

Euskararen erabilera Eusko Legebiltzarreko debateetan (2012-2020)

(The use of Basque language in debates in the Basque Parliament (2012-2020))

Escribano Garcia, Nayla¹; González, Jon Ander²; Orbegozo-Terradillos, Julen³; Larrondo-Ureta, Ainara⁴; Peña-Fernández, Simón⁵; Perez-de-Viñaspre, Olatz⁶; Agerri Gascón, Rodrigo⁷

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV-EHU). Sarriena s/n. 48940. Leioa.

nayla.escribano@ehu.eus¹; julen.orbegozo@ehu.eus³;

ainara.larrondo@ehu.eus⁴; simon.pena@ehu.eus⁵;

olatz.perezdevinaspre@ehu.eus⁶; rodrigo.agerri@ehu.eus⁷

Recep.: 10.03.2022

BIBLID [eISSN 1988-3935 (2022), 19; 163-178]

Acep.: 01.12.2022

Legebiltzarretan, deliberazio demokrazien testuinguruan egindako hitzaldien azterketa garrantzitsua da, hitzaldi horiek hertsiki lotuta baitaude ekintza politikoarekin eta legegintza ekimenak garatzeko arrazoien azalpenarekin. Bestalde, azken urteotan, azterketa automatizatuei esker gehitu egin dira diskurtsoa aztertzeko aukerak, eta informazio bolumen handiak prozesatu daitezke. Testuinguru horretan, euskarak Eusko Legebiltzarrean duen erabilera aztertzen da artikulu honetan, bi legegintzaldiko (hots, 2012-2016 eta 2016-2020 legegintzaldietako) bilkuren akten azterketa automatizatuan oinarrituta. Emaitzek adierazten dutenez, euskaraz egindako agerraldien guztizko bolumena eguneroko erabilerari dagokiona baino handixeagoa da, baina adierazi beharra dago euskarari agerikotasuna emateko helburua duten protokolo erabilerak eta erabilera sinbolikoak ere identifikatu direla.

Gako hitzak: Legebiltzarra. Elebitasuna. Euskara. Hizkuntza naturalaren prozesatzea. Euskadi.

El análisis de las intervenciones parlamentarias es relevante en el contexto de las democracias deliberativas, pues están íntimamente relacionadas con la acción política y con la exposición de las razones por las que se desarrollan las iniciativas legislativas. Por otro lado, en los últimos años, los análisis automatizados de los textos han permitido aumentar las posibilidades de estudiar los discursos y procesar grandes volúmenes de información. En este contexto, este artículo analiza el uso del euskera en el Parlamento Vasco a partir del análisis automatizado de las actas de las sesiones de dos legislaturas (2012-2016 y 2016-2020). Los resultados indican que el volumen total de intervenciones realizadas en euskera es algo mayor que su uso cotidiano, y también se han identificado usos protocolarios y simbólicos que tienen por objeto dar visibilidad al euskera.

Palabras clave: Parlamento. Bilingüismo. Euskera. Procesamiento de la lengua natural. Euskadi.

The analysis of interventions in parliament, in the context of deliberative democracies, is relevant because they are closely linked to political action and to the exposition of the reasons why legislative initiatives are developed. For its part, in recent years, automated content analysis has made it possible to process large volumes of information. In this context, this article studies the use of Basque in the Basque Parliament based on the automated analysis of the minutes of the sessions during two legislatures (2012-2016 and 2016-2020). The results indicate that, in the total volume of interventions, those carried out in Basque are slightly above their daily use, although tendencies towards their formal and symbolic use that seek to make their presence visible are also identified.

Keywords: Parliament. Bilingualism. Euskara. Natural language processing. Euskadi.

L'analyse des interventions au sein du Parlement, dans le contexte des dénommées démocraties délibératives, s'avère particulièrement intéressante parce qu'elles sont étroitement liées à l'action politique et à l'exposition des motifs qui sont à l'origine du développement des initiatives législatives. Pour leur part, les analyses automatisées ont permis durant les dernières années d'accroître les possibilités d'étude du discours, en permettant le traitement de gros volumes d'information. Dans ce contexte, cet article étudie l'utilisation de la langue basque au sein du Parlement basque à partir de l'analyse automatisée des comptes-rendus des sessions de deux législatures (2012-2016 et 2016-2020). Les résultats indiquent que le volume total d'interventions en basque est légèrement supérieur à son utilisation quotidienne, même si on identifie aussi des tendances à son emploi protocolaire et symbolique qui cherchent à visibiliser sa présence.

Mots clés: Parlement. Bilinguisme. Euskera. Basque. Traitement du langage naturel. Euskadi. Pays basque espagnol.

1. SARRERA

Komunikazio politikoaren esparruan, parlamentuko debateek arreta berezia jaso izan dute beti, erabakiak eraikitze prozesuan hautetsien artean sortzen den ideia konfrontazioaren eredu politikoaren adierazgarriak baitira. Ekintza politikoaren azterketa ezin da hizkuntzaren azterketatik berezi. Izan ere, hizkuntzaren azterketak tradizio erretoriko klasikoaren legatuan du jatorria (Dunmire, 2012), eta gune dialektikoetan egindako ideia konfrontazioa demokrazien funtsezko ezaugarritzat hartu izan da beti.

Bestalde, parlamentuko erretorikaren azterketari esker, eta ordezkariak nola hitz egiten duten aztertzearen bidez, gizarte bakoitza kezkatzen duten interesak eta gaiak nola bilakatzen diren eta gizartea bera nola eraldatzen den ere ikus daiteke. Erabilitako hizkuntza, debatearen xede diren gaiak eta une historiko bakoitzeko legegintza ekimenak gizarteak une jakin batean dituen kezken eta premien isla dira.

Bourdieu ekintza politikoaz ematen duen definizioari (1991) jarraikiz, parlamentuetan eragileek mezuak sortzen dituzte etengabe, eta, mezu horien bitartez, mundu sozialaren ikuspegi berezi bat eraiki eta ezartzea eta, aldi berean, azken batean beren boterearen euskarri direnen babesa mobilizatzea dute helburu (Sokolovska, 2016). Legegintza ganberak ekimenak eztabaidatzeko foroak ez ezik, ekimenen ezarpen soziala herritarrei motibazioak helaraztearen bidez defendatzeko lekuak ere badira.

Azken hamarkadetan, parlamentuetako debateak hertsiki lotuta egon dira deliberazio demokratiaren ikuskerarekin (Guttmann & Thompson, 1996). Ikuskera horren arabera, ordezkari politikoen eta herritarren arteko elkarriketa ere garrantzitsua da, erabakiak hartzeko prozesuan zein ondorengo ezarpenean. Hori dela eta, demokrazian ekintza politikoak legegintza ekimenak garatzeko arrazoiak azaltzea eta foro publikoetan argudioak elkartrukatzea dakarrenez gero, argudio horien azterketa zeregin garrantzitsua da (Bara, Weale & Biquelet, 2007).

Hori guztia dela eta, kontuan hartu behar da parlamentuko diskurtsoaren ezaugarri bereizgarria ez dela lexikoa, sintaxia edo hitzaldiaren edukia bera, elementu horiek guztiak beste formatu eta foro askotan ere agertzen baitira, hala nola komunikabideetan (Iñigo-Mora, 2007). Parlamentuko hizkuntzaren protokolo berezitasun batzuek salbuetsita, politikarien hitz egiteko modua ez da oso desberdina, funtsean, tribunan eta tribunatik kanpo. Ordezkaritza ganbera batean, diskurtso politikoaren elementu bereizgarria izaera politikoa da, hots, gizarte baten premiei aurre egitean ideiak aurrez aurre jartzeko eta ideia horiek gizarteari helarazteko modua. Era horretan, legitimatu eta, aldi berean, indartu egiten da erabakitzeko botere politikoa.

Hori guztia dela eta, informazio horren prozesatzeak eta hedapen publikoak interes berezia dute, parlamentuko jardueraren funtsa bera baitira.

Debateen azterketak gizartea gehien kezkatzen duten arazoak identifikatzea ahalbidetzen du, eta jarduera politikoaren inguruko gardentasuna eskatzen du. Debate horiek gizartera eramateko, komunikabideak erabili izan dira, eta komunikabide horiek herritarren eta klase politikoaren arteko bitartekaritza lana egin izan dute. Nolanahi ere, komunikazio tresna berriek sare sozialetara eramanez zeregin hori (Griffith & Leston-Bandeira, 2012), eta, horren ondorioz, plataforma horiek politikarientzat eremu publiko berri bat ote diren planteatu da (Kruse et al., 2018).

Parlamentuko jardueraren garrantzi eta eragin soziala dela eta, diskurtso politikoaren azterketak interes handia sortzen du ikertzaileengan. Ildo horretan, hainbat ganberatan egindako debateak aztergai izan dira hainbat herrialdetan. Adibidez, Britainia Handiko parlamentuan (House of Commons) (Salah, 2014; Abercombrie & Bastista-Navarro, 2018), Diputatuen Kongresuan (Iñigo-Mora, 2007) eta Autonomia Erkidegoetako ganberetan (hala nola Andaluziako parlamentuan) (Alcaide, Carranza & Catalina, 2016) diputatuek egindako hitzaldiei buruzko azterketak daude.

Debate corpus horiek aztertzeke azken hurbilketetako bat big datako azterketa tresnen erabilera izan da (algoritmoak, data mining eta machine learning), sentimenduaren azterketari dagokionez bereziki. Izan ere, John Hyde-k (2002) dioenez, parlamentuko diskurtsoa aztertzeke beharrezkoa da alde aurretik jakintza arloaren arteko mugak desagerraraztea. Hizkuntza naturalaren prozesatzea parlamentuko debateetan aplikatu zuten lehen azterlanak Ameriketako Estatu Batuetan egin ziren (Thomas, Pang & Lee, 2006; Burfoot, Bird & Baldwin, 2011).

Hurbilketa horiek bereziki interesgarriak dira parlamentuetako debateen kasuan, bertan egindako hitzaldi guztiek eskuz nekez prozesatu daitezkeen informazio bolumen handia biltzen baitute. Hortaz, bereziki egokiak dira azterketa tresna automatizatuekin edo erdi-automatizatuekin aztertzeke (Bara, Weale & Biquelet, 2007). Tresna horiei esker, aztertu daitezkeen hitzaldien bolumena esponentzialki gehitu daiteke, eta azterketa egiteko ikuspegi ugari eskaintzen dira.

Azken urteotan, azterketa automatizatuen esparru garrantzitsu bat diskurtso politikoaren sentimenduari eta polaritateari lotuta egon da, hau da, diskurtsoaren orientazio subjektiboa modu automatizatuan identifikatzeko moduari. Esparru publikoaren eta pribatuaren arteko mugak desagertu eta komunikazio politikoa emozioetara ireki den heinean (Richards, 2010), areagotu egin da emozioak erabiltzeko moduarekiko interesa. Diskurtsoetan agertutako polaritateari eta sentimenduari buruzko hainbat azterlan egin dira, hala nola Erresuma Batuko Komunien Ganberako diskurtsoei (Onyimadu et al., 2013; Salah, 2014; Duthie et al., 2016; Rheault et al., 2016; Abercombrie & Batista-Navarro, 2018) eta Holandako Parlamentukoari (Grijzenhout, Marx & Jijkoun, 2014) buruzkoak.

Bestalde, hizkuntza aniztasunaren eta aniztasun hori erakundeetan islatzeko moduaren azterketak ere piztu du ikertzaileen interesa. Europako Kontseilua hizkuntza aniztasunaren aldekoa da, eta, dioenez, Europa eleaniztuna da eta hizkuntza guztiak dira baliagarriak komunikaziorako (Europako Kontseilua, 2022).

Printzipio hori parlamentuko jarduerara eraman da, eta hizkuntza ofizial bat baino gehiago erabiltzen dituzten herrialdeetan edo ganberetan hizkuntzei eman zaien erabilerari buruzko azterketak egin dira. Askotariko egoerak aztertu dira, hala nola ganbera plurinazionalak (Europako Parlamentua, adibidez), hizkuntza ofizial bat baino gehiago dituzten herrialdeak eta hizkuntza gutxituen presentziari buruzko azterketa espezifikokoak. Halaber, parlamentu elebidunetako hizkuntzen erabilerari buruzko corpusak daude, hala nola frantsesa eta alemana erabiltzen diren Valaisko (Suitza) Kantonamenduko parlamentukoak (Imseng et al., 2012). Era berean, hizkuntza nagusiak eta gutxituak erabiltzen dituzten beste parlamentu batzuk ere aztertu dira, hala nola Eskoziako parlamentua (McLeod, 1998), Ipar Irlandakoa (Wilson & Stapleton, 2003), Hegoafrikakoa (Masumboka, 2017), Australiakoa eta Zeelanda Berrikoa (Murphy, 2020).

Parlamentuetan, hizkuntzen erabilerak, jakina, beste adar batzuk izango litzuke, tribunako erabileraz gain. Adibidez, Wright-ek (2007) Europako Parlamentuko eleaniztasunari buruz egindako azterlanean adierazten duenez, eleaniztasunarekiko konpromisoa parlamentarietako batzuek baino ez dute hartzen denboraren zati batean. Izan ere, berme formalak gorabehera, eleaniztasunaren erabilerak baldintzatu egiten ditu informaziorako eskuragarritasuna, batzordeetako negoziazio prozesuak eta parlamentuko jarduera informaleko beste alderdi batzuk. Hizkuntzek parlamentuetan duten erabileraren interesa balio sinboliko eta politikoan datza, ekintza politikoaren emaitza herritarrei jakinarazteko fasean.

Testuinguru horretan, artikulua honetan BasqueParl proiektuaren emaitzak aurkezten dira. Proiektu hori Eusko Legebiltzarreko bi legegintzaldiko (hots, 2012-2016 eta 2016-2020 legegintzaldietako) debateen corpusari buruz egindako ikerketa bat da, eta metodo konputazionala egindako aurrerapenak baliatzen ditu erretorika politikoa diziplina anitzeko ikuspegi batetik aztertzeko (Escribano et al., 2022). Eusko Legebiltzarren kasuan, debaterako foro publikoa den heinean, hitzaldiek Autonomia Erkidegoko bi hizkuntzen erabilera neurtzeko aukera ematen dute. Corpus horrez gain, parlamentuetako jarduera eleaniztunari buruzko beste batzuk daude (Roukos, Graff & Melamed, 1995; Koehn, 2005; Marx & Schuth, 2010; Hajlaoui et al., 2016).

Ikerketaren helburu nagusia da euskarak Eusko Legebiltzarrean duen erabilera ezagutzea, bilkuren akten azterketa automatizatuan oinarrituta. Helburu hori abiapuntutzat hartuta, ikerketarako honako galdera hauek ezartzen dira:

RQ1. Nolakoa izan da euskararen erabilerearen maila eta bilakaera Eusko Legebiltzarrean 2012-2020 aldian?

RQ2. Nolakoa izan da euskararen erabilerearen eta gaztelaniaren erabilerearen banaketa, generoaren eta alderdi politikoaren arabera?

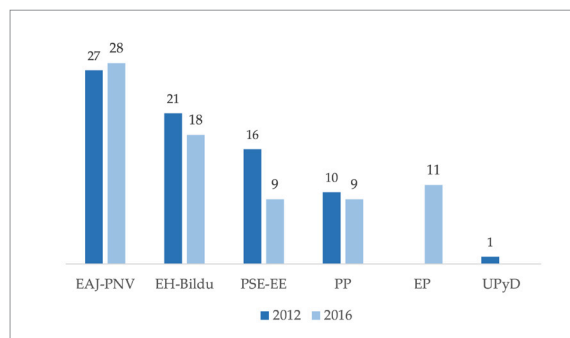
2. METODOLOGIA

Azterlan hau egiteko, Eusko Legebiltzarreko bilkuren transkripzioak oinarritzat hartuta, diskurtsoaren hainbat atributu aztertu dira, hala nola erabilitako hizkuntza (euskara edo gaztelania). Hizkuntza identifikatzeko (RQ1), eta corpusa elebiduna denez eta hizkuntza batetik bestera askatasunez aldatzen denez, diskurtsoko unitate bakoitza langdetect hizkuntza detektatzeko liburutegiarekin identifikatu da (Agerri et al., 2020).

Hizkuntzak paragrafoka esleitu dira, eta paragrafo bakoitzeko hitz guztiak hizkuntza bakarrari esleitu zaizkio, beste hizkuntzako hitz edo esaldiren bat bildu badezake ere. Adierazi beharra dago hitzaldi berean euskarazko zein gaztelaniarazko paragrafoak egon daitezkeela. Parte-hartze bakoitza hizkuntza bateko zein besteko gisa kontabilizatu da hizkuntza hori gutxienez paragrafo batean detektatu bada. Halaber, hitzaldi bera euskarazko eta gaztelaniarazko hitzalditzat har daiteke, eta, beraz, euskarazko eta gaztelaniarazko parte-hartzeen batura hitzaldien guztizko kopurua baino handiagoa da.

Legebiltzarkideak identifikatzeko eta haien hitzaldiak esleitzeko (RQ2), entitateen izenen ezagutzea (NER) eta lematizazioa ere egin da Flair-en bidez (Agerri & Rigau, 2020; Escribano et al., 2022), eta beste metadatu batzuk gaineratu dira, hala nola hitz egin duen pertsonaren izen osoa, jaio zen urtea, generoa eta ordezkatzeko duen alderdi politikoa. Horri guztiari esker, legebiltzarkidearen generoaren edo alderdi politikoaren arabera bereizi ahal izan da hizkuntzen erabilera.

1. irudia. Alderdi politiko bakoitzeko legebiltzarkideen kopurua X. eta XI. legegintzaldietan



Aztertutako aldian, hots, Eusko Legebiltzarraren X. eta XI. legegintzaldietan, lau alderdi politikok izan zuten ordezkari batzuetan: Eusko Alderdi Jeltzaleak (EAJ-PNV), Euskal Herria Bilduk (EH-Bildu), Euskadiko Alderdi Sozialista – Euskadiko Ezkerrak (PSE-EE) eta Alderdi Popularrak (PP). Bestalde, Unión, Progreso y Democracia (UPyD) alderdiak ordezkari bat lortu zuen 2012an, baina galdu egin zuen ondorengo hauteskundeetan, eta Elkarrekin-Podemos (EP) alderdia, berriz, 2016an sartu zen Legebiltzarrean, hamaika ordezkariarekin.

Halaber, Legebiltzarreko diskurtsoaren berezitasunak direla eta, Legebiltzarreko lehendakari Bakartxo Tejeriaren hitzaldiak eta gainerako legebiltzarkideenak bereizi egin dira corpusean, debateak zuzentzeko eta moderatzeko betetzen duen zeregina dela eta lehendakariak askotan hartzen baitu hitza bilkura bakoitzean. Zehazki, aztertutako aldian, aztertutako hitzaldien % 55 Legebiltzarreko lehendakariarenak izan dira (oso laburrak gehienak). Hitzaldi horiek aztertutako paragrafoen % 18 eta esandako hitzen % 5 izan dira.

Guztira, euskal legebiltzarkideek bi legegintzaldietan egindako hitzaldien transkripzioetatik ateratako 14 milioi hitzez osatutako corpus elebidun bat aztertu da. Erabilitako corpusaren demo bat dago eskuragarri honako helbide honetan: <http://legebiltzarra.ixa.eus/>

3. EMAITZAK

Legebiltzarkideek Eusko Legebiltzarraren X. eta XI. legegintzaldietan egindako hitzaldiei buruzko azterketaren emaitzek lehenik eta behin adierazten dutenez, bi hizkuntza ofizialen erabilera desberdina da debateetan. Hitzen identifikaziotik abiatuta parte-hartze gehienetan bi hizkuntzak erabiltzen badira ere, diskurtsoaren bolumena nabarmen desberdina da hizkuntza batean eta bestean.

Hizkuntza ofizial bakoitzaren erabileraren bolumena neurtzeko erreferentzia gisa hizkuntza bakoitzean esandako hitzen guztizko kopurua hartzen badugu (1. irudia), bi hizkuntzen artean hirurogei puntutik gorako aldea dagoela ikus dezakegu. Aztertutako aldian, legebiltzarkideek erabilitako hitzen % 81,6 gaztelaniakoak izan dira, eta % 18,4, berriz, euskarazkoak. Aitzitik, hitzaldiei erreparatzen badiegu, haien % 86,4k gaztelaniarazko paragrafo bat zuten gutxienez, eta % 65ek euskarazko paragrafo bat gutxienez. Beraz, legebiltzarkideek bi hizkuntzak erabili ohi dituzte hitzaldien zati batean gutxienez, baina hitzaldien zatirik handienaren edukia gaztelaniaraz garatzen da.

Legebiltzarreko lehendakariaren hitzaldiak salbuetsita daude aurreko datuetatik, debateko moderatzaile gisa betetzen duen zeregina dela eta askotan hartzen baitu hitza bilkura guztietan, hitza emateko edota ohartarazteko. Aztertutako aldian, Legebiltzarreko lehendakariak euskara

erabili du nagusiki. Legebiltzarreko lehendakariaren hitzaldien % 97,7k euskarazko paragrafo bat zuten gutxienez, eta gutxienez gaztelaniazko paragrafo bat zuten hitzaldiak % 24,1 izan dira. Hitzaldi laburrak izan direnez gero, haren parte-hartzea hizkuntza bakoitzaren erabilerearen kalkuluan sartzen badugu, Eusko Legebiltzarrean euskararen erabilera erabilitako hitzen guztizko bolumenaren % 21,2ra eta hitzaldi guztien % 83,5era iritsi da.

Datu horiek guztiak –Legebiltzarreko lehendakariaren hitzaldiak barnean hartuta– adierazten dutenez, euskararen erabilera handia da aldi laburretan edo protokolo edo testigantza izaerarekin, baina gaztelaniaren erabilera orokorragoa da hitzaldi luzeetan edo eduki handiagokoetan. Euskaraz erabilitako hitzen kopurua gutxienez euskarazko paragrafo bat duten hitzaldien kopuruaren herena baino txikixeagoa dela kontuan hartuta, esan daiteke euskararen erabilera askotan ez dela eraginkorra, sinbolikoa baizik.

1. taula. Corpusaren datuak

Hizkuntza	Hitzaldiak	Paragrafoak	Hitzak	Lemak	Erakundeak
Euskara	34.571	133.599	2.938.061	1.375.686	122.439
Gaztelania	18.016	209.067	10.934.044	3.714.887	227.451
Guztira	41.417	342.666	13.872.105	5.090.573	349.890

Iturria: geuk landua

Hitzaldiak erabilitako hizkuntzaren arabera banatuz gero ikus daitekeenez, legebiltzarkideen generoaren eta alderdi politikoaren aldagaiek ere desberdintasun esanguratsuak dituzte.

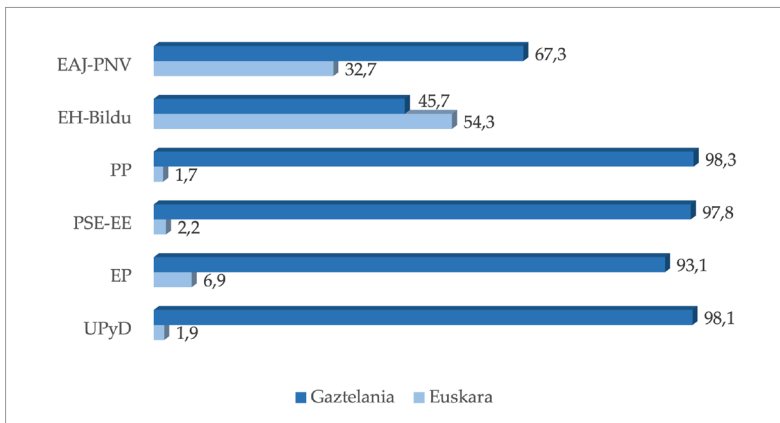
Alderdi politikoei dagokienez, eta hitzen guztizko kopurua kontuan hartuta, EH Bilduk baino ez du erabiltzen euskara gaztelania baino neurri handiagoan bere parte-hartzetan (% 54,3), EAJ/PNVren aurretik (% 32,7). Gainerako alderdiek neurri txikian erabiltzen dute euskara (2. irudia).

Aitzitik, neurketa unitate gisa bi hizkuntzak erabiltzen diren hitzaldiak erabiltzen baditugu, ikusiko dugu EAJko legebiltzarkideek aldien % 94,2tan erabili dutela euskara paragrafo batean gutxienez, eta % 86,5etan erabili dutela gaztelania (3. irudia). Bestalde, euskaraz soilik egindako hitzaldien kopururik handiena EH Bilduko ordezkariak egin dute. Haien hitzaldietan, euskararen erabilera % 95,5era iritsi da, eta gaztelaniarena, berriz, % 59,9ra. Elkarrekin-Podemos taldearen kasuan, alde handia dago euskararen erabilerearen guztizko bolumen txikiaren (% 6,9) eta diskurtsoaren zati batean gutxienez euskara erabili den hitzaldien ehuneko handiaren (% 82,6) artean. Horrek agerian uzten du euskarari agerikotasuna emateko ahalegina. Azkenik, Legebiltzarrean ordezkariak duten gainerako alderdi politikoek gaztelania erabiltzen dute ia kasu guztietan. PSE-EE alderdiak bere hitzaldien % 35,4tan

erabili du euskara, PP alderdiak % 21,7tan eta UPyD alderdiak, berriz, % 15,5etan.

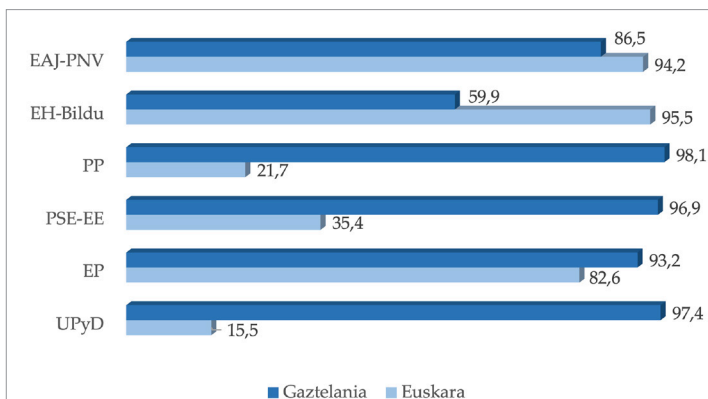
Generoa kontuan hartuta, emakumeek gizonak baino neurri handiagoan erabiltzen dute euskara (% 22,9 eta % 15, hurrenez hurren). Kasu horretan ere, Legebiltzarreko lehendakariaren hitzaldiak kontuan hartzen baditugu (haren zereginaren berezitasunekin), euskararen erabilera % 28,3ra iritsi da emakumeen kasuan.

2. irudia. Hizkuntzaren erabilera alderdi politikoaren arabera (hitz kopuruaren %)



Iturria: geuk landua

3. irudia. Hizkuntzaren erabilera alderdi politikoaren arabera (hitzaldien %)



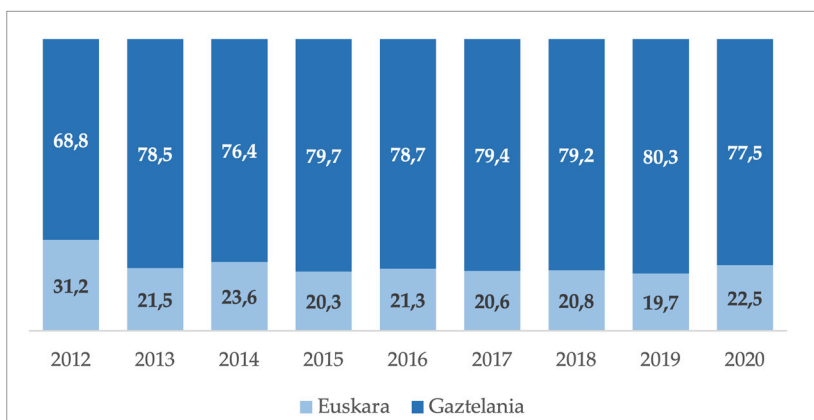
2. taula. Euskararen erabilerearen konparazioa

	Eusko Legebiltzarra	Kale neurketa	Hizkuntza gaitasuna
Kopurua	% 18,4 (% 21,2, legebiltzarreko lehendakaria barne)	EAE (% 15,47) Gipuzkoa (% 31,1) Bizkaia (% 8,8) Araba (% 4,6)	EAE (% 33,9) Gipuzkoa (% 50,6) Bizkaia (% 27,6) Araba (% 19,2)
Data	2012-2020	2016	2016
Iturria	BasqueParl	Soziolinguistika Klusterra	Inkesta Soziolinguistikoa

Euskararen erabilera aztertzeko BasqueParl corpusaren azterketaren (2012-2020) eta aldi berean egindako beste azterlan batzuen arteko konparazioa eginez gero (2. taula) ikus dezakegunez, euskararen erabilera eguneroko erabilerearen neurketek (Soziolinguistika Klusterra, 2016) emandakoa baino ia hiru puntu handiagoa da, baina herritarren hizkuntza ezagutza (Eusko Jaurlaritzak, 2016) baino hamabost puntu txikiagoa. Beste modu batera esanda, legebiltzarkideek herritarrek baino neurri handiagoan erabiltzen dute euskara, baina hizkuntzaren ohiko erabilera Legebiltzarrean ere gaztelaniarena baino askoz txikiagoa da.

BasqueParl corpusak, halaber, Eusko Legebiltzarrean euskararen erabilereak izan duen bilakaera historikoa aztertzeko aukera ematen du. Kasu honetan, datuek diote euskararen erabilerearen bolumenik handieneko urtea (2012) aztertutako serieko lehena izan dela (4. Irudia). Harrezkero, euskararen erabilereak egonkor jarraitu du, eta zifrak antzekoak izan dira urtez urte, alde esanguratsurik gabe. Izan ere, parte-hartzaileen (aulkia gutxienez lau urtetan bete duten legebiltzarkideen) lagina ere egonkorra izan da.

4. irudia. Hitzen erabilerearen urteko bilakaera (%)



Iturria: geuk landua

4. ONDORIOAK ETA EZTABAIDA

Legebiltzarreko debateen azterketaren interesa ez datza hizkuntza berezitasunetan, izaera politikoan baizik. Diskurtsoak Legebiltzarrean egiteak ematen die balio diferentziala, leku horretan legeak egin eta politikoki jarduten baita (Van Dijk, 2001). Alderdi bereizgarria, beraz, ez da hizkuntza izaera bera, diskurtsoak izan dezakeen eragin soziala eta diskurtsoa egiten dutenen ordezkari izaera baizik.

Erabilera politiko horrek ematen dio interesa Legebiltzarreko hitzaldien azterketari, hizkuntzaren erabilera elkarri lotutako testuinguruetan uler baitaiteke. Izan ere, aldi berean adostutako eta formalki aitortutako komunikazio mota bat, kultura edo herentzia indibiduala islatzen duen eskubide natural bat, eskubide formal eta legezko bat edota sinbolo politiko bat har daiteke kontuan (Wilson eta Stapleton, 2003).

Testuinguru horretan, BasqueParl proiektuan hizkuntza naturala prozesatzeko tresnen bidez Eusko Legebiltzarreko hitzaldiei buruz egindako azterketari esker, bi hizkuntza ofizialen erabilera bereizia ebaluatu ahal izan da, diskurtsoaren garrantzia izaera politikoan datzan testuinguru batean.

Lehenik eta behin, egindako azterketaren bidez, aztertutako aldi euskararen erabilera sinbolikoa edo protokolo erabilera erabilera praktikoa baino handiagoa izan dela identifikatu ahal izan da. Izan ere, euskara hitzaldi gehienetan erabiltzen bada ere, diskurtsoaren bolumen osoan erabilera nabarmen murrizten da. Era berean, hitzaldiaren iraupena zenbat eta luzeagoa izan, orduan eta maizago erabiltzen da gaztelania. Datu horiek bat datoz, halaber, aldi berean egindako beste azterketa batzuekin (Berria, 2013), eta, oro har, euskarari agerikotasuna emateko ahalegina egiten dela erakusten dute, baina beti ez zaio erabilera eraginkorra ematen debate tresna gisa edo argudioak edo ikuspuntuak aurkezteko tresna gisa.

Bigarrenik, Eusko Legebiltzarrean euskararen erabilera erabilera sozialaren gainetik badago ere, legebiltzarkideek ordezkatzan duten gizartearen isla dela esan daiteke. Zehazki, kaleko erabilera % 15,5era iritsi da (Soziolinguistika Klusterra, 2016), eta, data horretan bertan, Eusko Legebiltzarreko erabilera % 18,4koa izan da (% 21,2ekoa ganberako lehendakariaren hitzak sartzen baditugu). Generoa kontuan hartuta, legebiltzarkideen artean emakumeek euskara neurri handixeagoan erabiltzen dutela ikusi bada ere, aldeak ez dira esanguratsuak.

Erabilera hori (neurri batean sinbolikoa) bat dator Wrightek (2007) prozesu politikoan hizkuntzari ematen zitzaion erabileraren (hau da, parlamentarien lanaren) eta politikaren emaitzan hizkuntzari ematen zaion erabileraren artean identifikatzen zuen aldearekin. Prozesu politikoan, hizkuntzaren erabilera pertsonen hizkuntza gaitasunari eta hizkuntzen ohiko erabilerari lotuta dago, eta batzordeen lanean edo foro informaletan geratzen da agerian. Bestalde,

hizkuntzaren erabilera politikoa legebiltzarkideen hitzaldietan eta haien emaitzetan islatzen da, hala nola diskurtso publikoetan eta legeetan.

Hori dela eta, Legebiltzarrean hizkuntzari ematen zaizkion beste erabilera mota batzuk identifikatu ezin badira ere (aztertutako corpusa hitzaldi publikoetan oinarrituta dagoelako, legebiltzarkideek hizkuntzei ematen dieten beste erabilera batzuk alde batera utzita), Eusko Legebiltzarrean identifikatu den erabilerak erakusten du ahalegin handiagoa egiten ari dela euskarari agerikotasuna emateko, batez ere Legebiltzarreko talde batzuen kasuan.

Protokolo erabilera horren adibiderik onena Eusko Legebiltzarreko lehendakariaren hitzak dira, ia beti euskara erabiltzen baitu debateak moderatzeko. Alderdi politikoak kontuan hartuta, EH-Bildu da euskara gaztelania baino neurri handiagoan erabiltzen duen talde bakarra, alde txikiagatik bada ere. EAJ-PNV eta Elkarrekin-Podemos taldeek euskara erabiltzen dute ia hitzaldi guztietan, baina erabateko erabilera desberdin samarra da. PSE-EE, PP eta UpyD taldeetako ordezkarien hitzaldietan, berriz, gaztelania erabiltzen da nagusiki. Hori guztia bat dator euskarak talde politikoaren arabera duen erabilera sinbolikoarekin. Ildo horretan, erabilerak protokolo izaera handiagoa du batzuetan, eta izaera eraginkor handiagoa beste batzuetan.

Bestalde, BasqueParl corpusak legebiltzarkideen diskurtsoan aplikatutako hizkuntzaren azterketa automatizatuaren ahalmena adierazten du. Ildo horretan, hizkuntzen erabileran agertzen diren joerak ondorioztatzeko aukera ematen du, datu bolumen handien azterketan oinarrituta.

ERREFERENTZIAK

ABERCROMBIE, G. & BATISTA-NAVARRO, R. (2018). 'Aye' or 'No'. Speech-level Sentiment Analysis of Hansard Parliamentary Debate Transcripts. Accesible at: <https://aclanthology.org/L18-1659.pdf>

AGERRI, R. & RIGAU, G. (2020). "Projecting heterogeneous annotations for named entity recognition". In *Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2020)*. http://ceur-ws.org/Vol-2664/capitel_paper2.pdf

AGERRI, R., SAN VICENTE, I., CAMPOS, J. A., BARRENA, A., SARALEGI, X., ETXABE, SOROA, A. & AGIRRE, E. (2020). "Give your text representation models some love: the case for basque". In *Proceedings of LREC*. <https://doi.org/10.48550/arXiv.2004.00033>

ALCAIDE, E., CARRANZA, A., & FUENTES-RODRÍGUEZ, C. (2016). Emotional argumentation in Political Discourse. In Fuentes-Rodríguez, C. eta Álvarez, G. (koord.), *A Gender-based Approach to Parliamentary Discourse*. The Andalusian Parliament, 129-160.

- BARA, J., WEALE, A. & BICQUELET, A. (2007), *Analysing Parliamentary Debate with Computer Assistance*. *Swiss Political Science Review*, 13: 577-605. <https://doi.org/10.1002/j.1662-6370.2007.tb00090.x>
- BERRIA (2013ko abuztuak 6). “Euskararen erabilera Eusko Legebiltzarreko ekaineko osoko bilkuretan”. <https://www.berria.eus/paperekoa/1701/003/002/2013-08-06/euskararen-erabilera-eusko-legebiltzarreko-ekaineko-osoko-bilkuretan.htm>
- BOURDIEU, P (1991). *Language and Symbolic Power*. Cambridge: Harvard University Press.
- BURFOOT, C., BIRD, S., & BALDWIN, T. (2011). Collective classification of congressional floor-debate transcripts. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 1506–1515. Association for Computational Linguistics.
- DUTHIE, R., BUDZYNSKA, K., & REED, C. (2016). Mining ethos in political debate. In *COMMA*, pages 299–310.
- DUNMIRE, P. Political Discourse Analysis: Exploring the Language of Politics and the Politics of Language. *Language and Linguistics*, 6(11), pp. 735-751. <https://doi.org/10.1002/lnc3.365>
- ESCRIBANO, N., GONZÁLEZ, J.A., ORBEGOZO-TERRADILLOS, J., LARRONDO-URETA, A., PEÑA-FERNÁNDEZ, S., PEREZ-DE-VIÑASPRE, O. & AGERRI, R. (2022). “BasqueParl: A Bilingual Corpus of Basque Parliamentary Transcriptions”. *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, Marseille, 20-25 June 2022, pp. 3382–3390. <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.361.pdf>
- EUROPAKO KONTSEILUA (2022). *Education and Languages, Language Policy*. Estrasburgo: Europako Kontseilua. http://www.coe.int/t/dg4/linguistic/division_EN.asp
- EUSKO JAURLARITZA (2016). *VI. Inkesta soziolinguistikoa*. Gasteiz: Eusko Jauraritza.
- GRIFFITH, J. & LESTON-BANDEIRA, C. (2012). How Are Parliaments Using New Media to Engage with Citizens?, *The Journal of Legislative Studies*, 18:3-4, 496-513. <https://doi.org/10.1080/13572334.2012.706058>
- GRIJZENHOUT, S.; MARX, M.; JIJKOUN, V. (2014). “Sentiment Analysis in Parliamentary Proceedings”. In *From Text to Political Positions: Text analysis across disciplines*. Edited by Bertie Kaal, Isa Maks and Annemarie van Elfrinkhof [Discourse Approaches to Politics, Society and Culture 55], pp. 117–134
- GUTMANN, A. & THOMPSON, D. F. (1996). *Democracy and Disagreement*. Cambridge: Harvard University Press.
- HAJLAOUI, N., KOLOVRATNIK, D., VÄYRYNEN, J., STEINBERGER, R., & VARGA, D. (2014). DCEP -digital corpus of the European parliament. In *Proceedings of the*

Ninth International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, Iceland, May. European Language Resources Association (ELRA).

- IMSENG, D., BOURLARD, H., CAESAR, H., GARNER, PHILIP N., LECORVÉ, A., NANCHEN, A. (2012). Mediaparl. Bilingual Mixed Language Accented Speech Database. IEEE Spoken Language Technology Workshop (SLT), 2012, pp. 263-268, doi: 10.1109/SLT.2012.6424233.
- IÑIGO-MORA, I. (2007). Estrategias del discurso parlamentario. La Cámara de los Comunes y El Congreso de los Diputados. *Discurso & Sociedad*, 1(3): 400-438.
- KOEHN, P. (2005). Europarl: A parallel corpus for statistical machine translation. In *Proceedings of Machine Translation Summit X: Papers*, pages 79–86, Phuket, Thailand, September 13-15.
- KRUSE, L. M., NORRIS, D. R. & FLINCHUM, J. R. (2018). Social Media as a Public Sphere? Politics on Social Media, *The Sociological Quarterly*, 59:1, 62-84. <https://doi.org/10.1080/00380253.2017.1383143>
- MASUMBOKA, E.L. (2017). *Critique of Multilingualism in South Africa's Post-Democratic Parliament with particular reference to the use of Selected Minority Languages*. Grahamstown: Rhodes University.
- MARX, M. & SCHUTH, A. (2010). DutchParl. The parliamentary documents in Dutch. In *Proceedings of the Seventh International Conference on Lan-Resources and Evaluation (LREC'10)*, Valletta, Malta, May. European Language Resources Association (ELRA).
- MCLEOD, W. (1998). "Scotland's Languages in Scotland's Parliament". *Scottish Affairs*, 24(1), 68-82.
- MURPHY, J. R. (2020). "Indigenous Languages in Parliament and Legislation: Comparing the Māori and Indigenous Australian Experience". *Māori Law Review*, <http://dx.doi.org/10.2139/ssrn.3712771>.
- ESCRIBANO, N., GONZÁLEZ, J.A., ORBEGOZO-TERRADILLOS, J., LARRONDO-URETA, A., PEÑA-FERNÁNDEZ, S., PEREZ-DE-VIÑASPRE, O. & AGERRI R. (2022). BasqueParl: A Bilingual Corpus of Basque Parliamentary Transcriptions. *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, pages 3382–339. <http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.361.pdf>
- ONYIMADU, O., NAKATA, K., WILSON, T., MACKEN, D., & LIU, K. (2013). Towards sentiment analysis on parliamentary debates in Hansard. In *Joint International Semantic Technology Conference*, pages 48–50. Springer
- RHEAULT L., BEELEN K., COCHRANE C. & HIRST G. (2016) Measuring Emotion in Parliamentary Debates with Automated Textual Analysis. *PLoS ONE* 11(12): e0168843. doi:10.1371/journal.pone.0168843.
- RICHARDS, B.(2010). The Emotional Deficit in Political Communication. *Political Communication*, 21(3), 339-352.

- Escribano, N. [et al.]: Euskararen erabilera Eusko Legebiltzarreko debateetan (2012-2020)
- ROUKOS, S., GRAFF, D., & MELAMED, D. (1995). Hansard French/English LDC95T20.
- SALAH, Z. I. (2014). Machine Learning and Sentiment Analysis Approaches for the Analysis of Parliamentary Debates [PhD. Thesis]. University of Liverpool.
- SOKOLOVSKA, Z. 2016. "Imagining Europe's linguistic diversity in the Council of Europe's Parliamentary Assembly". *Language & Communication*, 51, 40-49. <https://doi.org/10.1016/j.langcom.2016.07.004>
- SOZIOLINGUISTIKA KLUSTERRA (2016). *Hizkuntzen erabileraren kale neurketa*. Andoain: Soziolinguistika Klusterra.
- THOMAS, M., PANG, B. & LEE, L. (2006). Get out the vote: Determining support or opposition from congressional floor-debate transcripts. In *Proceedings of the 2006 conference on empirical methods in natural language processing*, pages 327–335. Association for Computational Linguistics.
- VAN DIJK, T. A. (2001). Texto y contexto de los debates parlamentarios, *Tonos Digital: Revista Electrónica de Estudios Filológicos*, 2, 1-47.
- WILSON, J., & STAPLETON, K. (2003). Nation-State, devolution and the parliamentary discourse of minority languages. In *Journal of Language and Politics* (2. bol., 1 zenbakia, 5-30. orr.). John Benjamins. <https://doi.org/10.1075/jlp.2.1.03wil>
- WRIGHT, S. (2007). "English in the European Parliament: MEPs and their Language Repertoires". *Sociolinguistica*, 21(1), 151-165. <https://doi.org/10.1515/9783484605060.151>