    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A04   | 1    | 0    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A05   | 1    | 0    | 1    | 0    | 0     | 1    | 0    | 0     | 0    | 0    | 1    |
| A06   | 1    | 0    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A07   | 0    | 1    | 1    | 0    | 0     | 0    | 1    | 0     | 0    | 0    | 1    |
| A08   | 0    | 1    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A09   | 0    | 1    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A10   | 0,5  | 0,5  | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A11   | 0    | 1    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A12   | 0    | 1    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A13   | 1    | 0    | 0    | 1    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A14   | 1    | 0    | 1    | 0    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A15   | 1    | 0    | 1    | 0    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A16   | 0    | 1    | 1    | 0    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A17   | 0    | 1    | 1    | 0    | 1     | 0    | 0    | 0     | 0    | 0    | 1    |
| A18   | 1    | 0    | 0,5  | 0,5  | 0,5   | 0    | 0,5  | 0     | 0    | 0    | 1    |
| A19   | 0    | 1    | 0    | 1    | 0     | 1    | 0    | 0     | 0    | 0    | 1    |

**3. irudia:** datu-matrizearen zatia.

Matrizea eginda, informatzaileak sailkatzeko prozedura hautatu behar da. Multzokatze-azterketetan mota bi bereizten dira gehienetan (Pardo eta Ruiz, 2002, Pérez López, 2005): multzokatze hierarkikoak eta ez-hierarkikoak. Multzokatze ez-hierarkikoetan, edo K-batezbestekoetan (KM), aldeaz aurretik ikertzaileak multzo kopurua finkatu behar du. Hauek gain, badago logika lausoan oinarritzen den multzokatzea ere (fuzzy C-means, FCM).

Multzokatze-metodoa aukeratu ondoren, kasuen artean erkaketak egin ahal izateko, antzekotasunen eta desberdintasunen distantzia-neurria aukeratu behar da (Pardo eta Ruiz, 2002, Pérez López, 2005, Goebel, 1992, 2010; Aurrekoetxea, 1995, 2005; Clua, 2010); gure lantxo honetan, kasu guztietan distantzia euklidearra aukeratu dugu. Multzokatze

hierarkikoa egiteko aukeratu dugun sailkatze-metodoa Warden a izan da (Goebel, 1992, 2010; Aurrekoetxea, 1995, 2005; Gaminde, 2007; Gaminde et al., 2016).

KM eta FCM multzokatze mota erabiltzen baditugu, gauza bera gertatuko zaigu, baina, oraingo honetan, elementuen banaketa egingo da guk alde aurretik aukeratu dugun multzo kopuruaren arabera. Teknika honek elementuen arteko berdintasunak eta multzoen arteko aldeak nabarmenarazten ditu.

Azkenik, dimentsio aniztun eskala izeneko teknika estatistikoa ere erabili dugu informatzaileen artean azpimultzorik egin daitekeen ikusteko.

### 3. Intonazioaren azterketa

Atal honetan, esaldien bigarren osagaiaren  $f_0$ -aren azterketa egingo dugu hizkuntza bietan; datuak argiago ikusteko, banan aurkeztuko ditugu. Horrela, bada, lehen azpiatalean, euskararen gaineko datuak emango ditugu; bigarren azpiatalean, gaztelaniaren gainekoak, eta hirugarrenean, bien arteko erkaketa egingo dugu.

#### 3.1. Euskararen datuak

Euskararen esaldietan lehen silaban gehien erabiltzen den tonua beheranzkoa da (45, % 56,25) eta gero goranzkoa (35, % 43,75). Informatzaileen ama-hizkuntzaren arabera banaketa 2. taulan ikus daiteke; bertan ikusten diren aldeak oso txikiak dira eta ez dira estatistikoki esanguratsuak.

| Tonua       | A taldea |      | B taldea |    |
|-------------|----------|------|----------|----|
|             | kopurua  | %    | kopurua  | %  |
| beheranzkoa | 23       | 57,5 | 22       | 55 |
| goranzkoa   | 17       | 42,5 | 18       | 45 |

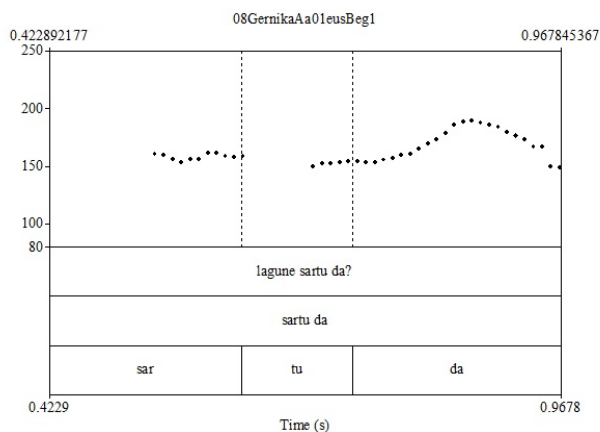
**2. taula:** lehen silabako tonu moten kopuruak eta ehunekoak informatzaileen ama-hizkuntzaren arabera.

Bigarren silaban gehien erabiltzen den tonua beheranzkoa da (65, % 81,25) eta gero goranzkoa (15, % 18,75). Datuak informatzaileen ama-hizkuntzaren arabera banatuta erakusten dira 3. taulan. Bertan ikusten den bezala, B taldeko informatzaileen beheranzko tonuen erabilera handiagoa da A taldeko informatzaileena baino. Alde hau khi karratu probaren arabera estatistikoki esanguratsua da ( $\chi^2 = (a.m.: 1) 9,928, p = 0,002$ ).

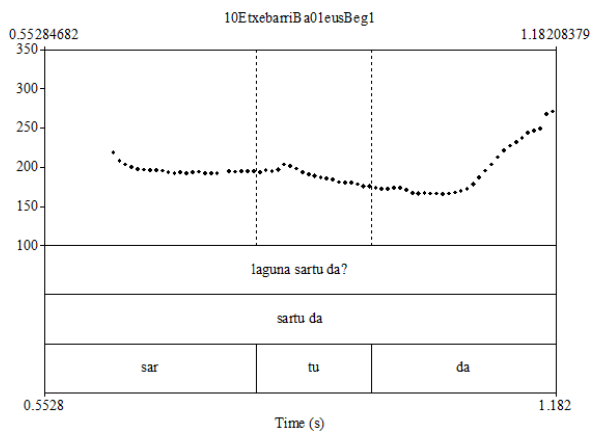
| Tonua       | A taldea |      | B taldea |    |
|-------------|----------|------|----------|----|
|             | kopurua  | %    | kopurua  | %  |
| beheranzkoa | 27       | 67,5 | 38       | 95 |
| goranzkoa   | 13       | 32,5 | 2        | 5  |

**3. taula:** bigarren silabako tonu moten kopuruak eta ehunekoak informatzaileen ama-hizkuntzaren arabera.

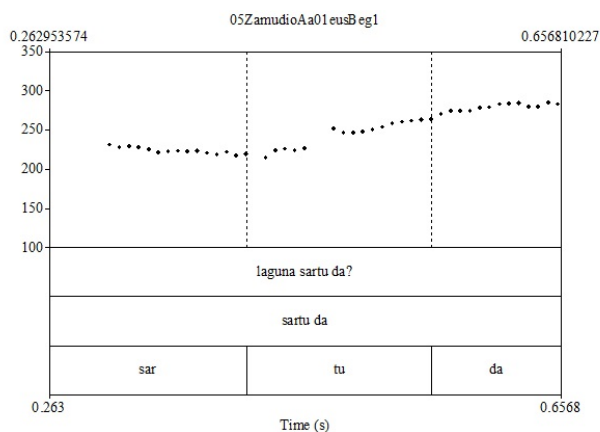
Hirugarren silaban lau tonu mota agertzen dira: goranzko/beheranzkoa (4. irudia), beheranzko/goranzkoa (5. irudia), goranzkoa (6. irudia) eta beheranzkoa (7. irudia).



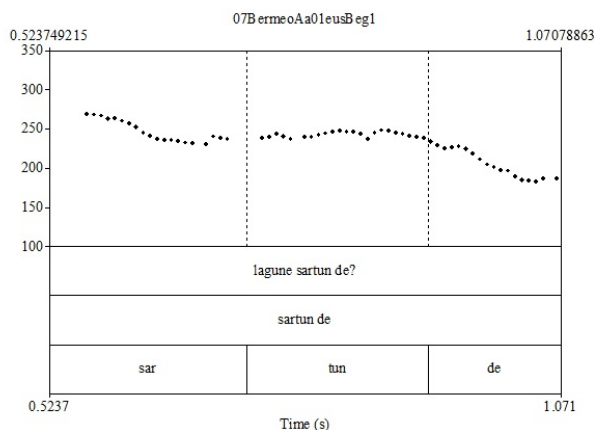
**4. irudia:** goranzko/beheranzko tonua azken silaban (A taldeko Gernikako informatzailea).



**5. irudia:** beheranzko/goranzko tonua azken silaban (B taldeko Etxebarriko informatzailea).



**6. irudia:** goranzko tonua azken silaban (A taldeko Zamudioko informatzailea).



7. irudia: beheranzko tonua azken silaban (A taldeko Bermeoko informatzailea).

Gehien erabiltzen dena goranzko/beheranzkoa (gb) da (52, % 65); gero, beheranzko/goranzkoa (bg) (12, % 15) eta goranzkoa (g) (11, % 13,75), eta gutxien, beheranzkoa (b) (5, % 6,25). Euron banaketa, informatzaileen ama-hizkuntzaren arabera, 4. taulan erakusten da. Bertan ikus daitekeen moduan, alderik handiena g/b eta b/g tonuen erabileran gertatzen da; aldeak estatistikoki esanguratsuak dira khi karratu probaren arabera ( $\chi^2 = (a.m.: 3) 14,214, p = 0,003$ ).

| Tonua | A taldea |      | B taldea |      |
|-------|----------|------|----------|------|
|       | kopurua  | %    | kopurua  | %    |
| gb    | 31       | 77,5 | 21       | 52,5 |
| bg    | 0        | 0    | 12       | 30   |
| g     | 6        | 15   | 5        | 12,5 |
| b     | 3        | 7,5  | 2        | 5    |

4. taula: Hirugarren silabako tonu moten kopuruak eta ehunekoak informatzaileen ama- hizkuntzaren arabera.

Silaba bakoitzeko agertzen diren tonuen konbinazio posibleak 16 dira; horietarik 10 erabili dira. 5. taulan tonu konbinazio bakoitzaren kopuruak eta ehunekoak ematen dira.

| 1. silaba      | 2. silaba | 3. silaba | kopurua | %     |
|----------------|-----------|-----------|---------|-------|
| g              | b         | gb        | 24      | 30    |
| b              | b         | gb        | 20      | 25    |
| b              | b         | bg        | 10      | 12,5  |
| b              | b         | g         | 9       | 11,25 |
| g              | g         | gb        | 4       | 5     |
| b              | g         | gb        | 4       | 5     |
| g              | g         | b         | 3       | 3,75  |
| g              | b         | bg        | 2       | 2,5   |
| g              | g         | g         | 2       | 2,5   |
| b              | g         | b         | 2       | 2,5   |
| <b>Guztira</b> |           |           | 80      | 100   |

5. taula: tonuen konbinazioen kopuruak eta ehunekoak euskaraz.

Tonuen konbinazioak informatzaileen ama-hizkuntzaren arabera sailkatzen baditugu (6. taula), ikusten dugu B taldekoek tonu-konbinazio gutxiago erabiltzen dituztela (6) A taldekoek baino (8), erabileraren aldetik talde homogeneousagoa osatuz.

| 1. silaba      | 2. silaba | 3. silaba | A taldea |      | B taldea |      |
|----------------|-----------|-----------|----------|------|----------|------|
|                |           |           | kopurua  | %    | kopurua  | %    |
| b              | b         | g         | 0        | 0    | 10       | 25   |
| b              | b         | gb        | 4        | 10   | 5        | 12,5 |
| g              | b         | gb        | 13       | 32,5 | 7        | 17,5 |
| b              | g         | gb        | 2        | 5    | 0        | 0    |
| g              | g         | gb        | 4        | 10   | 0        | 0    |
| g              | g         | b         | 0        | 0    | 2        | 5    |
| b              | g         | b         | 10       | 25   | 14       | 35   |
| b              | g         | g         | 1        | 2,5  | 2        | 5    |
| g              | b         | bg        | 2        | 5    | 0        | 0    |
| g              | g         | g         | 4        | 10   | 0        | 0    |
| <b>Guztira</b> |           |           | 40       | 100  | 40       | 100  |

**6. taula:** Tonuen konbinazioen kopuruak eta ehunekoak euskaraz, informatzaileen ama-hizkuntzaren arabera.

### 3.2. Gaztelaniaren datuak

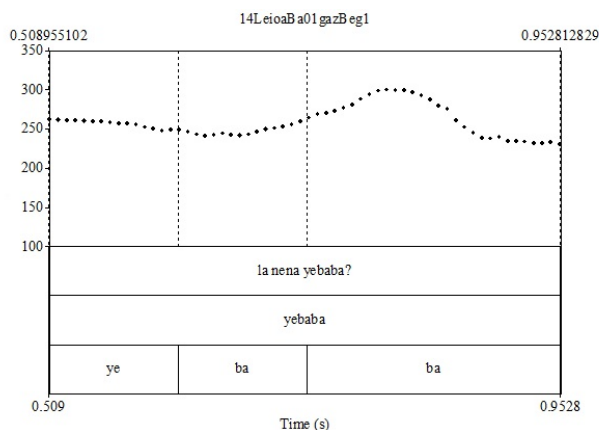
Gaztelaniaren esaldietan lehen silaban erabiltzen den tonu bakarra beheranzkoa da; honetan ez dago ezelako alderik informatzaileen artean.

Bigarren silaban gehien erabiltzen den tonua beheranzkoa da (62, % 77,5) eta gero goranzkoa (18, % 22,5). Datuak informatzaileen ama-hizkuntzaren arabera banatuta erakusten dira 7. taulan. Bertan ikusten den bezala, B taldeko informatzaileen beheranzko tonuen erabilera handiagoa da A taldeko informatzaileena baino; aldea ez da estatistikoki esanguratsua.

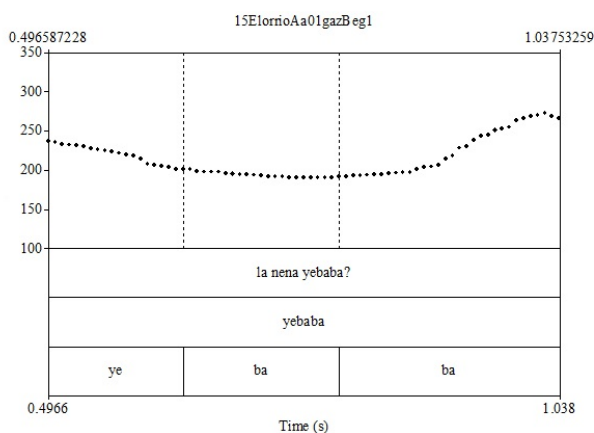
| Tonua       | A taldea |    | B taldea |    |
|-------------|----------|----|----------|----|
|             | kopurua  | %  | kopurua  | %  |
| beheranzkoa | 28       | 70 | 34       | 85 |
| goranzkoa   | 12       | 30 | 6        | 15 |

**7. taula:** bigarren silabako tonu moten kopuruak eta ehunekoak informatzaileen ama hizkuntzaren arabera.

Hirugarren silaban tonu mota bi agertzen dira: goranzko/beheranzkoa (8. irudia) eta goranzkoa (9. irudia).



**8. irudia:** goranzko/beheranzko tonua azken silaban (B taldeko Leioako informatzailea).



**9. irudia:** goranzko tonua azken silaban (A taldeko Elorrioko informatzailea).

Goranzko/beheranzkoa (gb) gutxiagotan erabiltzen da (21, % 26,25); goranzkoa (g) da gehien erabiltzen dena (59, % 73,75). Bibliografia gaingiroki arakatuta, goranzko tonua da normaltzat jotzen dena gaztelaniazko bai/ez galderetan (Navarro, 1980; Gili, 1978, Quilis, 1988, 1993; D’Introno et al., 1995; Hualde, 2003, Sosa, 2003; besteak beste).

Tonuen hedadura, informatzaileen ama-hizkuntzaren arabera, 8. taulan erakusten da. Bertan ikus daitekeen aldea ez da estatistikoki esanguratsua.

| Tonua | A taldea |      | B taldea |    |
|-------|----------|------|----------|----|
|       | kopurua  | %    | kopurua  | %  |
| gb    | 13       | 32,5 | 8        | 20 |
| g     | 27       | 67,5 | 32       | 80 |

**8. taula:** hirugarren silabako tonu moten kopuruak eta ehunekoak informatzaileen ama hizkuntzaren arabera.

Silaba bakoitzeko agertzen diren tonuen konbinazio posibleak 8 dira gaztelaniaz; horietarik, 4 erabili dira: b-b-gb (57, % 71,25), g-g-gb (16, % 20), g-b-gb (5, % 6,25) eta b-b-bg (2, % 2,5). Tonuen konbinazioak informatzaileen ama-hizkuntzaren arabera sailkatzen baditugu (9. taula), ikusten dugu B taldekoek tonu-konbinazio gutxiago erabiltzen dituztela (3) A taldekoek baino (4); honetan ere B taldea homogeneousagoa dela esan dezakegu.

| 1. silaba      | 2. silaba | 3. silaba | A taldea |      | B taldea |     |
|----------------|-----------|-----------|----------|------|----------|-----|
|                |           |           | kopurua  | %    | kopurua  | %   |
| b              | b         | gb        | 25       | 62,5 | 32       | 80  |
| g              | b         | gb        | 3        | 7,5  | 2        | 5   |
| b              | b         | bg        | 2        | 5    | 0        | 0   |
| g              | g         | gb        | 10       | 25   | 6        | 15  |
| <b>Guztira</b> |           |           | 40       | 100  | 40       | 100 |

**9. taula:** Ttonuen konbinazioen kopuruak eta ehunekoak euskaraz informatzaileen ama-hizkuntzaren arabera.

### 3.3. Hizkuntza bien erkaketa

Azken azpiatal honetan, hizkuntza bietan lortu ditugun emaitzak erkatuko ditugu aldeak agerian jartzekotan.

Lehen silabako tonua, euskaraz, beheranzkoa (45, % 56,25) eta goranzkoa (35, % 43,75) izan daiteke; gaztelaniaz, kasu guztietan beheranzkoa da (80, %100). Khi karratuaren probaren arabera, aldea estatistikoki esanguratsua da ( $\chi^2 =$  (a.m.: 1) 44,800,  $p = 0,000$ ). Hizkuntza bakoitzeko emaitzak erkatzen baditugu informatzaileen ama-hizkuntzaren arabera, 10. taulako datuak lortzen dira. A taldearen hizkuntzaren arabera aldea estatistikoki esanguratsua da ( $\chi^2 =$  (a.m.: 1) 21,587,  $p = 0,000$ ), gauza bera gertatzen da B taldearen aldeei dagokienez ( $\chi^2 =$  (a.m.: 1) 23,226,  $p = 0,000$ )

| Taldea | Hizkuntza  | Goranzkoa |      | Beheranzkoa |      |
|--------|------------|-----------|------|-------------|------|
|        |            | kopurua   | %    | kopurua     | %    |
| A      | euskara    | 17        | 42,5 | 23          | 57,5 |
|        | gaztelania | 0         | 0    | 40          | 100  |
| B      | euskara    | 18        | 45   | 22          | 55   |
|        | gaztelania | 0         | 0    | 40          | 100  |

**10. taula:** lehen silabako tonu moten kopuruak eta ehunekoak hizkuntza bakoitzeko informatzaileen ama-hizkuntzaren arabera.

Bigarren silabaren tonuari dagokionez, euskaraz, beheranzkoa (65, % 81,25) gehiagotan gertatzen da goranzkoa (15, % 18,75) baino; gaztelaniaz ere beheranzkoa (62, % 77,5) goranzkoa (18, % 22,5) baino gehiagotan erabiltzen da. Aldea ez da estatistikoki esanguratsua. Hizkuntza bakoitzeko datuak informatzaileen ama-hizkuntzaren arabera erakusten dira 11. taulan. Aldeak ez dira estatistikoki esanguratsua, ez A taldean ez eta B taldean ere.

| Taldea | Hizkuntza  | Goranzkoa |      | Beheranzkoa |      |
|--------|------------|-----------|------|-------------|------|
|        |            | kopurua   | %    | kopurua     | %    |
| A      | euskara    | 13        | 32,5 | 27          | 67,5 |
|        | gaztelania | 12        | 30   | 28          | 70   |
| B      | euskara    | 2         | 5    | 38          | 95   |
|        | gaztelania | 6         | 15   | 85          |      |

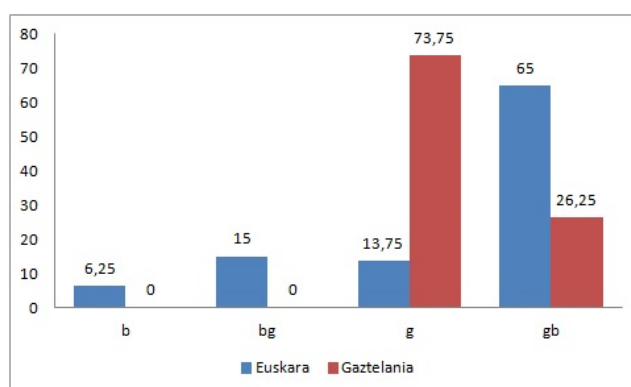
**11. taula:** Bigarren silabako tonu moten kopuruak eta ehunekoak hizkuntza bakoitzeko informatzaileen ama-hizkuntzaren arabera.



Hirugarren silabaren tonu motei dagokienez, lau aukera agertu izan zaizkigu: goranzko (g), goranzko/beheranzkoa (gb), beheranzkoa (b) eta beheranzko/goranzkoa (b/g). Euskaraz, laurak agertu dira, eta gaztelaniaz, ostera, “gb” eta “g” baino ez. 12. taulan mota bakoitzaren erabileraren kopuruak eta ehunekoak erakusten dira hizkuntza bakoitzeko. Bertan ikusten den bezala, gaztelaniaz gehien erabiltzen den muga-tonua “g” da eta euskaraz, ostera “gb” (10. irudiko grafikoa). Aldeak estatistikoki esanguratsuak dira ( $\chi^2 =$  (a.m.: 3) 63,079,  $p = 0,000$ )

|                   | <b>b</b><br>kopurua | %    | <b>bg</b><br>kopurua | %  | <b>g</b><br>kopurua | %     | <b>gb</b><br>kopurua | %     |
|-------------------|---------------------|------|----------------------|----|---------------------|-------|----------------------|-------|
| <b>Euskara</b>    | 5                   | 6,25 | 12                   | 15 | 11                  | 13,75 | 52                   | 65    |
| <b>Gaztelania</b> | 0                   | 0    | 0                    | 0  | 59                  | 73,75 | 21                   | 26,25 |

12. taula: lehen silabako tonu moten kopuruak eta ehunekoak hizkuntza bakoitzeko.



10. irudia: hirugarren silabako tonu mota bakoitzaren ehunekoak hizkuntzako.

Hizkuntza bakoitzeko datuak informatzaileen ama-hizkuntzaren arabera erakusten dira 13. taulan. Aldeak estatistikoki esanguratsuak dira A taldean ( $\chi^2 =$  (a.m.: 2) 23,727,  $p = 0,000$ ) eta B taldean ( $\chi^2 =$  (a.m.: 3) 39,530,  $p = 0,000$ ).

| Taldea   | Hizkuntza         | <b>b</b><br>kopurua | %   | <b>bg</b><br>kopurua | %  | <b>g</b><br>kopurua | %    | <b>gb</b><br>kopurua | %    |
|----------|-------------------|---------------------|-----|----------------------|----|---------------------|------|----------------------|------|
| <b>A</b> | <b>Euskara</b>    | 3                   | 7,5 | 0                    | 0  | 6                   | 15   | 31                   | 77,5 |
|          | <b>Gaztelania</b> | 0                   | 0   | 0                    | 0  | 27                  | 67,5 | 13                   | 32,5 |
| <b>B</b> | <b>Euskara</b>    | 2                   | 5   | 12                   | 30 | 5                   | 12,5 | 21                   | 52,5 |
|          | <b>Gaztelania</b> | 0                   | 0   | 0                    | 0  | 32                  | 80   | 8                    | 20   |

13. taula: hirugarren silabako tonu moten kopuruak eta ehunekoak hizkuntza bakoitzeko informatzaileen ama-hizkuntzaren arabera.

Azkenik, agertu izan zaizkigun tonu-konbinazioak 11 izan dira denetara; 14. taulan, bakoitzaren kopuruak eta ehunekoak ematen ditugu ordenaturik.

| 1. silaba | 2. silaba | 3. silaba | kopurua | %     |
|-----------|-----------|-----------|---------|-------|
| b         | b         | g         | 66      | 41,25 |
| b         | b         | gb        | 25      | 15,63 |
| g         | b         | gb        | 24      | 15    |
| b         | g         | gb        | 20      | 12,5  |
| b         | b         | bg        | 10      | 6,25  |
| g         | g         | gb        | 4       | 2,5   |
| g         | g         | b         | 3       | 1,88  |
| b         | g         | b         | 2       | 1,25  |
| b         | g         | g         | 2       | 1,25  |
| g         | b         | bg        | 2       | 1,25  |
| g         | g         | g         | 2       | 1,25  |

14. taula: tonuen konbinazioen kopuruak eta ehunekoak denetara.

15. taulan, konbinazio bakoitzaren kopuruak hizkuntzako eta informatzaileen ama-hizkuntzaren arabera sailkatuta agertzen dira. Bertan argi agertzen da euskaraz askozaz tonu-konbinazio gehiago erabiltzen dela gaztelaniaz baino. Era berean, B taldea homogeneoagoa da hizkuntza bietan: euskaraz 6 tonu-konbinazio eta gaztelaniaz 3 erabiltzen ditu; A taldeak, ostera, euskaraz 8 eta gaztelaniaz 4.

| Tonuak silabako |           |           | Euskara  |          | Gaztelania |          |
|-----------------|-----------|-----------|----------|----------|------------|----------|
| 1. silaba       | 2. silaba | 3. silaba | A taldea | B taldea | A taldea   | B taldea |
| b               | b         | g         | 0        | 10       | 0          | 0        |
| b               | b         | gb        | 4        | 5        | 25         | 32       |
| g               | b         | gb        | 13       | 7        | 3          | 2        |
| b               | g         | gb        | 2        | 0        | 0          | 0        |
| b               | b         | bg        | 0        | 0        | 2          | 0        |
| g               | g         | gb        | 4        | 0        | 10         | 6        |
| g               | g         | b         | 0        | 2        | 0          | 0        |
| b               | g         | b         | 10       | 14       | 0          | 0        |
| b               | g         | g         | 1        | 2        | 0          | 0        |
| g               | b         | bg        | 2        | 0        | 0          | 0        |
| g               | g         | g         | 4        | 0        | 0          | 0        |
| <b>Guztira</b>  |           |           | 40       | 40       | 40         | 40       |

15. taula: tonuen konbinazioen kopuruak denetara informatzaileen ama-hizkuntzaren arabera.

#### 4. Informatzaileen sailkapena

Atal honetan silaba bakoitzeko lortu ditugun tonu moten arabera sortu dugu informatzaile bakoitzeko bektore bat matrize batean, metodologiari buruzko atalean azaldu dugun bezala. Informatzaileen artean, bektore horien arteko distantziak kontuan hartuta, multzo homogeneorik egin daitekeen ikusteko, atal honetan, informatzaileen sailkapena egingo dugu lau metodoen arabera. Erabili ditugun metodoak honako hauek dira:

- Multzokatze hierarkikoa
- K-batezbestekoen multzokatzea (KM)
- Dimentsio aniztun eskala (DAE)
- FCM algoritmoaren araberako multzokatzea (FCM)

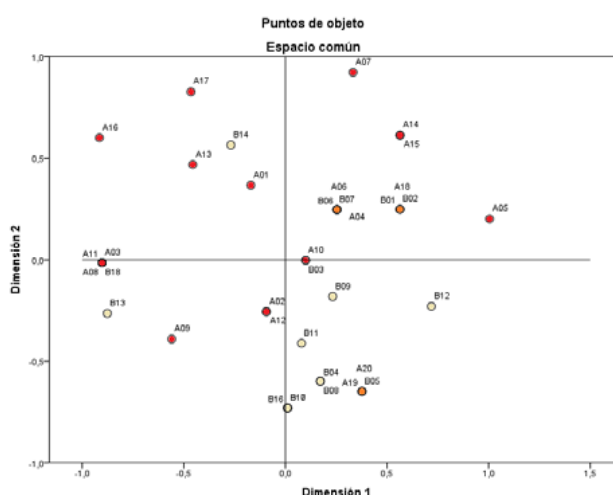


Taula aztertzen badugu, ikusten dugu A multzoan 18 informatzaile kokatzen direla (A taldeko 9 eta B taldeko 9), B multzoan 9 (A taldeko 6 eta B taldeko 3) eta C multzoan 13 (A taldeko 5 eta B taldeko 8). Banaketan dauden aldeak ez dira estatistikoki esanguratsuak, multzoen arteko distantziak txikiak dira eta (18. taula).

|   | A     | B     | C |
|---|-------|-------|---|
| A | 0     |       |   |
| B | 2,081 | 0     |   |
| C | 1,499 | 1,981 | 0 |

**18. taula:** multzoen arteko distantziak.

DAE metodoaren bidez informatzaileak lau batera proiektatzean, 12. irudiko dispersioa lortzen dugu; horretan, argi agertzen den bezala, ez dago azpimultzo argirik egiterik informatzaileen ama-hizkuntzaren arabera, denak nahiko nahasirik agertzen dira eta.



**12. irudia:** informatzaileen dispersioa DAE metodoaren arabera.

Azkenik, FCM algoritmoaren arabera informatzaileak hiru multzo nagusitan sailkatzen ditugunean, 19. taulako emaitzak lortzen ditugu. Informatzaile bat multzo batekoa onartzeko, haren barnekotasun-indizea 0,85 eta 1 artekoa izan behar da; euren barnekotasun-indizea 0,85 baino txikiagoa bada, trantsiziokotzat joko dugu.

| Multzoa | Informatzaileak   |
|---------|---|
| A       | A01, A04, A05, A06, A10, A14, A15, A18<br>B01, B02, B03, B06, B07, B12, B14, B15, B19 |
| B       | A03, A08, A09, A11, A16<br>B13, B18, B20  |
| C       | A02, A07, A12, A19, A20<br>B04, B05, B08, B10, B11, B16, B17                          |

**19. taula:** informatzaileen multzokatzea FCM algoritmoaren arabera.

Emaitzak KM metodoaren bidez lortutakoen antzekoak dira. Aldaketa bakarrak B09, A17 eta A13 informatzaileei dagozkie; hauek multzo biren arteko mugan kokatuta agertzen

dira. A13 informatzailea A multzoaren barnean 0,51 indizearekin agertzen da eta B multzoan 0,48 indizearekin; A17 informatzailea, B multzoan 0,82 indizearekin eta C multzoan 0,09 indizearekin. Azkenik, B09 C multzoan agertzen da 0,77 indizearekin, eta A multzoan 0,23 indizearekin.

## 5. Ondorioak eta hurrengo lanak

Azken atal honetan datuen azterketaren emaitzetatik atera ditzakegun ondorioarik nabarmenenak laburbilduko ditugu.

Intonazioaren azterketari dagokionez, lehen eta hirugarren silabetan agertzen den tonu motaren erabilerari hizkuntzak estatistikoki modu esanguratsuan eragiten diola ikusi dugu. Horrela, bada, euskaraz, lehen silaban goranzko eta beheranzko tonuak erabiltzen dira, eta gaztelaniaz, beheranzkoak baino ez. Hirugarren silaban, edo esaldiaren muga-tonuari dagokionez, euskaraz, lau mota agertu zaizkigu: beheranzkoa, beherazko/goranzkoa, goranzkoa eta goranzko/beheranzkoa; azken hau da gehien erabiltzen dena. Gaztelaniaz, ostera, goranzko/beheranzkoa eta goranzkoa agertu zaizkigu, eta gehien erabiltzen dena goranzkoa da.

Euskaraz, lehenengo eta hirugarren silabetan erabiltzen diren tonu kopuruetan agertzen diren aldeetan informatzaileen ama-hizkuntzak eragin estatistikoki esanguratsua dauka. Haatik, gaztelaniaz, informatzaileen ama-hizkuntzak ez dauka ezelako eraginik.

Tonu-konbinazioak aztertzean, argi agertu zaigu gaztelania homogeneousagoa dela euskara baino; izan ere, gaztelaniaz lau konbinazio erabiltzen dira, eta euskaraz, berriz, 10. Era berean, A taldeko informatzaileek konbinazio gehiago erabiltzen dute B taldekoek baino, bai euskaraz, bai gaztelaniaz. Beraz, esan dezakegu euskara sakabanatuago eta ñabarrago agertzen dela gaztelania baino.

Azkenik, informatzaileak sailkatu nahi izan ditugu euron artean ama-hizkuntzagatik azpimultzo nabarmenik egin daitekeen jakiteko, eta ikusi dugu, metodoa edozein dela ere, ezin egin daitekeela azpimultzo garbirik irizpide hori erabilita eta informatzaileak oso nahasirik agertzen direla.

Intonazioaren garrantziaz jabetu ostean, intonazioaren didaktikaren beharra azpimarratu gura genuke azken lerrootan, etorkizunera begirako proposamenak luzatu orduko.

Gazte euskaldunen bai/ez galderen intonazioan indarguneak eta ahulguneak identifikatu ostean, intonazioaren didaktika beharrezkotzat jotzen dugu, ikasleen trebakuntzan, irakasleen formazioan eta irakasleen irakasleenean.

Horretarako, gure proposamena, batetik, irakasleen hasierako formazioan intonazioaren didaktika eranstea da eta, bestetik, eskolara begira, intonazioaren irakaskuntza curriculumean txertatzea. Lan honetan egindako deskribapenenetan eta lortutako ondorioetan oinarrituta inplikazio didaktikoak doitzea eta eskola zein unibertsitateetara

eramatea litzateke hurrengo pausoa. Esku-hartze zuzenak eta egokiak diseinatzea eta praktikan jartzea.

Lantxo honetan agerian gelditu den moduan, intonazioaren irakaskuntzaren eta didaktikaren beharraz ohartuta, irakatsi beharreko hori zelan irakatsi behar den zehaztu beharra dago. Horrela, bada, agerian gelditzen da hizkuntzaren eta didaktikaren uztarketaren premia, ez baitira aparteko aldagaiak, biek eskutik joan behar baitute.

Hemen aurkeztutakoa hizkuntzaren zati minimo bat baino ez da, erakusgarri bat, bai/ez galderak kasu, intonazioa aztergai edukita. Honek guzti honek, intonaziori dagozkion beste 6 galdera motei eragin dakiekeen moduan, hizkuntzaren beste zenbait alderditan ere eragin lezake. Beraz, honetan sakontzen jarraitzeari beharrezko irizten diogu.

Hortaz, gure ikerketa-eremua zabaltzea litzateke etorkizunera begira proposatzen dugun lana, lagina ere handitu liteke eta generoa aldagai gisa sartu.

Komunikazio-sareak aberasteko bideak urratzen has gaitezen intonazioa erdigunetzat hartuta (Gaminde, Etxebarria, Garay & Romero, 2012).

## Bibliografia

- Aurrekoetxea, G. (1995) *Bizkaieraren egituraketa geolinguistikoa*, UPV/EHU-ren Argitalpen Zerbitzua, Leioa.
- Aurrekoetxea, G. (2005) Nafarroako Euskararen Sailkapenaz. in Etxebarria, P. eta Knör, H. (2005) *Nerekin yaió nun. Txillardegiri omenaldia*. Euskaltzaindia eta EHU, Iker 17, Bilbo. (109-124)
- Boersma, P. eta Weenink, D. (2016) *Praat: doing phonetics by computer* [Computer program]. <<http://www.praat.org>>
- Clua, E. (2010) Relevancia del análisis lingüístico en el tratamiento cuantitativo de la variación dialectal. in Aurrekoetxea, G. eta Ormaetxea, J. L. (2010) *Tools for Linguistic Variation*. UPV/EHU, Bilbo.
- Clua, E.; Iglesias, A.; Usobiaga, I. eta Salicrú, M. (2017) Nola identifikatu ezaugarri esanguratsuenak bariazio dialektalean: Dia Tech nabarmentzeko aukera in Iglesias, A. eta Ensunza, A. (2017) *Gotzon Aurrekoetxea lagunarterik hara*. UPV/EHU. Bilbo.
- D’Introno, F.; Del Teso, E. eta Weston, R. (1995) *Fonética y Fonología actual del Español*. Cátedra. Madrid
- Eguskiza, N.; De Pablo, I. eta Gaminde, I. (2017) Hizkuntz hizkeren sailkapena metodo kuantitatiboaren bidez. agertzeko Psikodidaktika kongresuan
- Elordieta, G. (1999). “Primer estudio comparativo de la entonación de tres variedades dialectales vascas”, in *Actas del I Congreso de Fonética Experimental*. Universitat Rovira i Virgili eta Universitat de Barcelona.
- Elordieta, G. (2000). Mendebaldeko intonazioaren inguruan, in *Mendebalde Kultura Alkartea: Mendebaldeko berbetearren formalizazioa*. Bilbo.
- Elordieta, G. (2003). Intonation, in J.I. Hualde eta J. Ortiz de Urbina (arg.), *A grammar of Basque*, Mouton de Gruyter: Berlin.

- Elordieta, G., (2007a), Constraints on intonational prominence of focalized constituents, in D. Büring., M. Gordon eta C.L. Lee (arg.), *Topic and focus: Papers from a workshop on intonation and meaning*. Springer: Dordrecht.
- Elordieta, G., (2007b), A constraint-based analysis of the intonational realization of focus in Northern Bizkaian Basque, in T. Riad eta C. Gussenhoven (arg.), *Tones and Tunes: Volume I, Typological Studies in Word and Sentence Prosody*. Mouton de Gruyter: Berlin.
- Elordieta, G., Gaminde, I., Hernáez, I., Salaberria, J. eta Martín de Vidales, I., 1999, Another Step in the Modeling of Basque Intonation: Bermeo. in V. Matousek, P. Mautner, J. Ocelíková eta P. Sojka (arg.), *Text, Speech and Dialogue*. Plzeň.
- Estebas, E.; Prieto, P. (2009): La notación prosódica en español. Una revisión del Sp\_ToBI, *Estudios de Fonética Experimental XVIII*, 263-283.
- Etxebarria, A. eta Eguskiza, N. (editoreak) (2018) *Bariazioa Esaldien Intonazioan*. UPV/EHU. Bilbo.
- Face, T.; Prieto, P. (2007): Rising accents in Castilian Spanish: a revision of Sp\_ToBI, *Journal of Portuguese Linguistics*, 6.1, 117-146.
- Gaminde, I. (2007). *Bizkaian Zehar: Euskararen Ikuspegi Orokorra*. Mendebalde Kultura Alkartea eta Bizkaiko Foru Aldundia: Bilbo
- Gaminde, I. (2010a) *Bizkaiko Gazteen Prosodiaz: Euskaraz eta Gaztelaniaz*. Mendebalde Kultura Alkartea eta Bizkaiko Foru Aldundia. Bilbo.
- Gaminde, I. (2010b): Bilboko gazte euskaldunen ezaugarri linguistikoez. in Mendebalde Kultura Alkartea: *Bilbon Mundua Ikusi*. Bilbo. (35-59)
- Gaminde, I. (2011). Noraka Euskal Azentuak?. in G. Aurrekoetxea, eta I. Gaminde (koor.), *Prosodiaz eta Hezkuntzaz I. Jardunaldiak / I. Jornadas sobre Prosodia y Educación*. UPV-EHU.
- Gaminde, I., Etxebarria, A., Garay, U., & Romero, A. (2012). Lehen hizkuntzaren eragina bizkaiko gazteen lexikoaren erabileran. *Ikastorratza, e-Revista de didáctica*, 9.
- Gaminde, I., Hernáez, I., Etxebarria, B. eta Etxeberria, P., (1997), An Analysis of the Intonation for a Pitch Accent Variety of the Basque Language. in A. Botinis, G. Kouroupetroglou eta G. Carayiannis (arg.), *Intonation: Theory, Models and Applications. Proceedings of an ESCA Workshop*. Atenas.
- Gaminde, I.; Romero, A. & Legarra, H. (2012). *Gramatika eta Hizkuntz Bariazioa Bermeon*. Bermeoko Udala eta Campos Hegaluzea.
- Gaminde, I.; Aurrekoetxea, G.; Etxebarria, A.; Garay, U. & Romero, A. (2014). *Ahoskera Lantzeko Argibideak eta Jarduerak. Laguntzarako materiala: Teoria eta Praktika*. UPV/EHU. Bilbo
- Gaminde, I., Etxebarria, A., Eguskiza, N., Romero, A. eta Unamuno, L. (2016) Lexikoaren bariazioa eta multzokatze-azterketa. *Euskalingua-28*, 19-3
- Gaminde, I.; Olalde, A.; Etxebarria, A.; Eguskiza, N. eta Gaminde, U. (2017a): *Hizkuntza Aldakortasuna Larrabetzun*. Larrabetzuko Udala. Bilbo.
- Gili, S. (1978) *Elementos de fonética general*. Gredos. Madril.
- Goebel, H. (1992) Problèmes et méthodes de la dialectométrie actuelle (avec application à l' AIS), in Euskaltzaindia/Académie de la Langue Basque éd., *Nazioarteko dialektologia biltzarra. Agiriak/Actes du Congrès International de Dialectologie* (Bilbo/Bilbao 1991), Bilbo/Bilbao, 429-475.

- Goebl, H. (2010) Introducción a los problemas y métodos según los principios de la Escuela Dialectométrica de Salzburgo. in Aurrekoetxea, G. eta Ormaetxea, J. L. (2010) *Tools for Linguistic Variation*. UPV/EHU, Bilbo.
- Hualde, J. I. (1997). *Euskararen azentuerak*. EHU eta GFA: Bilbo, Donostia. *ASJU*, gehigarriak, 42.
- Hualde, J. I. (2003). El modelo métrico y autosegmental in P. Prieto (koor.), *Teorías de la Entonación*. Ariel: Bartzelona.
- Hualde, J. I. (2006). Estado actual de las investigaciones sobre la acentuación en lengua vasca, *Oihenart* 21.
- Ito, K. (2002a). *The interaction of focus and lexical pitch accent in speech production and language comprehension: Evidence from Japanese and Basque*. Doktorego tesia, University of Illinois at Urbana-Champaign.
- Ito, K. (2002b). The effect of focus on lexical pitch peak alignment in Tokyo Japanese and Bermeo Basque, *LP2002 Proceedings*. Tokio.
- Ito, K. (2003). The focus-independent effect of tonal proximity on the realization of lexical pitch accent in Tokyo Japanese and Bermeo Basque in M.J. Solé, D.Recasens eta J. Romero (arg.), *Proceedings of the 15th International Congress of Phonetic Sciences*. Bartzelona. CD-Rom.
- Jun, S. A. & Elordieta, G. (1997). Intonational Structure of Lekeitio Basque. in A. Botinis, G. Kouroupetroglou eta G. Carayiannis (arg.), *Intonation: Theory, Models and Applications, Proceedings of an ESCA Workshop*. ESCA: Atenas.
- Legarra, H. (2011). *Muxikako euskararen azterketa linguistikoa*. Doktorego Tesia, Deustuko Unibertsitatea.
- Navarro, T. (1980) *Manual de pronunciación española*. CSIC. Madril.
- Pardo, A. eta Ruiz, M. A., (2002) *SPSS 11 Guía para el análisis de datos*. McGraw-Hill, Madril.
- Pérez López, C. (2005) *Métodos estadísticos avanzados con SPSS*. Thomson, Madril.
- Prieto, P. (coor.) (2003) *Teorías de la entonación*. Ariel. Bartzelona.
- Prieto, P. eta Roseano, P. (eds.) (2010) *Transcription of Intonation of the Spanish Language*. Lincom Europa: München.
- Quilis, A. (1988) *Fonética acústica de la lengua española*. Gredos. Madril.
- Quilis, A. (1993) *Tratado de fonología y fonética españolas*. Gredos. Madril
- Sosa, J. M. (2003) La notación tonal del español en el modelo Sp-ToBI. in Prieto, P.: *Teorías de la entonación*. Ariel. Bartzelona.



## EUSKARAREN ERABILERA NEURTZEKO KALE-NEURKETA BILBOKO HEZKUNTZA FAKULTATEAN<sup>6</sup>

UPV/EHU

Aiora Aginaga, Ibai Bilbao, Irati Etxebarria, Amaia Izarzugaza, Eider Larrabeiti, Ibai Urrutia eta Lorea Zulaika

**Laburpena:** Euskararen egoera administratiboak EAEn izandako aldaketek azken urteetan hizkuntzaren ezagutzan eta erabileran bilakaera bat egotea eragin dute. Bilakaera hori kontuan hartu behar bada ere, oraindik euskarak hizkuntza gutxitua izaten jarraitzen du, eta diglosi egoera batean aurkitzen da, EAEko kasuan gaztelaniak pisu handiagoa izaten jarraitzen duelarik. Halaber, ezagutzak gora egin duen arren, ahoz aho dabil erabilerak goia jo izanaren kezka. Geroz eta elebidun gehiago dagoen arren, oraindik erabilera gaztelania erraztasun gehiagoz darabilte hiztunek. Testuinguru hau kontuan izanda, Bizkaiko Hezkuntza Fakultatean egoera nola islatzen den ikertu dugu. Horretarako, fakultateko euskararen erabilera neurtzeko kale-neurketak burutu ditugu.

**Abstract:** The administrative situation of the Basque in the Basque Country, has changed the use and knowledge of the language. In spite of the evolution, the Basque is a minority language and is suffering from a diglossia situation, due to the influence of Spanish. Although the knowledge has developed, there is concern that the use of language is decreasing. Nowadays in the aforementioned autonomous community, are more bilingual people than years ago but they have more fluency with Spanish. Taking into account this context, we done an investigation in the Bizkaia Education Faculty to examine the use of Basque, by street measurement.

### 1. Sarrera

Bide luzea. Mendi ibilbide bat. Malda. Erronka. (...) Malda izan da bidea. Nekea eragin du; urradura batzuk oinetan. Egiteko asko dagoela erakutsi du, tontorra goian dagoela oraindik. Eta errealismoa ekarri du: ezagutzaren oinetakoak ez direla aski tontorrera iristeko. Oinetakoak edukitzeak ez duela esan nahi oinez ibiltzea. Erabiltzea. Baina egindakoari begiratzea ere komeni da, perspektiba ez galtzea. Abiapuntua non geratu den ohartzea. Aurrera egiteko baldintzak hobetu direla konturatzea. (Goikoetxea, 2016, 5-6.or)

Garikoitz Goikoetxeak 2016an kaleratutako *Euskara Irabazteko bidean* liburuko hitzaurrean ageri da aipu hau. Ondo adierazten du euskararen ibilbidea nolakoa izan den gaur egunera arte, eta egun dugun erronkari egiten dio erreferentzia. Hain zuzen, txosten honetan Bilboko Hezkuntza Fakultatean euskararen erabilera neurtzeko egindako neurketen inguruan arituko gara. Lehendik bagenuen euskarak fakultatean duen erabilerari buruzko kezka bat, kontuan hartu behar baita egun ikasten ari garenok etorkizuneko irakasle izateko prestatzen ari garela. Horregatik, interesgarria izan da ikerketa hau egitea. Modu objektibo batean dagoen gabeziak eta erronkari ohartzeko balio izan digulako.

---

<sup>6</sup> Lan hau Euskaralari 2017/2018 deialdian jasotako dirulaguntzari esker egin da.

## 2. Marko Teorikoa

Marko teorikoa egiteko, testuingurua kontuan hartu dugu: Bilboko Hezkutza Fakultatean egin ditugu kale-neurketak, Bizkaian kokatzen gara, eta, beraz, administratiboki Euskal Autonomia Erkidegoan. Hori, kontuan hartu dugu, bai baitakigu euskararen egoera administratiboa ez dela berdina Euskal Herriko lurralde guztietan, eta horrek hizkuntzan ere eragina baitu.

### 1.1. Euskararen egoera administratiboa EAEn

Euskararen aldeko gizarte ekimenak indarra hartu zuen 80. hamarkadan, ikastolak ireki ziren eta helduen euskalduntze- alfabetatzea ere ematen hasi zen (Goikoetxea, 2016). 1982an EAEn, Euskararen Erabilera Normalizatzeko Legea onartu zen, gehiengo zabalaz (Goikoetxea, 2016). Alde batetik, legearen oinarri diren printzipio orokorretatik abiatuta, euskara Euskal Herriko Autonomia Erkidegoaren hizkuntza propio gisa aitortzen du eta euskara eta gaztelania bere lurraldean hizkuntza ofizial gisa. Titulu berean hizkuntza-arrazoiengatik bereizkeria galarazten da (Eusko Jaurlaritza, 2012). Horrez gain, hezkuntza-sistema jartzen du euskara bultzatzeko ardatz nagusitzat (Goikoetxea, 2016); 17. artikuluan, hain zuzen, zera dio legeak: “Jaurlaritzak, derrigorrezko ikastaldia bukatzerakoan ikasleek, aukera berdinetan, bi hizkuntza ofizialak benetan erabiltzeko adina menperatuko dituztela ziurtatzera bideratutako neurri guztiak hartuko ditu; halaber, euskal giroa ziurtatuko du, euskara [ikastetxeetako] barne- nahiz kanpo-jardueretan eta administrazio ekintza eta dokumentuetan ohiko tresna bihurtuz” (Eusko Jaurlaritza, 2012).

Euskararen aurrerabiderako ezinbesteko zutabea da hezkuntza. Azken hamarkadetan, garapen handiena bertan eman dela esan daiteke. Bilakaerari erreparatuta, 1976-1977ko datuekin alderatuta, lehen irakasleen % 5ek zekien euskaraz, eta egun, % 80 baino gehiagok (Goikoetxea, 2016).

### 1.2. Euskara: hizkuntza gutxitua eta diglosia

Euskara hizkuntza gutxitua dela esaten da. Hizkuntza gutxitua definitzeko orduan, honako irizpide kuantitatiboak hartu behar izaten dira kontutan: biztanleriaren proportzioa, biztanleriaren banaketa edo hizkuntzaren erabilera. Era berean, gizartearen antolaketarekin edo irizpide soziopolitikoekin zerikusia duten alderdiak ere hizkuntza gutxituen definizioan integratu behar direla diote adituek: hizkuntzaren estatusa, bazterketa, balioa galtzea, gutxiagotasuna, menderakuntza, dominazioa eta boterea, esaterako; halaber, boterea erabiltzeko aukera, talde-kontzientzia, tentsioa, erresistentzia eta gatazka. Parametro hauetan denetan kokatu behar da hizkuntza gutxituen definizioa (Manterola eta Berasategi, 2011).

Era berean, euskara diglosi egoera batean bizi dela ere esaten da. Diglosia kontzeptua honela definitzen du Coyosek:

Hizkuntza-komunitate batean bi hizkuntza edo bi hizkuntza- barietate desberdinen erabiltzea da. Hizkuntza edo barietate horiek esparru eta funtzio desberdinetarako erabiltzen dira: bata funtzio formaletarako (...) eta bestea informal eta pribatuetarako (...). Maizenik egoera diglosikoa

desorekatua da funtzio horien banaketa bidezkoa ez delako. Gainera egoera diglosikoa dinamikoa da, hizkuntza batek, menderatzaileak, hain zuzen, besteren funtzio eta erabilera-eremu berri batzuk bereganatzen dituelako denboran zehar. (Manterola eta Berasategi, 2011: 22)

Irizpide kuantitatiboak kontuan hartzerakoan, argi geratzen da euskara hizkuntza gutxitua dela. Hain zuzen ere, *Inkesta Soziolinguistikoak 2016n* EAEn egindako VI. edizioak biltzen dituen datuei erreparatu diegu. Bertan ikusten den legez, EAEn bizi den 16 urte edo gehiagoko biztanleriaren % 33,9 euskalduna da, % 19,1 euskaldun hartzailea eta % 47 erdaldun elebakarra. Kontuan hartu behar da, EAE osatzen duten hiru lurraldeak alderatuz, emaitzen arteko ezberdintasunak nabariak direla. Izan ere, Gipuzkoako biztanleriaren erdia euskalduna da (% 50,6), Bizkaiko lurraldean, aldiz, euskaldunen ehunekoa % 27,6koa da eta, azkenik, Arabako biztanleriaren bostena da euskalduna (% 19,2) (Irekia, 2016).

### *1.3. Euskararen erabilera urria*

Ikasleek beren egunerokotasunean erabiltzen duten euskararen erabilera ulertu ahal izateko, kontuan hartu beharreko irizpideetako bat beraien testuingurua dugu; modu zuzenean edota zeharka, etxeko hizkuntzaren erabileraren inguruko datuek zerikusi handia izaten baitute honetan. Azken urteotako datuei erreparatuz gero, ikus dezakegu etxearen barne-espazioan aldaketa nabarmenak izan direla. Izan ere, 1991n EAEKo euskaldunen % 63 etxe guttiz euskaldunetan bizi zen. Orain datu horiek gaur egungo datuekin alderatuz gero, kopuru hori % 40ra jaitsi dela ikusi daiteke (Berria, D.G).

Motelaldi hau gehienbat ahozko jardun informalean nabari da, eta kalea da agertoki garbienteko bat. 1980ko hamarkadatik 2000. urtera bitartean euskararen erabilerak gora egin zuen arren, apurka-apurka igoeraren erritmoa apaltzen joan zen. Izan ere, 2000. urteaz geroztik, igoera hori eten egin dela nabaritu ahal izan dugu.

Bilboko Hezkuntza Fakultatea Bizkaian kokatuta dago; beraz, Bizkaiko euskararen egoerara hurbilduko gara, euskararen egoera aztertzeko. Probintzia honetan, euskararen erabilerak hogeitun urtean 3 puntuko igoera baino ez du izan; igoera motela da, kontuan hartuta gaur egun, esaterako, Bilboko biztanleriaren % 22 inguruk euskaraz dakiela, eta 1981ean % 6 baino ez zela kopuru hori. Euskaraz dakiten gehienak umeak eta gazteak dira, euskaraz eskolatuak izan direlako eta horren ondorioz ikasi dutelako (Goikoetxea, 2016).

### *1.4. Elebitasuna*

Euskal Herriak eta gure hizkuntzak mendeetan zehar jasan duten egoera sozio-politikoaren ondorioz, esan genezake, maila handiagoan edo txikiagoan, denok egin behar izaten dugula gaztelaniaz edo frantsesez geure eguneroko bizitzan, elebiduna baita euskaldunon gaur egungo errealitatea (Duñabeitia, 2016).

EAEko kasuan, gizarteko elebitasuna desorekatua dela esan daiteke, batzuk hizkuntza ofizial biak jakin eta hitz egiteko gai badira ere gainontzekoek gaztelania baino ez baitakite. Horrek hizkuntza erabiltzeko hiztunek duten eskubidea mugatzen du (Irekia, 2013).

Honetaz gain, norberaren elebitasun desorekatua gailentzen da: elebidunen ehuneko handi batek errazago egiten du gaztelaniaz euskaraz baino, eta berezko joera izaten da errazagoa egiten zaien hizkuntzan hitz egitea (Irekia, 2013). Bizkaian bizi diren elebidunei euskaraz edo gaztelaniaz hitz egiteko erraztasunari buruz galdetzerakoan, 2011n ateratako datuek zera diote: % 19,6k euskaraz errazago egiten du, % 28k bietan berdin, eta % 52k erdaraz errazago. Euskaraz eta gaztelaniaz dakitenen artean, elebidun kopuru handiena 24 urtetik beherakoena da, eta elebidunak Bizkaian % 30,3 dira guztira (Goikoetxea, 2016).

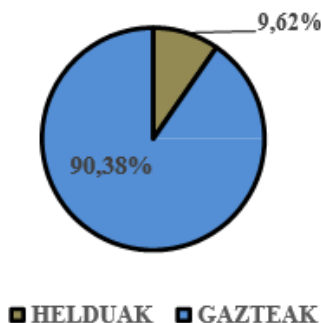
### 3. Metodoa

Bizkaian kokatuta dagoen Euskal Herriko Unibertsitateko Bilboko Hezkuntza Fakultatean hizkuntzen erabilera zer-nolakoa den denbora-tarte batez aztertzen ibili ondoren, hipotesi horien guztien atzean dagoena zer den aztertzeari ekingo diogu. Horretarako, taldekideok hiru eta lau taldeetan banatu gara eta BIEUn aurkitzen diren pertsonen elkarrizketak aztertu ditugu.

#### 3.1. Parte-hartzaileak

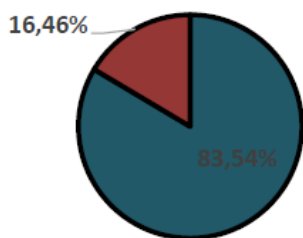
Ikerketa hau Leioako EHUko campusean kokatzen den Bilboko Hezkuntza Fakultatean egin dugu, 2018ko apirileko bigarren astean zehar. Taldekideok eraikineko toki ezberdinetatik aritu gara hiztun bila, hala nola ikasgelan, jantokian, tabernan edota atsedenean. Ikasgelarena argitze aldera, eskolak ematen ari diren bitartean ez dugu daturik hartu; izan ere, eskolak euskaraz ematen direnez, irakasleari hizkuntza honetan egiten diote. Behin saioa bukatuta, bakoitza bere kontura dabilela, une horiek aprobe txatu ditugu informazioa biltzeko.

**PARTAIDE KOPURUA ADINAREN ARABERA**



Ohikoa den moduan, unibertsitateko testuinguruan, ikasleak dira nagusi, gehienak gazteak (15-24 urte bitartekoak), eta gutxiago badira ere, irakasleak, helduen tartean sartu ditugunak (24-64 urte bitartean). Esan beharra dugu kale-neurketa honetako lagina 395 partaidekoa izan dela. Hori dela eta, adin tarte hauetako laginak bakarrik izan dira gure azterketan; hain justu ere, % 90,38 gazteak eta % 9,62 helduak.

**AZTERTUTAKO PARTAIDEAK GUZTIRA**

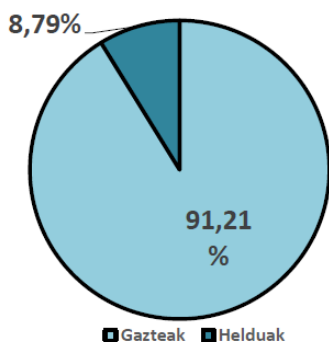


■ EMAKUMEAK ■ GIZONAK

Generoari erreparatuz, emakumeak gehiago izan dira gizonaezkoak baino: % 83,54 eta % 16,46.

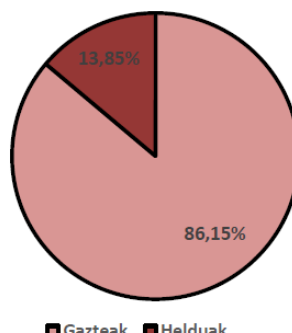
Andrazkoetan zentratuz, ehuneko laurogeita hamar baino gehiago gazteak dira (% 91,21), eta helduak, aldiz, % 8,79. Gizonezkoetan ere antzeko datuak ditugu: % 86,15 gazteak eta % 13,85 helduak.

EMAKUMEAK ADINAREN ARABERA



■ Gazteak ■ Helduak

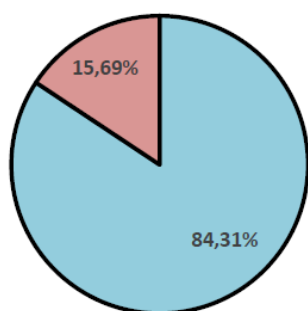
GIZONAK ADINAREN ARABERA



■ Gazteak ■ Helduak

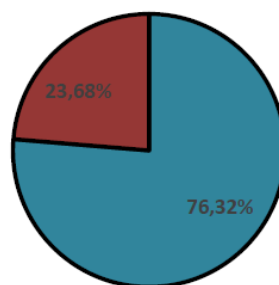
Ezaugarriak bukatzeko, adin-tarte bakoitzean generoaren arabera dagoen desberdintasuna aipatuko dugu: gazteen artean, ia ehuneko laurogeita bost emakumezkoak dira (% 84,31), eta % 15,69 soilik, gizonaezkoak; alde nabarmena ikasleen artean. Irakasleetan ere antzeko ehunekoak aurkitu ditugu: gizonaezkoak ez dira laurdenera iristen (% 23,68), eta emakumezkoak % 76,32 dira.

GAZTEAK GENEROAREN ARABERA



■ Gazte emakumeak ■ Gazte gizonak

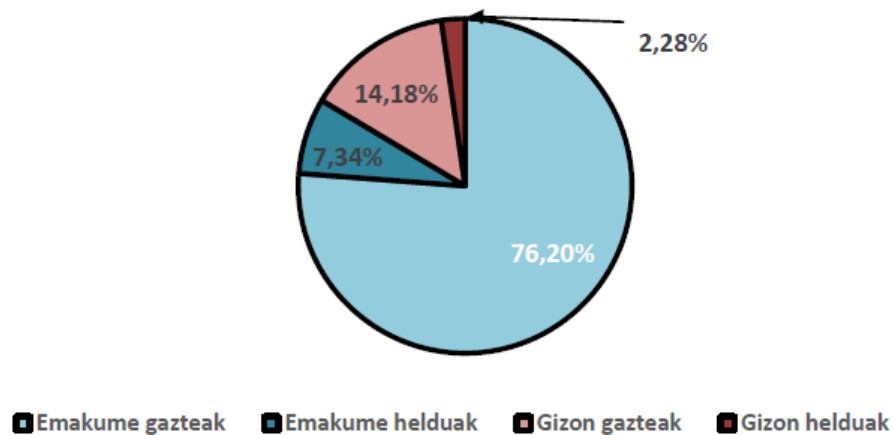
HELDUAK GENEROAREN ARABERA



■ Heldu Emakumeak ■ Heldu gizonak

Laburbiltzeko, grafiko honen bitartez, partaide guztiak kontuan izanik, generoa eta adina aintzat hartuz taldeka dagoen kopurua ikus daiteke. Horrela, ikus dezakegu parte-hartzaileen hiru laurden baino gehiago (% 76,2) emakume gazteak direla, % 14,18 gizonaezko gazteak, % 7,34 emakume helduak eta, gutxien, gizonaezko helduak, % 2,28.

**PARTAIDE GUZTIAK KONTUAN IZANIK, SEXU ETA ADINAREN ARABERA DAGOEN KOPURUA TALDEKO**



### 3.2. Baliabidea: inkesta soziolinguistikoa

Inkesta soziolinguistikoa euskararen bilakaera sozialari buruz egiten den azterketa sakon, jarraitu eta sistematikoa da. Ikerketa honen bitartez, biztanleen hizkuntza-gaitasuna, hizkuntzen ezagutza erlatiboa, lehen hizkuntza, transmisioa, hizkuntzaren galera eta irabaziak, eta eremu bakoitzeko euskararen erabilera aztertzen dira (Baztarrika, 2016). Behaketa hau aurrera eramateko, honako oharrak izan beharko genituzke kontuan:

- Ahalik eta elkarrizketa gehien jasoko dira.
- Neurtzaileekin hitz egiten dutenean, ez dira datuak jasoko.
- Normalean, elkarrizketetan pertsona bakarrari entzungo diogu.
- Elkarrizketa elebidunetan, bi aldiz jasoko dira elkarrizketa bera.
- Pertsona batek mintza-jardunean bi hizkuntza erabiltzen baditu, nagusitzen dena jasoko da.
- Neurraldia hastean, fitxa-orri berri batean hartuko ditugu datuak.
- Mugikorretik hitz egiten dagoena, pertsona bakar bezala jasoko dugu.

Behin baldintza hauek aintzat hartuta, kale-neurketari ekin diezaiokegu. Gure kasuan, Leioako EHUKo Hezkuntza Fakultatean egin ditugu; honek gure laginen ezaugarriak (adina, jatorria.. ) eta kopurua mugatu ditu. Herrietako kaleetan aniztasun handiagoa aurki dezakegu jendearen joan-etorrian, eta neurketa gehiago lortu.

### 3.3. Prozedura

Arestian aipatu modura, datuak jasotzeko momentu, toki eta teknika desberdinak erabili ditugu. Tokiei erreparatuz gero, batik bat, korridore eta atsedeen-guneetan aritu gara: hor dago jende kopuru gehien, eta hiztunen arteko komunikazioa informalagoa da. Honetaz gain, ikasgeletara sartu izan gara, betiere eskolak ematen ari ez diren bitartean; hain zuzen ere irakaslearekin hitz egiteko hizkuntza denez, euskara baita komunikatzeko hizkuntza

testuinguru formal horretan. Tokiak bukatzeko, irakasleak zein administrazioko langileak dabilzan espazioetatik ibili gara, kale-neurketetan bildutako laginen adin-tartea zabalagoa izateko, baita jantokian ere.

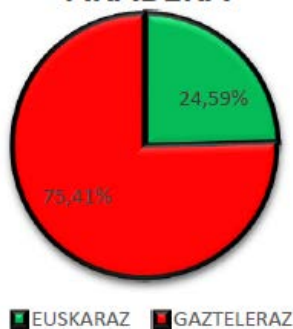
Aztertze erabili dugun momentuari dagokionez, eguneko une desberdinak erabili izan ditugu: autobusetik heldu bezain pronto (7:30-8:30), atsedendietan, klaseko tartetan eta etxerako bidean.

Prozedurak bukatzeko, erabilitako teknikak aipatuko ditugu; izan ere, anitzak izan dira egoeraren arabera. Unibertsitatera heldu bezain pronto, taldekideok bakarka aritu gara, ikasgela desberdinetan sartuz eta lanean ariko bagina bezala antzetzuz. Gero, atsedendietan zein klase-tartetan, taldea bizpahiru multzotan banatuta, korridoreetatik zehar ibili gara, eta bai jantokian eta tabernan ere. Bukatzeko, hiruzpalau pertsonako taldeak sortuta, fakultateko "arrainontzietan" datuak jaso ditugu, lana egiten geniharduen bitartean.

### 3.4. Emaitzak

Aste osoan zehar kale-neurketak egin ostean, paperean jasotako datuak ordenagailura pasatzeko ordua da, horretan informazio guztia biltzeko eta emaitzak ateratzeko. Zifra hauek ehunekoan adierazi ditugu grafiko bidez, irakurlearentzat interpretatzeko errazagoak eta erabilgarriagoak izango direlakoan baikaude.

#### KOPURUA HIZKUNTZAREN ARABERA



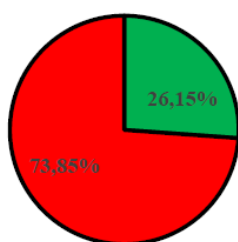
Lehenik eta behin, lagin guztiak kontuan hartuta, ikus dezakegu lautik batek (% 24,59) soilik hitz egiten duela euskaraz; gaztelaniaz, aldiz, % 75,41ek. Leioako campuseko Hezkuntza Fakultateko hizkuntzen erabilera orokorra ezagututa, sexuaren eta adin-tartean arabera zehaztuko ditugu emaitzak.

Gizonezkoen erreparatur, batez besteko datu orokorrak ehuneko hobeak ditu euskararidagokionez; izan ere, % 26,15i entzun diogu euskaraz eta % 73,85i gaztelaniaz.

Emakumezkoetan, ordea, gaztelaniari buruzko datuek zertxobait gora egin dute;

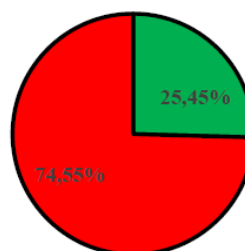
hain zuzen ere, % 74,55ek egin du gaztelaniaz, eta % 25,45ek euskaraz.

#### GIZONAK ETA HIZKUNTZA



■ EUSKARAZ ■ GAZTELERAZ

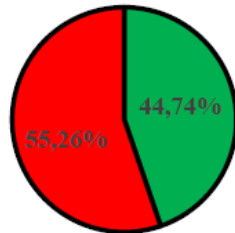
#### EMAKUMEAK ETA HIZKUNTZA



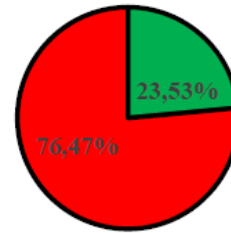
■ EUSKARAZ ■ GAZTELERAZ

Adin-tarteetan erabiltzen den hizkuntzari erreparatuz gero, ikus dezakegu helduetan nahiko parekatuta daudela; hain justu ere, % 44,74k euskaraz, eta gainontzekoak, % 55,26k, gaztelaniaz. Gazteetan, ordea, gaztelania askoz pisu handiagoa dauka; izan ere, hiru laurden baino gehiagok (% 76,47k) gaztelania erabiltzen du komunikatzeko, eta % 23,53k.

**HELDUAK ETA HIZKUNTZAK**



**GAZTEAK ETA HIZKUNTZAK**



■ EUSKARAZ ■ GAZTELERAZ

■ EUSKARAZ ■ GAZTELERAZ

Datuak eta emaitzak emanda, ondorioak ateratzeko unea da. Argi geratu da euskararen erabilera baxua dela: % 25 inguru, esparru gehienetan; helduen kasuan izan ezik, non ia erdiek euskaraz hitz egiten duten. Jasotako heldu gehienak irakasleak izan dira, eta sarritan ikusi ditugu elkarren artean euskaraz mintzatzen; beraz, esan daiteke eskolak euskaraz emateaz aparte hizkuntza hau komunikatzeko ere erabiltzen dutela. Hala ere, esan beharra dago ez dugula ikerketa-laginik nahikoa hori ziurtatzeko, gehienbat jasotako datuak gazteen arteko elkarrizketei baitagozkie.

Gazteen edo ikasleen kasuan, kezkarriak iruditzen zaizkigu datu hauek. Izan ere, % 23,53k soilik hitz egiten du euskaraz. Gure fakultatea hein handi batean hezkuntzara bideratuta dagoela ikusita eta gradu gehienak euskara hutsekoak direla jakinda, erabilera oso urria da. Honek ez du esan nahi gazteen maila baxua denik, baina bai elkarren artean komunikatzeko orduan erraztasun handiagoa daukatela gaztelaniaz. Elebidunek gaztelaniaz edo euskaraz hitz egiteko duten erraztasunari buruz aurretiaz aipatutako datuekin pareka genezake datu hau; Bizkaiko kasuan, % 19,6k baino ez baitu euskaraz hitz egiteko erraztasun handiagoa. Kaleko errealitatea fakultatean islatzen dela argi dago.

Beste alde batetik, datuak jasotzen genbiltzala, konturatu gara euskaraz hitz egiten zuten pertsonak talde txikietan zeudela, hiruzpalauko taldeetan, alegia. Talde handiagoetan egonda, gaztelaniaren erabilera nagusitu da, beharbada talde horretako kide guztiek uler zezaten. Hau diglosiaren fenomenoarekin lot genezake.

Bukatzeko, esan beharra dugu kale-neurketa hau egiteko lortutako elkarrizketa kopurua ez dela nahikoa datu zehatzak ateratzeko. Gure mugak izan ditugu, hala nola irakasleen euskara-maila neurtzeko orduan, ez dugulako nahikoa datu jaso. Hori gertatzearen arrazoia irakasleak goiko pisuetako bulegoetan egotea da, eta, sarri, nahikoa bakartuta ibiltzea. Aipatzekoa da irakasleei buruz jasotako datu asko telefonoz hitz egiten ari zirela jaso ditugula. Azkenik esan behar da ikerketa hau egiteak bestelako galderak ere sortu dituela. Halaber, generoaren ikuspegi batetik erreparatuta, deigarria da irakasle edo gizarte hezitzaile izateko prestatzen ari direnen artean gehiengoa emakumea izatea.



#### 4. Bibliografia

- Duñabeitia, J.A. (2016). *Elebitasuna, bi hizkuntza hitz egitea baino gehiago*. 2018ko maiatzaren 3an jasoa URL honen bidez:  
[http://www.ikasten.ikasbil.eus/pluginfile.php/1945003/mod\\_habecms/doc\\_attachment\\_eu/1878/H89\\_Anton%20eta%20Du%C3%B1abeitia-1.pdf](http://www.ikasten.ikasbil.eus/pluginfile.php/1945003/mod_habecms/doc_attachment_eu/1878/H89_Anton%20eta%20Du%C3%B1abeitia-1.pdf)].
- Eusko Jaurlaritza (2012) Euskararen Legeak 30 urte. 2018ko maiatzaren 8an URL honetatik berreskuratua: [http://www.euskara.euskadi.eus/contenidos/informacion/argitalpenak/eu\\_6092/adjuntos/EUSKARARENLEGEA\\_WEB.pdf](http://www.euskara.euskadi.eus/contenidos/informacion/argitalpenak/eu_6092/adjuntos/EUSKARARENLEGEA_WEB.pdf)
- Goikoetxea, G. (2016). Lehen Atala: Orain arteko ibilbidea. *Euskara irabazteko bidean*.(5-47. orr) Donostia: Elkar-Berria-Jakin.
- Goikoetxea, G. (2016). *Euskararen erabilera jaitsi egin da kalean*. Berria egunkaria. 2018ko maiatzaren 5ean URL honen bidez berreskuratua:  
[https://www.berria.eus/albisteak/136816/euskararen\\_erabilera\\_jaitsi\\_egin\\_da\\_kalean.htm](https://www.berria.eus/albisteak/136816/euskararen_erabilera_jaitsi_egin_da_kalean.htm)
- Irekia (2013) *Baztarririk ziurtatu du egiazko elebitasuna gaur oraindik helburua dela eta ez errealitatea*. 2018ko maiatzaren 8an URL honetatik berreskuratua:  
<https://www.irekia.euskadi.eus/eu/news/15559-baztarririk-ziurtatu-egiazko-elebitasuna-gaur-oraindik-helburua-dela-eta-errealitatea?track=1>
- Manterola eta Berasategi (2011) *Hizkuntza Gutxituak. Hizkuntz Gutxituen erronkak* (19-22 orr.) Bilbo: UEU.

## HELDUEN PROSODIAREN ERAGINA IRAKURRITAKO TESTUETAN HAURREN ARRETAN

UPV/EHU

Eudia ikerketa-taldea

Irati de Pablo, Asier Romero, Iñaki Gaminde eta Naia Eguskiza

[irati.depablo@ehu.eus](mailto:irati.depablo@ehu.eus), [a.romero@ehu.eus](mailto:a.romero@ehu.eus), [inaki.gaminde@gmail.com](mailto:inaki.gaminde@gmail.com),  
[naia.egusquiza@ehu.eus](mailto:naia.egusquiza@ehu.eus)

**Laburpena:** Ikerketa honetan aztertu nahi izan da motherese hizkuntza motak eraginik duen haurren izaera multimodalean. Horretarako, irudi berdinekin osaturiko bideo bi erakutsi zitzaizkien umeei: batean, motherese hizkuntza erabilia eta, bestean, narratzaile moduan. Corpusa Bizkaiko hamahiru eta hamalau hilabete bitarteko 12 hurrek osatu zuten; erdiak neskak izan ziren, eta beste erdiak mutilak. Datuen analisiari dagokionez, batetik, bideo bien arteko ezaugarri prosodikoen aldeak aztertu genituen eta, bestetik, hurrek bideo bakoitzaren aurrean izandako izaera multimodala. Bideoen ezaugarri prosodikoen aldeak esanguratsuak izan ziren, bai pitch-ean bai intentsitatean. Eta haurren izaera multimodala ezberdina agertu zen bideo motaren arabera, keinu gehiago eginik haurren bideoarekin.

**Resumen:** En esta investigación se ha observado si el tipo de habla motherese tiene influencia en el comportamiento multimodal de los niños/as. Para ello, se les presentaron dos vídeos a los niños/as compuestos por las mismas imágenes pero que diferían en el tipo de habla utilizada, mientras que en uno se utilizaba el habla motherese en el otro se les contaba el cuento de forma narrativa. El corpus lo compusieron 12 niños/as de entre trece y catorce meses, siendo la mitad niños y la otra mitad niñas. En lo que respecta al análisis de los datos, por una parte, se observaron las diferencias de las características prosódicas en los dos vídeos y, por otra parte, el comportamiento multimodal que tuvieron los niños/as delante de cada vídeo. La diferencia entre las características prosódicas de ambos vídeos fue significativa en el pitch así como en la intensidad. Y el comportamiento multimodal de los niños/as fue diferente en función del vídeo, produciendo mayor cantidad de gestos con el vídeo en el que se utilizaba el habla de tipo motherese

**Abstract:** In this investigation it has been observed whether the type of motherese speech has influence in the multimodal behavior of the children. For this, two videos were presented to the children, composed of the same images but which differed in the type of speech used, while in one the motherese speech was used in the other, the story was narrated. The corpus was composed of 12 children between thirteen and fourteen months old, half of them children and half of them girls. Regarding the analysis of the data, on the one hand, the differences of the prosodic characteristics in the two videos were observed and, on the other hand, the multimodal behavior that the children had in front of each video. The difference between the prosodic characteristics of both videos was significant in the pitch as well as in the intensity. And the multimodal behavior of the children was different depending on the video, producing more gestures with the video in which motherese speech was used.

### 1. Sarrera

Motherese hizkuntza eta haurren izaera multimodalaren eragina hainbat hizkuntzatan ikertu da. Gure kasuan, helburua izango da, batetik, analizatzea helduaren

prosodiaren influentzia hurrek ekoitzitako keinuetan eta, bestetik, aztertzea helduaren prosodiak oihartzuna duen produktutako bokalizazioetan eta helduaren prosodiak ondorioa duen begiradari eustean. Bide batez, zenbait bideo erakutsi zaizkio laginari: batean, motherese hizkuntza erabilia, eta bestean, narratzaile moduan.

Lana lau atal nagusitan banatuta aurkezten dugu. Sarrera honen osteko bigarren atalean, motherese hizkuntzaren ezaugarri nagusiak aztertzen dira. Hirugarren atalean, erabili dugun corpusaren ezaugarriak eta hori jasotzeko metodologia aurkezten dira. Laugarren atalean, datuen azterketa ematen da. Azkenik, bosgarren atalean, ondorio nagusiak laburbiltzen dira.

## 2. Marko teorikoa

Askotan, haur batekin hitz egiten dugunean, berarekin elkarrekintzan gaudenean, gure hizkera moldatu ohi dugu. Hortaz, helduok hurrekin eta helduekin erabilitako hizkerak, askotan, desberdinak dira. Helduok hurrekin erabiltzen dugun hizkera mota hori terminologia desberdinez ezagutu da: baby talk, motherese eta CDS “Children’s Directed Speech” (Snow, 1986; Diez-Itza, 1993). Euskaraz honetaz idatzi denean, “haur-hizkuntza”z hitz egin da; beste batzuen artean, Juan José Zubiriren “Euskarazko haur-hizkuntzaren berezitasunak” artikuluan (Zubiri, 1995) .

Riverok (1993) adierazten du haur-hizkuntza hizkuntzaren jabekuntza-prozesuan lagungarria dela; izan ere, antolaturiko input bat da, egitura aldetik sinpleagoa, erredundanteagoa, helduekin erabiltzen den hizkuntzaren aldean, eta testuinguruko orainaldian zentratuagoa. Hortaz, helduok gure kulturaren murgilduko diren eta gure hizkuntza ikasiko duten horientzako euskarri lagungarria izan gaitezke (Diez-Itza, 1993); izan ere, kontuan hartu behar dugu hizkuntza-jabekuntza gauzatzen dela helduekin elkarrekintzan dagoen testuinguruan (Hernández-Pina, 1984). Baina denbora aurrera joan ahala, eta haurraren garapen linguistikoa aurreratzen doan heinean, apurka-apurka, hizkera hori helduok erabiltzen dugun horretara hurbiltzen joango da (Zubiri, 1995; Hernández-Pina, 1984); beraz, esan liteke proportzionalki inbertsoak direla bideak: zenbat eta txikiagoa izan haurra, orduan eta egokitzapen handiagoa izango du, eta umea adinean aurreratu ahala, egokitzapen hori txikituz joango da.

Zubirik (1995:178) hizkuntza mota honen inguruan hainbat ikerketak aipaturiko ezaugarriak laburbiltzen ditu bere artikuluan:

- a) Hizkuntza zehatzagoa da helduekin erabiltzen dena baino: gehienetan haurraren ingurune hurbilari eta errealitate zehatzari dagokio.
- b) Hizkuntza enfatikoagoa da, markatuagoa.
- c) Hizkuntza errepikakorra
- d) Hizkuntza sinpleagoa
- e) Esaldi-motei dagokienez, galderazko eta agintezko perpaus asko erabiltzen dira.
- f) Heldua haurrari zuzentzen zaionean, sarritan, aditzaren hirugarren pertsona erabiltzen du, lehenengoa edo bigarrena erabili beharrean.

O’Gradyk (2010:187-188) ere hizkera mota honen ezaugarri eta abantaila batzuk aipatzen ditu:

| <b>Ezaugarriak</b>  | <b>Abantailak</b>   |
|---|---|
| <ul style="list-style-type: none"> <li>▪ Artikulazio geldoa eta zaindua</li> <li>▪ Intonazio puztua, puntu bateraino</li> <li>▪ Hiztegia oinarrizko elementuz osatua</li> <li>▪ Ohiko esaldi baino laburragoak</li> </ul> | <ul style="list-style-type: none"> <li>▪ Diskurtoaren pertzepzio hobea eta hizkuntza haren barneko osagaietan zatitzeko erraztasuna.</li> <li>▪ Ahozko diskurtoaren ulermena erraztea.</li> <li>▪ Haurrentzat errepikapenak helduek esandakoa ulertzeko bigarren aukera bat litzateke.</li> </ul> |

Honez gain, ataltxo bat utzi nahi izango genieke euskarazko haur-hizkuntzaren ezaugarriei. Ezaugarri hauek Zubirik (1995) aipatzen ditu:

1. Berezitasun fonetikoak. Txistukari afrikatuak eta frikariak sabaikari bihurtzen dira: “tx” eta “x” berdin gelditzen dira. Laburbilduz, aldaketa fonetikoak hurrengoak dira:  $z \rightarrow x$ ,  $s \rightarrow x$ ,  $x \rightarrow x$ ,  $tz \rightarrow tx$ ,  $ts \rightarrow tx$ ,  $tx \rightarrow tx$ ,  $l \rightarrow ll$ ,  $r \rightarrow ll$  edo  $dd$ ,  $t \rightarrow tt$ .
2. Berezitasun morfosintaktiko batzuk: hitz batzuk izen berezi gisa erabiltzen dira; “apapa” hitzak erabilera berezia dauka; aditza laguntzailerik gabe joan ohi da; sintagmak errepikatzeke eta hitz onomatopeikoak erabiltzeko joera dago; hitz batzuk absolutiboan soilik esaten dira, eta, kasu batzuetan, haur-hizkuntzako hitzen kategoria gramatikal berezia erabiltzen da.
3. Berezitasun lexikalak. Haur-hizkuntza erabiltzean, berezko lexikoaren zenbait hitz erabiltzen dira; beste askoren artean, apapa (kanpora), fu (sua/beroa), papa (ogia), pipia (txoria)...

Badira ikerketa asko gai honen inguruan (Díez-Itza, 1993; Thiessen, Hill, eta Saffran, 2005; Trainor eta Desjardins, 2002; Spinelli, Fasolo, eta Mesman, 2017), baina gutxi dakigu motheresek haurren produkzio multimodalean duen eraginaz, eta hutsune bibliografiko hau bete nahian gauzatu da ikerketa hau.

Ikerketa honen helburu nagusia da ezagutzea ea testu-estiloagatik prosodian gerta daitezkeen aldaketek haurren arretan eraginik duten. Planteatu ditugun bigarren mailako helburuak honakoak izan dira: (1) analizatzea helduaren prosodiaren eragina haurrek ekoiztutako keinuetan, (2) aztertzea helduaren prosodiak oihartzunik duen produzitutako bokalizazioetan eta (3) ikertzea helduaren prosodiak ondoriorik duen begiradari eustean.

### 3. Corpora eta metodologia

Atal honetan, lortutako corpora eta erabilitako metodologia azalduko dira. Corpora 12 umeren datuek osatu dute: sei neska eta sei mutil. Ume hauen adina, grabaketa-aldian, 13-14 hilabete bitartekoa zen, eta guztien ama-hizkuntza euskara zen. Bizkaiko haur-eskoletakoak izan dira. Corpusaren lorpenerako, ikertzaileok Haurreskoak Patzuergoko zuzendariarekin harremanetan jarri ginen, eta hark haur-eskola guztietara mezu bat helarazi zuen, gure ikerkuntzaren berri emanez. Gero, parte hartu nahi izan zuten haur-eskolak gurekin jarri ziren harremanetan, eta, eurak gure bitartekariak izanik, gurasoen, haur-hezkuntzako koordinatzailearen eta eurekin grabaketa mementoan egongo zen helduaren (tutore edo laguntzailearen) baimena eskuratu zen.

Metodologiari dagokionez, grabaketa hasi baino lehenagoko aldia eta grabaketen denboraldian izandakoa desberdinduiko ditugu.

Lehenengoari dagokionez, ikertzaileok erabaki genuen *Txanogorritxu* ipuina kontatzea hurrei, tradizioz kontatu den ipuina izan delako. Horretarako, ahotsa aldagai bakarra hartzeko, irudiz osaturiko bideo bat sortu da. Horren ostean, ahotsa txertatu genuen Movie Maker programa erabiliz. Horrela, ipuin bera, argazki berdinekin, ahots mota ezberdinez kontatzen zen: motherese ahots motarekin eta narratzaile-ahotsarekin. Ahots-grabaketa egin zuen pertsona emakumea da; euskara du ama-hizkuntza eta, unibertsitateko irakasle izateaz gain, haur-hezkuntzako irakasle ere bada. Bideoek antzeko iraupena zuten: minutu erdi ingurukoak ziren.

Metodologiaren bigarren partea grabaketa-aldian aurrera eramandakoa da. Kasu honetan, ikertzaileok grabaketarako testuingurua prestatzen genuen, tutorea/laguntzailea haurrarekin etorri baino lehen: ordenagailua zoruan jartzen zen, gela erdian, eta kamera, tripode bati helduta, ordenagailuaren atzealdean, haurraren keinu eta begiradak grabatu ahal izateko. Heldua haurraren alboan esertzen zen, haurra eroso sentiaraziz, baina ez zuen berarekin elkarrekintzarik izaten bideoa martxan zegoen bitartean. Bideo motak haurraren izaera multimodalean eraginik ez izateko, sei haur haurren bideoarekin hasi ziren, eta beste seiak helduaren bideoarekin; guztira 24 bideoko lagina izan zen. Grabaketa-aldian, haur gehienek denbora batetik aurrera ez zuten kasu handiegirik egiten, ez bideo bati, ez besteari, eta, ikerketa honetan naturaltasunari garrantzia eman nahi diogunez gero, erabaki genuen laginaren lehen minutua soilik aztertzea, hots, 24 minutu.

Kodifikazioa ELAN (Lausberg & Sloetjes, 2009) eta PRAAT (Boersma & Weenink, 2015) programekin egin zen. Horretan, bideoen kasuan talde prosodikoak anotatu ziren, eta, haurren grabaketen kasuan, bokalizazioak, keinuak eta begirada. Lortutako datuak segundotan kuantifikatu ziren, eta SPSS (SPSS Inc., Chicago IL) programa estatistikoarekin aztertu ziren.

## 4. Emaidzak

Emaidzak atal bitan aurkeztuko dira. Lehen atalean, bideoetan erabilitako ezaugarri prosodikoak aztertuko ditugu: iraupena, F0 eta intentsitatea. Horren ostean, hurrek bideoen aurrean izandako izaera multimodala aztertuko dugu: bokalizazioak, keinuak eta begirada; eta horrek sexuaren eta bideoaren arabera desberdintasunik erakusten duen.

### 4.1. Bideoen ezaugarri prosodikoak

Bideo bien arteko ezberdintasuna ikertzailearen ahotsa izan zen: batean, motherese bezala hitz egiten zuen, eta, bestean, narratzaile moduan. Hortaz, aztertu genuen lehen aspektua bideoen ezaugarri prosodikoak izan ziren. Hasteko, bideo bien talde prosodikoen iraupenari erreparatu genion, eta, emaidzak estatistikoki esanguratsuak ez izan arren, talde prosodikoen iraupenen batezbestekoen desberdintasuna agerikoa zen: motherese bideoen talde prosodikoen batezbestekoa 1560,9 ms-koa zen, eta narratzailearena, berriz, 1763,5 ms-koa. f0-ari dagokionez, bideo biren arteko ezberdintasuna esanguratsua izan zen t-student lagin independenteen proba aplikatuta,  $t = (a.m:187) 2,850$ ;  $p = .005$ ; motherese bideoko batez besteko f0-a altuagoa zen bestean baino. Intentsitateari dagokionez, ezberdintasunak esanguratsuak izan ziren,  $t = (a.m: 188) 10,294$ ;  $p = .000$ ; narratzaile-bideoan intentsitate altuagoa erabiliz.

### 4.2. Haurren izaera multimodala

Horren ostean, haurren izaera multimodala aztertu zen bideo bakoitzaren aurrean. Horretarako, bokalizazioak, keinuak eta begirada sexuaren eta bideo motaren arabera ezberdintasunik dagoen aztertu zen. Konbinazio guztietatik (keinu-sexu, keinu-bideo, bokalizazio-sexu, bokalizazio-bideo, begirada-sexu, begirada-bideo), emaitza esanguratsuak eman zituen bakarra keinu-bideo konbinazioa izan zen, t-student proba aplikatuta,  $t = (a.m: 25) 2,413$ ;  $p = .023$ . Umeek keinu-produkzio handiagoa izan zuten motherese bideoarekin bestearekin baino.

## 5. Ondorioak

Ikerketa honetan konprobatu ahal izan dugu haurrari zein helduari zuzendutako entzumen-estimuluak eragina daukala haurren keinuen produkzioan.

Jakin badakigu gure ikerketak mugak dituela (izan ere, corpora 12 hurrekin egin da soilik), eta, lortutako emaitza hau orokortzeko, uste dugu corpus handiagoa behar dugula. Hortaz, gure etorkizunerako proposamena litzateke corpora handitzea; eta, horretaz gain, gatzelania ama-hizkuntza duten hurrekin konparatzea.

Gainera, interesgarria litzateke lortutako emaitzak f0-arekin edota intentsitatearekin korrelazionatzen diren ikustea.

## 6. Erreferentzia bibliografikoak

- Boersma, P. eta Weenink, D. (2016) *Praat: doing phonetics by computer, Version 5.1*. [Computer program]. <<http://www.praat.org>>
- Diez-Itza, E. (1993). Variaciones tonales en el habla a los niños y adquisición del lenguaje. *Estudios de Psicología*, 1993, 50, 33-47.
- Hernández-Pina, F. (1984). Actitudes lingüísticas parentales y desarrollo del lenguaje infantil. *Infancia y aprendizaje*, 25, 35-60.
- IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.
- Lausberg, H. & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, 41 (3), 841-849.
- O'Grady, W. (2010). *Cómo aprenden los niños el lenguaje*. Madrid: Akal.
- Thiessen, E.D.; Hill, E. A. eta Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy*, 7(1), 53-71
- Trainor, L. J. eta Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review* 9 (2), 335-340.
- Rivero, M. (1993). La influencia del habla de estilo materno en la adquisición del lenguaje: valor y límites de la hipótesis del input. *Anuario de psicología*, 57, 45-64.
- Snow, C. (1986). Conversations with children. In Garman, M. and Fletcher, P. (eds.), *Language acquisition*, 69-89. Cambridge: Cambridge University Press.
- Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review*, 44, 1-18.
- Zubiri, J. J. (1995). Euskarazko haur-hizkeraren berezitasunak. *Huarte de San Juan*, 1, 175-203.

**Zabalduz**

Jardunaldi, kongresu, sinposio, hitzaldi  
eta omenaldien argitalpenak

Publicaciones de jornadas, congresos,  
simposiums, conferencias y homenajes

**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](http://www.ehu.eus/argitalpenak)

eman ta zabal zazu



Universidad  
del País Vasco

Euskal Herriko  
Unibertsitatea