

Ahoskeraren Dibertsitatea Arratia-Nerbioin (1981-2000)¹

Diverse forms of pronunciation in Arratia-Nerbioi (1981-2000)

Iñaki Gaminde*

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

ABSTRACT: This paper shows the results obtained by applying unusual techniques in the quantitative analysis of linguistic variation. On the one hand, the materials analyzed have been free oral texts and, on the other hand, instead of dichotomous variables, the probability of the facts analyzed has been taken into account. The second section describes the methodology used for data collection, processing and analysis. The third section presents the data analysis. Finally, the fourth section shows the results obtained; most notably, the difference due to gender is statistically significant.

KEYWORDS: variation; classification; phonology.

LABURPENA: *Lan honetan bizkuntz aldakortasunaren azterketa kuantitatiboan ohikoak ez diren teknikak aplikatzean lortutako emaitzak erakusten dira. Alde batetik, aztertutako materialak ahozko testu libreak izan dira eta, beste alde batetik, aldagai dikotomikoak erabili beharrean aztertu diren moldeen probabilitateak kontuan hartu dira. Bigarren sekzioan datuak jasatzeko, prozesatzeko eta aztertzeko metodologia deskribatzen da. Hirugarren sekzioan datuen azterketa aurkezten da. Azkenik, laugarren sekzioan, lortu diren emaitzak erakusten dira; nabarmenena da generoaren araberako aldea estatistikoki esanguratsua izatea.*

HITZ GAKOAK: *aldakortasuna; sailkapena; fonologia.*

¹ Bihoazkio nire eskerrik beroenak Naia Eguskizari honen zirriborroa irakurri izanagatik eta oharrengatik.

* **Harremanetan jartzeko / Corresponding author:** Iñaki Gaminde. EUDIA - Bilboko Hezkuntza Fakultatea (UPV/EHU). Sarriena, z/g (48960 Leioa). – inaki.gaminde@gmail.com

Nola aipatu / How to cite: Gaminde, Iñaki (2023). «Ahoskeraren Dibertsitatea Arratia-Nerbioin (1981-2000)», *ASJU*, 57 (1-2), 343-355. (<https://doi.org/10.1387/asju.25957>).

Jasoa/Received: 2022-07-22; Onartua/Accepted: 2022-10-23.

ISSN 0582-6152 - eISSN 2444-2992 / © UPV/EHU Press



Lan hau *Creative Commons Aitortu-EzKomertziala-LanEratorririkGabe 4.0 Nazioartekoa* lizentzia baten mende dago

1. Sarrera

Has nadin egitasmo honen eragileei eskerrak ematen bertan parte hartzeko parada eta aukera emateagatik. Ohore handitzat daukat Jose Ignacio Hualderen omenez nire ekarpena egin ahal izatea.

Jose Ignacio 1987. urtean ezagutu nuen Iruñean, UEUko linguistika sailak azentuaren gainean antolatu zituen jardunaldi batzuetan (UEUko Linguistika Saila 1987); orduan azentuerak, tonua, *pitch* eta enparauak lehenengoz entzun nituen. Gerora, 1991. urtean, irakasle eduki nuen doktorego ikastaro haietako batean; ikasleok, filologia desberdinetakoak, asko genekielako uste genuen, nik aitortu behar dut han entzundakoak beste galaxia batekoak begitandu zitzaizkidalako. Momentuan ez nuen gauza handirik ulertu, baina ikastaroaren edukiek eta emateko erak liluratu ninduten. Nazioartean egiten zena eta teoriaren antapararen konportak diren bestean zabaldu zituen gure aurrean; uholde hark sorgindu eta erabaki nuen aurrerantzean hortik joango zela nire bidea.

Irakasleak asko etorri eta joan egiten dira, bakar batzuek, baina, markatzen gaituzte eta irauten dute gure gogo bihotzetan; horiek maisuak dira eta horietakotzat daukat Jose Ignacio.

Gauza jakinak dira bere jakituria, buru argitasuna eta lan egiteko gaitasun ikaragarriak, nik hemen nabarmendu nahi nituzke bere apaltasuna eta eskolak eta hitzaldietan erakutsitako asioa.

Urteen joan etorrian, ahal izan dudana ginoan, segitu dut haren lan eskerga; nire iritzian euskal fonologian eta bereziki azentuan eta prosodian egineko ekarpenak ezin gaituztekoak izango dira luzaroan.

Orain, urteen zamak eta pandemiak abailduta eta gaztetako grinak apalduta, zer-bait aukeratu behar izan dut adiskideari idazteko; mendebaldeko euskararen barrietate bat eta ahoskerarekin lotutakoa izan behar izan da ezinbestean; oroimena lausoa da baina oker ez banago, bere euskararen gaineko hastapeneko lanak mendebaldean egin zituen (Hualde 1986, 1987, 1989). Balioko ahal du ahal izan dudantxo honek nire miresmenaren testigantzat!

Nire lantxo honen helburua xumea da, testu labur askeen bidez lortutako ahoskeraren gaineko datuak erabilia, arlo honetan egon daitekeen dibertsitatea agirian jarri nahi da eta beronen azterketa kuantitatiboaren bidez jakin nahi da azpian leutzakeen egiturak azaleratu litezkeen ala ez.

Gure artean informatzaileak edo hizkerak sailkatzeko metodo kuantitatiboak erabili diren gehienetan (ikus besteak beste Aurrekoetxea *et al.* 2019) aukeratu diren ezaugarriak galdekizun zehatzen bidez jaso dira eta denek balio bera eduki dute, hau da, ez da kontuan hartu bakoitzaren maiztasuna. Sailkapenerako matrizeak dikotomikoak izaten dira, hots, ezaugarria erabiltzen da bai (1) edo ez (0).

Hemen urrats batzuk nahi ditugu harago egin; batetik, ez da inkestarik egin datuak jasotzeko, testu askeen informazioa erabili da matrizeak eraikitzeke eta, bestetik, ezaugarrien (fonemak gure kasuan) maiztasunak kontuan hartu dira, euren arteko eragina islatzeko.

Lana lau ataletan banatuta aurkezten dugu; sarrera honen osteko bigarren atalean datuak jasotzeko, prozesatzeko eta aztertzeko erabili dugun metodologia azalduko dugu; hirugarren atalean datuen azterketa aurkeztuko da eta, azkenik, laugarren atalean ondorio nagusiak laburbilduko ditugu.

2. Metodologia

Atal honetan datuak jasotzeko, prozesatzeko eta aztertzeko erabili den metodologiak jardungo dugu. Testuak Bizkaiko Arratia-Nerbioi izeneko eskualdeko hamabi herritan jaso dira: Galdakao, Bedia, Lemoa, Igorre, Dima, Arantzazu, Artea, Areatza, Zeanuri, Orozko, Zeberio eta Zollo (ikus bedi 1. irudiko mapa). Herri bakoitzeko neska bat eta mutil bat erabili dira. Informatzaileek euskara familia transmisioaren bidez jaso dute eta euron ana hizkuntza izan da; bestalde, informatzaile guztiak 1981 eta 2000 urte bitartean jaiotakoak izan dira.



1. irudia

Laginak jaso diren eskualde honetako herrien kokaguneak

Testuak grabatzeko sei binetako komiki bat erabili da (Gaminde 2010), komikia informatzaileei eman eta istorioa ulertzeko denbora tarte txiki baten ostean, bertan agertzen zena kontatzeko eskatu zitzaizkien grabatzeko.

Behin datuak grabatuta, euron transkripzio fonetikoak egin da; transkripzioa egiteko audio seinaleak kontuan hartu dira eta *Praat* izeneko programa erabili da (Boersma & Weenink 2019). 2. irudian transkripzioaren adibide bat erakusten da bertsiostandarrago batekin batera.

<p>komiki oneðan ekosteg dana ða (v) H % (eh) aite ðad bere semeayaz dau HL \$ (tx) par1ke baten (v) H \$ patoay eyogten H % (eh) eta (v) \$ okur2iteni jako ar1yaški ðet atertea (v) L \$ bieiji (v) LH % (eh) baje (v) % (eh) (tx) ar1yaški tetako momentuen kar2aka ðoa semeen ondoan ipiñteko (v) H % (eh) etar2i ðeteyaş tr1obeşau iteñ da H \$ eta (v) % (eh) xauşteñ da (v) L \$ uretara HL \$ eta (v) \$ momentor2eðan ar1yaški (v) L % (eh) ayeñ1tañ da H \$ ðaunia ðeyire H \$ etajite HL \$ ða (v) % (eh) ujetan HL \$ (tx) (eh) xauşite L %</p> <p>komiki honetan ikusten dena da, aita bat bere semeagaz dago, parke baten, patoak ikusten, eta, okurritzen zaio argazki bat ateratzea, bieji, baina, argazkia ateratzeko momentuan karraka ðoa semearen ondoan ipintzeko, eta harri bategaz tropezatu egiten da, eta, jausten da, uretara, eta, momentu horretan argazkia, agertzen da, ba, umea begira, eta aita, ba, uretan, jausita.</p>

2. irudia

Transkripzio ereduak

Behin transkripzio fonetikoak edukita azterketa fonologikoa egin da fonema bakoitzaren maiztasunak kalkulatu ahal izateko eta matrizeak egiteko (ikusi 5. irudiko matrizea adibidetzat).

Fonemen maiztasunen matrizeak edukita datuen azterketa kuantitatiboa egiteko estatistiko batzuk erabili dira. Batetik, informatzaileen sailkapena egiteko multzokatze azterketa hierarkikoa (MAH) eta K batezbestekoen multzokatzea (KM) izeneko teknikak erabili dira. Teknika hauen oinarrian matrize dikotomikoak egon daitezke (Aurrekoetxea *et al.* 2019) edo arrazoizkoak edo eskalakoak. Aldagai dikotomikoek ezaugarri baten erabilera edo erabilera eza baino ez dute adierazten; haatik, eskalako aldagaiek ezaugarri bakoitzaren probabilitatea adierazten dute eta berori kontuan hartuta egiten da sailkapena.

MAH multzokatzeak balio du datuak zelan multzokatzen diren eta zelan hierarkizatzen diren ikusteko. Multzokatze hierarkikoa egiteko, kasu guztietan *Ward* metodoa erabili da distantzia euklidearrarekin; egia da metodo asko daudena (Pardo & Ruiz 2002; Pérez 2005) baina bariazioaren azterketetan gailendu dena *Ward* metodoa izan da (Goebel 1992, 2010; Aurrekoetxea 1995, 2009; Gaminde 2007; Gaminde *et al.* 2017). *Ward* metodoak aldeak eta kidetasunak maximizatzen dituelako izan da aukeratua gure hemengo lan honetarako. Distantzia linguistiko batzuk dauden arren, distantzia euklidearra aukeratu dugu konputatzeko unibertsalagoa eta zuzenagoa delako. MAH multzokatzearen emaitza dendrograma bat izaten da eta bertan multzo nagusiak agertzen dira; elementu bakoitza multzo bakarrean agertuko da gainontzeko elementuekiko daukan distantziaren arabera.

KM multzokatzea erabilia gauza bera gertatzen da, baina kasu honetan elementuen banaketa egingo da ikertzaileak aurretik aukeratutako multzo kopuruaren arabera. Teknika honek elementuen arteko antzekotasunak eta aldeak nabarmentzen ditu.

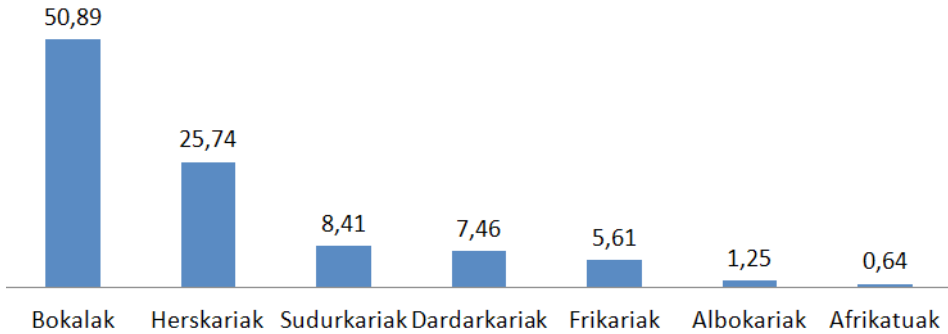
Fonemen maiztasunetan informatzaileen generoaren arabera egon daitezkeen aldeak estatistikoki esanguratsuak diren ala ez aztertzeke Mann-Whitneyren U izeneko estatistiko ez-parametrikoa (Pérez 2008) erabiliko dugu. Informatzaileen sailkapenean euron generoak eragina daukan aztertzeke erabiliko den estatistikoa Khi karra-tua izango da (χ^2). Proba hauetan guztietan emaitzak esanguratsuak diren ala ez adieraztean erabili dugun « α » 0,05 izan da.

Sailkapen horiez gain, interesgarri iritzi diogu ezaugarrien azalpen indarra aztertzeke Shannonen entropia (1948) indizearen bidez (Roberto *et al.* 2012; Wojciechowski & Gorzynski 2016). Kalkulu hau egin ahal izateko PAST izeneko programa erabili dugu (Hammer *et al.* 2001).

3. Datuen azterketa

Jaso ditugun testuetan 9069 soinu agertu dira, azterketa burututa 25 fonema dira denetara. Bokalak dira gehien erabiltzen diren fonemak, denetara 4615 izan dira, hau da, corpus osoaren % 50,89; gehien erabiltzen dena alde handiarekin /a/ da eta gutxien, berriz, /u/. Herskariak 2334 izan dira, fonema guztien % 25,74; gehien erabiltzen dena /t/ da eta gutxien erabiltzen dena, berriz, /j/. Sudurkariak denetara 763 izan dira, fonema guztien % 8,41; gehien erabil-

tzen dena /n/ da, gero /m/ eta gutxien /ɲ/. Frikariak 509 izan dira, fonema guztien % 5,61; gehien erabiltzen dena alde handiz /s/ da. Afrikatuak denetara 58 izan dira, fonema guztien % 0,64. Dardarkariak 677 izan dira, fonema guztien % 7,46. Azkenik albokariak 113 izan dira % 1,25 (3. irudian sail bakoitzeko ehunekoak ematen dira).



3. irudia

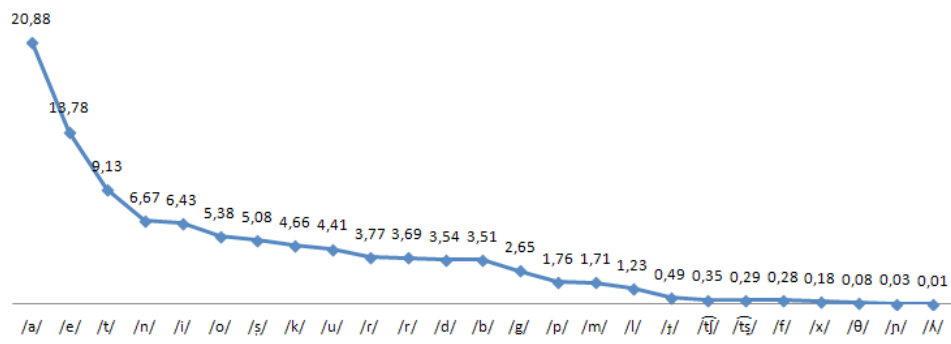
Sail bakoitzeko ehunekoak

1. taulan fonema bakoitzaren kopuruak eta ehunekoak ematen dira handienetik txikienera ordenaturik (4. irudiko grafikoa). Bertan ikusten den moduan, maiztasunik handiena /a/ fonemak dauka eta txikiena /k/-k.

1. taula

Fonema guztien kopuruak eta ehunekoak

Fonema	Kopurua	%	Fonema	Kopurua	%
/a/	1894	20,88	/g/	240	2,65
/e/	1250	13,78	/p/	160	1,76
/t/	828	9,13	/m/	155	1,71
/n/	605	6,67	/l/	112	1,23
/i/	583	6,43	/j/	44	0,49
/o/	488	5,38	/tʃ/	32	0,35
/s/	461	5,08	/tʃ̃/	26	0,29
/k/	423	4,66	/f/	25	0,28
/u/	400	4,41	/x/	16	0,18
/r/	342	3,77	/θ/	7	0,08
/r/	335	3,69	/ɲ/	3	0,03
/d/	321	3,54	/k/	1	0,01
/b/	318	3,51			



4. irudia

Fonemen ehunekoak ordenaturik

Informatzaileen sailkapen hierarkikoa egin ahal izateko fonema bakoitzaren ko-
 puruekin matrize bat eraiki dugu eta berorren gainean fonema bakoitzaren probabi-
 litateak kalkulatu dira bakoitzeko, 5. irudian agertzen den matrize zatiak erakusten
 duen bezala.

Lekukoa	a	e	i	o	u	p	t	k	b	d	j	g	m
GaldakaoAa	0,213	0,119	0,072	0,052	0,052	0,015	0,081	0,048	0,026	0,051	0,005	0,019	0,014
GaldakaoAg	0,185	0,156	0,061	0,069	0,046	0,023	0,075	0,046	0,023	0,049	0,000	0,020	0,017
ZolloAa	0,183	0,125	0,083	0,100	0,025	0,025	0,075	0,042	0,033	0,017	0,008	0,050	0,008
ZolloAg	0,218	0,130	0,063	0,042	0,049	0,011	0,092	0,035	0,035	0,042	0,011	0,039	0,018
OrozkoAa	0,185	0,167	0,069	0,058	0,033	0,015	0,113	0,055	0,036	0,029	0,004	0,025	0,029
OrozkoAg	0,179	0,164	0,057	0,061	0,064	0,021	0,089	0,043	0,025	0,036	0,000	0,014	0,018
BediaAa	0,181	0,145	0,072	0,041	0,054	0,016	0,088	0,052	0,032	0,027	0,002	0,034	0,014
BediaAg	0,232	0,143	0,055	0,042	0,032	0,016	0,073	0,053	0,036	0,018	0,005	0,028	0,018
LemoaAa	0,216	0,133	0,049	0,038	0,066	0,019	0,074	0,047	0,057	0,045	0,004	0,036	0,013
LemoaAg	0,180	0,139	0,065	0,071	0,041	0,020	0,095	0,051	0,017	0,034	0,010	0,024	0,031
IgorreAa	0,225	0,141	0,068	0,045	0,047	0,013	0,110	0,031	0,029	0,039	0,003	0,026	0,016
IgorreAg	0,221	0,153	0,058	0,043	0,035	0,025	0,105	0,050	0,040	0,030	0,003	0,020	0,023
DimaAa	0,214	0,149	0,065	0,048	0,033	0,024	0,098	0,054	0,015	0,036	0,003	0,030	0,018
DimaAg	0,213	0,129	0,056	0,067	0,028	0,017	0,101	0,070	0,042	0,034	0,003	0,028	0,025
ArantzazuAa	0,195	0,126	0,048	0,084	0,060	0,015	0,102	0,018	0,039	0,036	0,006	0,024	0,015
ArantzazuAg	0,231	0,146	0,058	0,045	0,040	0,020	0,101	0,045	0,058	0,036	0,002	0,020	0,016
ZeberioAa	0,213	0,130	0,079	0,065	0,032	0,022	0,119	0,029	0,043	0,040	0,011	0,029	0,004
ZeberioAg	0,179	0,138	0,082	0,052	0,057	0,016	0,098	0,045	0,032	0,050	0,005	0,020	0,014

5. irudia

Informatzaileen probabilitateak fonema bakoitzeko

Balio handien efektua ekiditeko matrizearen normalizazioa burutu dugu erro ka-
 ratuaren bidez; 6. irudian matrize honen zati bat erakusten da.

Lekukoa	a	e	i	o	u	p	t	k	b	d	j	g	m	n
GaldakaoAa	0,462	0,345	0,269	0,228	0,228	0,123	0,284	0,219	0,161	0,225	0,074	0,138	0,117	0,286
GaldakaoAg	0,430	0,395	0,246	0,263	0,215	0,152	0,274	0,215	0,152	0,222	0,000	0,142	0,132	0,274
ZolloAa	0,428	0,354	0,289	0,316	0,158	0,158	0,274	0,204	0,183	0,129	0,091	0,224	0,091	0,316
ZolloAg	0,467	0,361	0,252	0,206	0,222	0,103	0,303	0,188	0,188	0,206	0,103	0,197	0,133	0,278
OrozkoAa	0,431	0,409	0,263	0,241	0,181	0,121	0,336	0,234	0,191	0,171	0,060	0,160	0,171	0,263
OrozkoAg	0,423	0,405	0,239	0,246	0,254	0,146	0,299	0,207	0,158	0,189	0,000	0,120	0,134	0,231
BediaAa	0,425	0,381	0,269	0,202	0,233	0,126	0,297	0,228	0,178	0,165	0,048	0,184	0,117	0,269
BediaAg	0,481	0,378	0,235	0,204	0,180	0,127	0,270	0,229	0,191	0,132	0,073	0,168	0,132	0,252
LemoaAa	0,465	0,364	0,222	0,195	0,257	0,138	0,272	0,218	0,238	0,213	0,062	0,190	0,115	0,257
LemoaAg	0,425	0,373	0,254	0,267	0,202	0,143	0,309	0,226	0,130	0,184	0,101	0,154	0,175	0,274
IgorreAa	0,474	0,376	0,261	0,211	0,217	0,114	0,332	0,177	0,170	0,198	0,051	0,162	0,125	0,234
IgorreAg	0,470	0,391	0,240	0,206	0,187	0,158	0,324	0,224	0,200	0,173	0,050	0,142	0,150	0,250
DimaAa	0,463	0,386	0,256	0,218	0,181	0,154	0,313	0,231	0,122	0,189	0,055	0,173	0,134	0,262
DimaAg	0,462	0,359	0,237	0,260	0,168	0,130	0,318	0,265	0,205	0,184	0,053	0,168	0,159	0,237
ArantzazuAa	0,442	0,355	0,219	0,290	0,245	0,123	0,320	0,134	0,198	0,190	0,077	0,155	0,123	0,263
ArantzazuAg	0,481	0,382	0,241	0,212	0,201	0,142	0,318	0,212	0,241	0,189	0,047	0,142	0,125	0,227
ZeberioAa	0,462	0,361	0,282	0,255	0,180	0,147	0,345	0,170	0,208	0,199	0,104	0,170	0,060	0,225
ZeberioAg	0,423	0,372	0,286	0,228	0,238	0,126	0,312	0,213	0,178	0,223	0,067	0,143	0,117	0,261

6. irudia

Erro rarratuaren bidez normalizatutako matrizearen zatia

Informatzaileen sailkapena egiteko euron arteko distantzien matrizea eraikitzen da 7. irudian agertzen den bezala. Bertan ikus daitekeen bezala distantziarik handiena (0,309) ZolloAa eta ArantzazuAa informatzaileen artean gertatzen da eta txikiena (0,100) berriz ArantzazuAg eta IgorreAg informatzaileen artean.

	GaldakaoAa	GaldakaoAg	ZolloAa	ZolloAg	OrozkoAa	OrozkoAg	BediaAa	BediaAg	LemoaAa	LemoaAg	IgorreAa	IgorreAg	DimaAa	DimaAg	ArantzazuAa	ArantzazuAg
GaldakaoAa	0,000	0,190	0,162	0,224	0,205	0,131	0,188	0,173	0,188	0,122	0,189	0,141	0,196	0,243	0,190	0,190
GaldakaoAg	0,190	0,000	0,270	0,230	0,217	0,166	0,159	0,197	0,204	0,160	0,219	0,171	0,191	0,185	0,266	0,205
ZolloAa	0,234	0,270	0,000	0,220	0,238	0,301	0,238	0,251	0,268	0,224	0,267	0,252	0,214	0,229	0,309	0,261
ZolloAg	0,162	0,230	0,220	0,000	0,175	0,233	0,182	0,167	0,151	0,165	0,157	0,172	0,138	0,164	0,256	0,178
OrozkoAa	0,224	0,217	0,238	0,175	0,000	0,216	0,218	0,205	0,211	0,150	0,220	0,199	0,170	0,164	0,273	0,213
OrozkoAg	0,205	0,166	0,301	0,233	0,216	0,000	0,186	0,198	0,176	0,187	0,202	0,187	0,205	0,204	0,244	0,180
BediaAa	0,131	0,159	0,238	0,182	0,218	0,186	0,000	0,149	0,167	0,166	0,149	0,167	0,151	0,188	0,248	0,182
BediaAg	0,188	0,197	0,251	0,167	0,205	0,198	0,149	0,000	0,167	0,170	0,190	0,145	0,169	0,158	0,286	0,171
LemoaAa	0,173	0,204	0,268	0,151	0,211	0,176	0,167	0,167	0,000	0,201	0,174	0,159	0,194	0,169	0,236	0,124
LemoaAg	0,188	0,160	0,224	0,165	0,150	0,187	0,186	0,170	0,201	0,000	0,210	0,154	0,150	0,138	0,270	0,203
IgorreAa	0,122	0,219	0,267	0,157	0,220	0,202	0,149	0,190	0,174	0,210	0,000	0,170	0,136	0,186	0,236	0,164
IgorreAg	0,189	0,171	0,252	0,172	0,193	0,187	0,167	0,145	0,159	0,154	0,170	0,000	0,157	0,101	0,287	0,100
DimaAa	0,141	0,191	0,214	0,138	0,170	0,205	0,151	0,169	0,194	0,150	0,136	0,157	0,000	0,151	0,261	0,186
DimaAg	0,196	0,185	0,229	0,164	0,164	0,204	0,188	0,158	0,169	0,138	0,186	0,101	0,151	0,000	0,281	0,194
ArantzazuAa	0,243	0,266	0,309	0,256	0,273	0,244	0,248	0,286	0,236	0,270	0,236	0,287	0,261	0,281	0,000	0,227
ArantzazuAg	0,190	0,205	0,261	0,178	0,213	0,180	0,182	0,171	0,124	0,203	0,164	0,100	0,186	0,134	0,237	0,000
ZeberioAa	0,179	0,244	0,244	0,180	0,226	0,261	0,195	0,223	0,216	0,229	0,149	0,211	0,181	0,205	0,236	0,197
ZeberioAg	0,162	0,173	0,249	0,156	0,209	0,162	0,164	0,190	0,145	0,153	0,171	0,137	0,170	0,155	0,232	0,125
ArteAa	0,217	0,302	0,209	0,237	0,293	0,289	0,234	0,230	0,261	0,255	0,246	0,259	0,233	0,252	0,247	0,233
ArteAg	0,193	0,156	0,245	0,212	0,210	0,187	0,165	0,154	0,183	0,168	0,206	0,124	0,186	0,117	0,303	0,159
ArantzazuAa	0,297	0,171	0,256	0,182	0,170	0,177	0,194	0,188	0,173	0,170	0,184	0,123	0,165	0,097	0,274	0,139
ArantzazuAg	0,196	0,199	0,267	0,176	0,223	0,207	0,181	0,144	0,156	0,168	0,185	0,125	0,195	0,160	0,284	0,154
ZeauriAa	0,187	0,237	0,287	0,203	0,250	0,234	0,191	0,241	0,208	0,237	0,174	0,233	0,181	0,238	0,161	0,207
ZeauriAg	0,163	0,206	0,269	0,151	0,213	0,213	0,147	0,177	0,173	0,214	0,148	0,192	0,150	0,190	0,198	0,167

7. irudia

Distantzia matrizea

Informatzaileen azterketa hierarkikoa eginez gero, 8. irudiko dendrograma eta sailkapena lortzen ditugu.



A: GaldakaoAg, OrozkoAa, OrozkoAg, BediaAg, LemoaAa, LemoaAg, IgorreAg, DimaAg, ArantzazuAg, ZeberioAg, ArteaAg, AreatzaAa eta AreatzaAg.

B: GaldakaoAa, ZolloAa, ZolloAg, BediaAa, IgorreAa, DimaAa, ArantzazuAa, ZeberioAa, ArteaAa, ZeanuriAa eta ZeanuriAg.

8. irudia

Informatzaileen sailkapen hierarkikoa fonemen arabera

Generoari dagokionez bertan ikus dezakegu bost herritako informatzaile biak agertzen direla batera multzo berean; Orozko, Lemoa eta Areatza A multzoan eta Zollo eta Zeanuri B multzoan. Gainerako herri guztietako informatzaileak multzo banatan sailkatuta agertzen dira generoka; 2. taulan informatzaileak euron generoaren arabera multzoetan zelan banatzen diren erakusten da; bertan agertzen den aldea estatistikoki esanguratsua da ($K_{hi} = 8,224$ (a.m: 1), $p = 0,004$).

2. taula

Informatzaileen banaketa multzoetan euron generoaren arabera

	A	B	Guztira
Neskak	3	9	12
Mutilak	10	2	12
Guztira	13	11	24

Nesken zein mutilen datuak bateratzen baditugu herrika eta KM multzokatzea egiten badugu multzo biren arabera, A multzoan zazpi herri kokatzen dira (Galdakao, Zeberio, Igorre, Arantzazu, Orozko, Areatza eta Zeanuri) eta B multzoan bost (Zollo, Bedia, Lemoa, Dima eta Artea); 9. irudiko mapan agertzen dira emaitzak proiektatuta.



9. irudia

Saillapenen emaitzak mapara proiektatuta

Hona helduta, generoaren araberako aldeak non gertatzen diren ikusi beharko genuke, horretarako sail bakoitzeko fonemen batezbestekoak aztertuko ditugu informatzaileen generoaren arabera. Tauletan batezbestekoekin batera desbiderapenak (S) eta Mann-Whitneyren U proba ez parametrikoren Z eta p estatistikoen balioak ematen dira.

3. taulan bokalen datuak ematen dira informatzaileen generoaren arabera; bertan ikusten den bezala, aldeak behin ere ez dira estatistikoki esanguratsuak.

3. taula

Bokalen datuak informatzaileen generoaren arabera

Fonema	Neskak		Mutilak		Z	p
	\bar{x}	S	\bar{x}	S		
/i/	0,066	0,012	0,064	0,009	-0,723	0,470
/e/	0,138	0,018	0,139	0,015	-0,260	0,795
/a/	0,204	0,015	0,209	0,023	-0,637	0,524
/o/	0,061	0,019	0,053	0,012	-1,012	0,312
/u/	0,044	0,013	0,042	0,010	-0,318	0,751

4. taulan herskarien datuak ematen dira informatzaileen generoaren arabera; bertan erakusten den legez, aldeak ez dira estatistikoki esanguratsuak.

4. taula

Herskarien datuak informatzaileen generoaren arabera

Silaba	Neskak		Mutilak		Z	p
	\bar{x}	S	\bar{x}	S		
/p/	0,018	0,004	0,018	0,005	-0,550	0,582
/t/	0,094	0,016	0,092	0,010	-0,404	0,686
/k/	0,044	0,013	0,049	0,012	-0,491	0,623
/b/	0,034	0,010	0,035	0,010	-0,203	0,839
/d/	0,034	0,011	0,036	0,009	-0,058	0,954
/j/	0,006	0,006	0,004	0,004	-1,078	0,281
/g/	0,029	0,009	0,026	0,007	-0,811	0,417

5. taulan sudurkarien datuak erakusten dira informatzaileen generoaren arabera; /m/ sudurkariarekin agertzen den aldea estatistikoki esanguratsua den bakarra da.

5. taula

Sudurkarien datuak informatzaileen generoaren arabera

Fonema	Neskak		Mutilak		Z	p
	\bar{x}	S	\bar{x}	S		
/m/	0,014	0,007	0,021	0,005	-2,467	*0,014
/n/	0,070	0,013	0,065	0,008	-1,013	0,311
/ɲ/	0,001	0,001	0,000	0,000	-1,808	0,071

6. taulan frikarien datuak agertzen dira informatzaileen generoaren arabera; bertan erakusten den moduan, aldeak ez dira estatistikoki esanguratsuak.

6. taula

Frikarien datuak informatzaileen generoaren arabera

Fonema	Neskak		Mutilak		Z	p
	\bar{x}	S	\bar{x}	S		
/s/	0,051	0,013	0,051	0,008	-0,463	0,644
/f/	0,004	0,008	0,002	0,002	-0,275	0,783
/x/	0,001	0,002	0,002	0,003	-0,884	0,377
/θ/	0,001	0,001	0,001	0,002	0,000	1,000

7. taulan afrikatuen datuak ematen dira informatzaileen generoaren arabera; bertan erakusten diren aldeak estatistikoki esanguratsuak dira. Mutilek sabaikari-aurrekoa askoz gehiagotan erabiltzen dute; neskek, oster, laminarra gehiagotan erabiltzen dute mutilek baino; datu hauek bat datoz Igorren agertutakoekin (Eguskiza & Gaminde 2022).

7. taula

Afrikatuen datuak informatzaileen generoaren arabera

Fonema	Neskak		Mutilak		Z	p
	\bar{x}	S	\bar{x}	S		
/tʃ/	0,001	0,001	0,006	0,004	-3,345	*0,001
/tʃ̃/	0,005	0,006	0,001	0,002	-2,519	*0,012

8. taulan dardarkarien datuak ikus daitezke informatzaileen generoaren arabera; bertan agertzen den moduan, aldeak ez dira estatistikoki esanguratsuak.

8. taula

Dardarkarien datuak informatzaileen generoaren arabera

Fonema	Neskak		Mutilak		Z	p
	\bar{x}	S	\bar{x}	S		
/r/	0,038	0,005	0,036	0,006	-1,333	0,182
/r/	0,032	0,009	0,040	0,008	-1,878	0,060

Azkenik, 9. taulan albokariaren datuak erakusten dira informatzaileen generoaren arabera; bertan ikusten den bezala, aldeak ez dira estatistikoki esanguratsuak.

9. taula

Albokariaren datuak informatzaileen generoaren arabera

Fonema	Neskak		Mutilak		Z	p
	\bar{x}	S	\bar{x}	S		
/l/	0,011	0,008	0,011	0,005	-0,116	0,908
/k/	0,000	0,000	0,000	0,001	-1,000	0,317

4. Ondorioak

Azken atal honetan datuen azterketan lortu ditugun ondorio nagusiak laburbilduko ditugu. Informatzaileen sailkapen hierarkikoa egitean ikusi dugun bezala, informatzaileen generoa estatistikoki esanguratsua da multzoen osaketan. Bestalde, fonemen maiztasunen azterketaren bidez ikusi dugu hiru fonemaren maiztasunak baino ez direla estatistikoki esanguratsuak informatzaileen generoaren arabera: /m/ sudurkaria eta /tʃ/ eta /tʃ̃/ afrikatuak. Hauek, beste batzuen aldean, maiztasun urrikoak dira eta ondorioz, Shannonen indizearen arabera balio handiak dauzkate entropian; horrela, bada, /a/ fonemaren entropia, maiztasunik handiena daukana, 5,5 da eta /m/ sudurkariarena 30,67, /tʃ̃/-rena 68,23 eta /tʃ̃/-rena 72,99.

Azkenik, esan dezakegu hemen erakutsi ditugun teknika hauek baliagarriak izan daitezkeela azaleratzeko azpian gertatzen diren fenomeno batzuk. Aurkeztu dugun ahoskeraren kasu honetan, lagin txikiakin ere, fonemen arteko maiztasunaren bidez lor daitezke emaitza interesgarriak.

Bibliografia

- Aurrekoetxea, Gotzon. 1995. *Bizkaieraren egituraketa geolinguistikoa*. Leioa: Argitarapen Zerbitzua, Euskal Herriko Unibertsitatea.
- Aurrekoetxea, Gotzon. 2009. Bariazio sinkronikoak aztertzeko metodologiak, *Lapurdum* 13. 43-59.
- Aurrekoetxea, Gotzon, Iñaki Gaminde, José Luis Ormaetxea & Xarles Videgain. 2019. *Euskalkien sailkapen berria*. Bilbo: UPV/EHU.
- Boersma, Paul & David Weenink. 2019. *Praat: doing phonetics by computer*. <https://www.fon.hum.uva.nl/praat/>.
- Eguskiza, Naia & Iñaki Gaminde. 2022. *Ahoskera Igorren*. Igorre: Igorreko Udala.
- Gaminde, Iñaki. 2007. *Bizkaian Zehar: Euskararen Ikuspegi Orokorra*. Bilbo: Mendebalde Kultura Alkartea & Bizkaiko Foru Aldundia.
- Gaminde, Iñaki. 2010. *Bizkaiko gazteen prosodiaz: euskaraz eta gaztelaniaz*. Bilbo: Mendebalde Kultura Alkartea & Bizkaiko Foru Aldundia.
- Gaminde, Iñaki, Naia Eguskiza, Aitor Iglesias & Lorea Unamuno. 2017. Sexuaren eragina Igorreko afrikatu bizkarkariaren tasun akustikoetan. *Euskalingua* 30. 6-12.

- Goebel, Hans. 1992. Problèmes et méthodes de la dialectométrie actuelle (avec application à l' AIS). In Gotzon Aurrekoetxea & Xarles Videgain (arg.), *Nazioarteko Dialektologia Biltzarra (IKER 7)*, 429-475. Bilbo: Euskaltzaindia.
- Goebel, Hans. 2010. Introducción a los problemas y métodos según los principios de la Escuela Dialectométrica de Salzburgo. In Gotzon Aurrekoetxea & José Luis Ormaetxea (arg.), *Tools for Linguistic Variation (ASJUren Gehigarriak 53)*, 3-39. Bilbo: UPV/EHU.
- Hammer, Øyvind, David A. T. Harper & Paul D. Ryan. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 4(1). 1-9. http://palaeo-electronica.org/2001_1/past/issue1_01.htm.
- Hualde, José Ignacio. 1986. Tone and Stress in Basque: A Preliminary Study. *ASJU* 20(3). 867-896.
- Hualde, José Ignacio. 1987. A Theory of Pitch-Accent, with Particular Attention to Basque. Eusko Jaurlaritza. *II. Mundu-Biltzarra*. 1. liburukia, 53-60.
- Hualde, José Ignacio. 1989. Acentos vizcaínos. *ASJU* 23(1). 275-326.
- Pardo, Antonio & Miguel Ángel Ruiz. 2002. *SPSS 11. Guía para el análisis de datos*. Madrid: McGraw Hill.
- Pérez, César. 2005. *Métodos estadísticos avanzados con SPSS*. Madrid: Thomson.
- Pérez, César. 2008. *Técnicas Estadísticas con SPSS 12: Aplicaciones al análisis de datos*. Madrid: Pearson Educación S.A.
- Roberto, John A., M. Antònia Martí & Maria Salamó. 2012. Análisis de la riqueza léxica en el contexto de la clasificación de atributos demográficos latentes. *Procesamiento de Lenguaje Natural* 48. 97-104.
- UEUko Linguistika Saila. 1987. *Euskara Batua: Iruñeko Proposamena Azentuari Buruz*. UEU.
- Wojciechowski, Adam & Krzysztof Gorzynski. 2016. A Method for Measuring Similarity of Books. In Zygmunt Vetulani, Hans Uszkoreit & Marek Kubis. *Human Language Technology. Challenges for Computer Science and Linguistics: 6th Language and Technology Conference, LTC 2013, Poznań, Poland, December 7-9, 2013, Revised Selected Papers*, 161-176. Berlin: Springer. https://doi.org/10.1007/978-3-319-43808-5_13.

