

100. zenb.

XL. URTEA
2022 UZTAILA-ABENDUA

ISSN 2530-3287

**Mikel
Osinalde
Agirre**

Eibarko Hizkuntza Eskola Ofiziala
Euskara Mintegia

mosinalde@irakasle.eus

**Mikel
Iruskieta
Quintan**

HiTZ zentrua, IXA Taldea
Euskal Herriko Unibertsitatea
Hizkuntzaren eta Literaturaren
Didaktika Saila. Euskara Institutua

mikel.iruskieta@ehu.eus

Ikerketak

Hizkuntza-ikasleen testu-corpus etiketatuaren analisia eta interpretazioa B2 eta C1 mailetan

Gero eta garrantzitsuagoa da ikerketan sorturiko datuak erakustea, are ikerketaren fida-garritasuna edota gai berean interesa duten ikerlariekin partekatzea ere, itxi gabeko ikerketa-ildoak garatzeko. Hori horrela, Humanitate Digitalen esparrutik eta ikerketa-azpiegiturak erabiliz, hizkuntzen irakaskuntzan datuetan oinarrituriko ebaluazioan urratsak emateko asmoz egin dugu lan hau. Lan honetan zinezko gizabanakoek benetako azterketatan ekoiztutako testuak aztertu dira, B2 eta C1 mailakoak. Lan hau HABEren eta Ixa Taldearen (UPV/EHU) elkarlanari esker burutu izan da. Egindako azterketa-lanaz gain, zenbait tresna eta bitartekoren berri ematen da eta etorkizunean horiek hizkuntza-gaitasunen ebaluazioan zein azterketan zenbaterainoko laguntza ekar lezaketen hausnartu da. Testu-zuzenketaren etiketatze-mailan Markin tresnarekin erabiltzeko etiketa-zerrenda ere proposatzen da eta hori baliatuta testuetako erroreak B2 eta C1 mailetan deskribatzeko saiakera ere egiten da, hein batean, halako ekimenek aurrera begira ekar ditzaketen onuren isla eskainita. Egindako irizpenak zein baliatutako datu gehienak eta bitartekoak arloko komunitatearen esku uzteko erabakiak lanaren zein jarduera esparruaren helburuekin bat egiten du, EuDat datuen azpiegitura europarra erabili dugu xede horretarako.

Gako-hitzak

analisia, corpusak, ebaluazioa, etiketatzea, hizkuntzen irakaskuntza, humanitate digitalak

1. SARRERA

Garapen teknologikoen eta berarekin ematen diren rol aldaketek eguneroko bizimoduaren eragina erabatekoa da (Kirschenbaum, 2010). Bai behinik behin, orain hamarkada bat edo bi genuen bizimoduarekin eta ikasteko moduarekin alderatuz gero. Egunero egiten ditugun ekintza gehienetan baliatzen ditugu gailu digitalak. Sakelakorik gabe bizitzeak barne herstura eta ezinegona eragin diezaguke. Lagunekin, familiako kideekin edo lankideekin hitz egiteko, mezularitza edota posta elektronikoen bidez komunikatzeko nahiz hainbeste hedatu diren sare sozialen bidez harremanetan eta informazio-bilketan aritzeko ezinbesteko bilakatu dira zenbait kasutan aipatutako teknologiak.

Hezkuntzan ere igarri da bilakaera teknologiko horren eragina eta horrek ekarri dituen moldaketa pedagogikoak eta diziplina berriak (Isasi eta Iruskietia, 2019). Esaterako, ikasleei dagokienez, ikasi behar duten edukiaren eta horretarako hartu beharreko denboraren kontrol eta erantzukizun handiagoa bereganatzeko testuinguruen egokitasuna handitu egin da.

Irakaslea informazio-iturri izatetik informazioa bideratu, ikaskuntza-ekosistemak baliatu eta prozesuan bidaide izatera aldatu da hainbat esparrutan (ohiko prozedurek hor jarraitzen badute ere, egoera berri horretara bidean gaudela esango genuke). Haatik, maiz betiko irakaskuntza-ohiturak errepikatzen ditugu bestelako baliabide teknologikoak erabiliz. Horregatik, Fyfek (2011) proposatu bezala, ikuspegi teknologikotik edo teknopedagogiatik hobe da ikastunitatea pantailak itzalita abiaraztea, zer, noiz, non eta nola erabili teknologia jakiteko.

Aldaketa horiek guztiek, gizartearen jokabidea, xedeak eta izaera ere bestelakotzea ekarri dute. Areago, esan genezake, aldaketok eta berrikuntzok funtsezkoak dira edo izan beharko lukete teknologia modu esanguratsuan erabiltzeko (Arriola et al., 2020); izan ere, teknologiak ere badakartza arazo ekonomikoak, autonomia erasaten du eta osasunean ere kalteak ekar ditzake. Baumanek (2003) zehaztu duen gizarte likidoarekin lotzen dugu aspalditik giza azturatan antzematen den berriarekiko grina neurrigabea, oinarriari eutsi ezina eta presa mugagabea, zeina eskolan ere sumatzen den. Alegia, arretari eutsi ezina, azalpen oro luzeegizat hartzea: “aspertu naiz”, “eskain iezadazu beste zerbait” edo “eskolan lana egiteko (guztia egiteko) denbora izango dugu?”. Hala ere, gaude jendarte horren araberrako irakas-ikaskuntza prozesuak antolatu beharrak eragozpenak sorrarazten dizkigula. Ohitura horien ondorioen berri badugu, nola ekonomian hala ingurumenean; ez ote genuke mundu hori aldatzeko ahalegina egin behar patxada eta sosegua gure jardun eta metodoetan txertatuta?

Sinetsita gaude datuen bilketa masiboak, prozeduren eta ohituren azterketa sistemati-koak eta guztien interpretazioak zoru trinkoagoa ezar dezaketela hezkuntzako eragile-lean jardunean. Zentzu horretan hartzen ditugu beharrezkotzat Humanitate Digitalak, ez nahitaez noraezean eta alderrai dabilen horri egokitzeko.

Hainbat aldiz eraiki, berregin eta egokitutakoari eusten diharduenak entzungor egiten dio bestelako jokabideak saiatzeari. Izan ere, aldaketak hemen badaude eta horietara moldatzen jarraitu behar badugu, eraikitzen hainbeste kosta zaion habiaren egonkortasuna zalantzan legoke. Olatu berritzaileak harrapatu aurretik ihes egin nahi luketen bali handiko irakasle asko ezagutu ditugu urteotan. Uhinaren azkartasunak egoera larritu egiten du. Batel formako gurdiak egiten zituen ontziolatik bidalitako arotzen irudia adierazgarria begitantzen zaigu (Pako Aristiren “Industri birmoldaketa” olerkiko irudia hartuta).

Bestalde, konputazio-gaitasunen areagotzeak hizkuntza-irakasleei eskaini dien onura zerrendari ezin zaio ezikusiarrena egin. Berrikuntzak onerako erabiltzeko kemena eta grina erakusten duten hezkuntzako kide ugarien adibidea itxaropen pizgarria da. Uretan dauden ikasleekin igeri egin eta haien beharretatik zein bitarteko ezagunetatik ekiten duenaren eredu. Ea etorkizuna halabehar tristearen atzamarretan ez uzteko gai garen eta etorkizun gizatiarragoa adosten asmatzen dugun geuk sortutako teknologia geure alde jarrita.

Zeinahi jarduera esparrutan sortzen den ekoizpena bilduko, sailkatuko eta aztertuko bagenu, egindako ibilbidea ulertzeko eta etorkizuna bideratzeko aukera areagotuko genuke. Horixe ikusten diogu corpusgintzari onura nagusizat. Egindakoa ulertzeko eta desbideraketak aztertu ostean hobekuntzak proposatzeko ezinbesteko urratsa delakoan gaude.

Hezkuntzaren alorrean ere hori ekarri du egin izan denetan. Informazioa biltzeak hura ezaugarritzen laguntzen du; esaterako, ohiko akatsen sailkapen edo deskribapen proposak egin daitezke. Jakin badakigu hainbat saiakera egin direla arlo honetan, nola eremu pribatua hala publikoan. Ahalegin horiek tresna komertzialak, zenbait enpresen oinarria eta ikertzaileentzako lanabes gisa erabilia ezagutza areagotzea ekarri dituzte.

Alabaina, datu-bilketa hori komunitate osoak partekatutako irizpideen arabera eta eragile ororen eskura utziz egingo balitz, ez al leharke iraultza (geure ustetan) onuragarria hezkuntzaren esparrura, izan ikaste- zein irakaste-jardunbideetan?

Unean uneko ikastaldeko ekoizpenetik (ahozkoa nahiz idatzia) abia genezake ikas-irakaskuntza prozesua. Ebaluazio jarraituari benetako izaera eskaini. Banakoen zein taldeen ibilbidean zinez lagungarriak liratekeen aholkuak eta proposamenak datu enpirikoetan oinarritu. Erroren sailkapenak ikasleen mailaren arabera egin litezke, jatorrizko hizkuntzei so, ikasten ari diren hizkuntzei begira, osatutako ariketa aintzat hartuta edota, besteak beste, ohiko akatsen maiztasun zerrendak sailkatuta.

Irakasleek eta ikasleek gain, ikerlariak ere izango lukete zeresanik eta egitekorik honetan guztian. Testuen ebaluazio automatikoan, akatsen tratamendu automatikoan zein, besteak beste, tresnen eta emaitzen interpretazioan.

Horretarako, jakina, adostasun handia behar litzateke hezkuntza-komunitateko eragile guztien artean eta lanerako ohiturak zein irizpideak zehazteko ahalegina. Datuen iturria agorgaitza da, kodetze-esperientziak badira eta corpusen egokitasuna neurtzen ere bada eskarmenturik.

Kultura eta ezagutza gordetzeko, iraunarazteko zein garatzeko ezinbestekoa den Humanitate Digitalen azpiegituretan corpusok denen eta, batik bat, behar dituztenen eskura izan behar lukete dagokion arloko aplikazioen eta tresnen oinarri direlako zein esperimentazio- eta proba-eremu aparta direlako.

Humanitate Digitalerako ikerketa-azpiegiturek erabiliz, hizkuntzen irakaskuntzan datuetan oinarrituriko ebaluazioan urratsak emateko asmoz egin dugu lan hau. Lan honetan zinezko ikasleek benetako azterketatan ekoiztutako testuak aztertu dira, HABE eta Ixa Taldearen (UPV/EHU) elkarlanari esker (Azpillaga, 2022). Egindako lanaz gain, Humanitate Digitalerako azpiegitura europarrak erabili ditugu eta erabilitako baliabide, tresna eta bitartekoen berri ematen da eta etorkizunean horiek hizkuntza-gaitasunen ebaluazioan zein azterketan zenbaterainoko laguntza ekar lezaketen hausnartu da. Testu-zuzenketaren etiketatze-mailan bateratze-lan bat ere proposatzen da, eta hori baliatuta testuak EEMB eta HEOCeke B2 eta C1 mailako ezaugarrien arabera deskribatu da. Egindako irizpenak zein baliatutako datu gehienak eta bitartekoak arloko komunitatearen esku utzi dira Humanitate Digitalerako azpiegiturek baliatuta.

Emaitzei dagokionez, irakasle batek zuzen ditzakeen testu-kopurua zuzendu da eta B2 eta C1 mailen arteko alderaketan nabarmenenak izan diren erroreak (ortografia, aditz laguntzailea, moldea, ergatiboa, singular/plural komunztadura, mugagabea eta aditz aspektuko markak) deskribatu dira. Horrez gain, etorkizunean interesgarria litzateke testu gehiago zuzentzea eta hemengo emaitzak esanguratsuak diren berrestea. Horrez gain, ildo honetan egindako lanak erraztea, beste irakasle batzuk ere antzeko emaitzak lortzea eta Humanitate Digitalerako azpiegiturek erabilgarriak diren aztertzea komenigarria dateke.

2. METODOA: AZTERKETETAKO OHARRAK AZTERGAI

Deskribatutako testuingurua gehiago zedarritze aldera, euskararen irakaskuntzari so aritu gara azpiatal honetan. B2 eta C1 mailetako azterketak egin dituzten ikasleen idatzia jaso ditugu HABEri eta Ixa Taldeari esker eta datu horien analisisan osatu da honako azterketa.

Helburuen artean daude mailetako idatzien tipologia jakin bati antzematea, beharrezko ohar-irizpideak bateratzea edota etiketa-multzoak biltzeko irizpideak adostea. Finean, analisi hauek irakasleen zein ikasleen eguneroko xedeetan lagungarri izan ote daitezkeen egiaztatu nahi dugu, hein batean bada ere.

Horretarako bildu ditugu, beraz, HABEko azterketetako zenbait idazlan aipatutako mailetan sailkatuta, B2 eta C1 mailetan. Modurik aproposenean antolatu ditugu, haientzako Markin tresnarekin erabiltzeko etiketa egokiak proposatu, lanak eskuz ebaluatu eta etiketatu, emaitzak jaso, antolatu eta aztertu.

Aipatu idatziek dagoeneko jaso dute dagokien irakaslearen edo ebaluatzailearen irizpena (HABEk egina) eta baita zenbait tresna automatikorena ere (Ixa Taldeak egina). Hori horrela, haien gaineko aberaste-lana ere abiatu dugu bata eta bestearen arteko alderatzeak fintzeko asmoz.

2.1. EGITASMOAREN GORDAILUA

Pare bat ideia hartu ditugu aintzat proiektuaren alderdi honi heltzerakoan. Batetik, ahal bezainbeste bitarteko eta datu etiketatzeak zein eskuragarri bezain erabilgarri izateak berebiziko garrantzia duela. Gordailuen xedea aurrez egin den hori guztia etorkizunetarako erabilgarri uztea izan behar luke. Egiteko horretan Humanitate Digitalean diharduten eragileek zein banakoek zeresan handia izango dute irizpideak ezartzen, bilaketa tresnak prestatzen edota bildutakoaren interpretazio aproposak egiten.

Bigarrenik, dena gordetzerik eta erakusterik ez dagoen arren, aukeraketa egiten den bitartean galtzerik nahi ez genituzkeen oroitzen komunitarioak haraindikotasuna iritsi dezaten lortuko da. Halaxe eskainiko zaio etorkizuneko begiraleari iraganaren ikuspegi ahalik eta zintzoena.

Horiek gogoan, Europa mailako ondarea gordetzeko asmoz eratu den EUDAT¹ (Lecarpentier et al., 2013) ekimena baliatzea erabaki da gure ahaleginean bildutako datuak zein irizpenak hezkuntza- zein ikerketa-komunitatearen esku uzte aldera. Elkarlanerako datu-azpiegitura horren helburuekin bat egiten dugu elkarrekintzaren alde egiten duelako eta Europa mailan diharduten hainbat diziplinako ikertzaileen beharrei zein gizartearen onari begira antolatua dagoen testuinguru publikoa baita.

Lanean zehar zein eranskinetan ezarri ditugun esteka guztiak gordailu horretan jaso ditugun geure dokumentuetara eramango dute irakurlea.

2.2. DATUEN BILKETA

HABEren eskutik jaso ditugu zenbait azterketa modu anonimoan, erdiak proba gainditu zutenenak eta beste erdiak gainditu ez zutenenak. Eskuz idatzitako testuak izaki, analisirako digitalizatu egin behar izan dira, OCR pasatuz, *Transkribus* tresna erabiliz. Ondoren, XML egiturako fitxategietan aztertzailearen datu pertsonalak eta idazlanaren azken kalifikazioa kodetu da (baita egokitasunari, koherentziari, kohesioari, aberastasunari eta zuzentasunari dagozkien kalifikazioak ere). Lan hori HABEri egin izan da. Ondoren, dokumentu horien kalifikazioak datu-base batean gorde eta testuak TXT formatura pasatu dira, Ixa Taldeko analizatzaileekin modu automatikoan aztertzeko eta sailkatzaile

¹ <https://www.eudat.eu/>

automatikoaren erantzunak lortzeko. Horra, Humanitate Digitalen beste alderdi onuragarrietako bi: elkarlana eta diziplinartekotasuna.

Nahiz eta corpus-lagin handiagoa lortzeko aukera izan, erabaki dugu irakasle batek epe jakin batean zuzen ditzakeen lanak aintzat hartzea. Horrela, corpusaren zati bat hartu eta hogeitaz idazlaneko bi multzo baliatu ditugu gure xederako; orotara, beraz, 40 testu. B2 eta C1 mailetak azterketa-sortak dira, EEMBaren eta HEOCaren arabera, hori baita hizkuntzaren erabiltzaile independenteen eta erabiltzaile gaituaren artean dagoen urratsa. Lanak banan-ban aztertu ditugu *Markin* tresnarekin eta ondorioetan multzokatzek eta datuen interpretazioa eskaini dira.

Bildutako dokumentuekin benetako datuz osatutako corpus txiki bezain baliagarria osatu da. Ataza jakin bati lotutako lanak dira, hots HABEren azterketetan eskatutako egitekoetara bideratuak. Euskara-ikasleen ekoizpenak direnez, hizkuntza bakarreko bilduma da baliatu duguna.

2.3. ETIKETATZEA

Badira testuak automatikoki aztertzen dituzten tresna eraginkorrak, baita hizkuntza-osagai bakoitzaren deskribapen zabala egiteko gai direnak ere: *CTAP* (Chen & Meurers, 2016), *Multiaztertest* (Gonzalez-Dios & Bengoetxea, 2021) eta *ANALHITZA* (Otegi et al., 2017), besteak beste. Haatik, hizkuntzaren deskribapena ez da lan honen xedea; aitzitik, hizkuntza horren erabilera okerrak edo nolabaiteko desbideraketak azaleratuz horien deskribapena egiten ahalegintzea du helburutzat lan honek.

Gauzak horrela, corpora eskuz etiketatzea erabaki genuen. Era horretara, akatsen deskribapena edota haien nolakotasunaren inguruko hausnarketa egiteaz gain, eskuarki agertzen diren hutsegiteen aurreikuspena neur genezakeen.

B2 eta C1 mailako azterketak ebaluatzeko baliatu izan diren marka-zerrenda eskuratzeko modua izan dugu eta egokitzapen gutxi batzuekin hura hartu dugu datu-bilketarako eredutzat. Horien iturria ikusi nahi bada jo Joseba Butroek sorturiko atalak.eus webgunera,² eta bertan C1 mailako Euskara Gaitasun Agirian erabiltzen ziren markak deskribatzen dira. Guk dakigula erabaki dugun mailarako ez dago bestelako marka-zerrendaren proposamenik.

Eskuz etiketatzeko, ordea, aro digitaleko tresna baten ekarpena baliatzea erabaki dugu. Alegia, hein batean bada ere, ebaluazio-prozesua automatikoxeagoa egiten saiatu gara errubrika baliatuz eta, batik bat, irizpide bateratuak ezartzeko lagungarria dena. Egiteko horretan *Markin* tresna hautatu dugu.³

2.3.1. Markin tresna

Online irakaskuntzan ebaluazio-lanetan lagungarri izan asmoz apailatutako lanabes honen, *Markinek*, baditu guretzat erabilgarriak diren hainbat ezaugarri.

Digitalizatutako dokumentuak banaka aztertzea eta etiketatzea ahalbidetzen du. Etiketak egokieraren arabera moldatzeko, aldatzeko zein emendatzeko modua eskaintzen du. Ikasleei haien ekoizpenak hobetzeko oharrak jasotzea errazten die HTML formatuan helaraz dakiekeen dokumentuen bidez (ikus 1. irudia). Irakasleei eta ikerlariei, berriz, corpus osoaren datuak biltzeko eta analisira bideratzeko modua ere eskaintzen die.

² <https://atalak.eus/ikasgela-materialak/idazlanak-kalifikatzea>

³ <https://www.cict.co.uk/markin>

1. irudia: Markin tresnaren feedback dokumentua HTML formatuan ikaslearentzat.

Normal view | Statistics | Feedback | All annotations | Marking key

BIZITZA DIGITALIZATUA

Azken urte hauetan sare sozialen eta komunikabide digitalen erabilera asko areagotu da. Noiztik ez zara etxetik irretzen mugikorra hartu gabe? Noiztik ez zara bankura joan diru trasferentzia bat egitera? Hau da, digitalizazioaren biktima al gara? **HUTSA**

Gure bizitzako arlo askotan erabiltzen ditugu tresna digitalak, **medikuan hitzordua hartzeko Postposizioa** **Dekl**, egunean zenbat pauso egin ditugun jakiteko, **erropa Ort** erosteko... Esan behar da, arlo batzuetan bizitza erraztu digutela. Adibidez, pandemia honetan etxean geunden, **baifa Ort**, ordenagailuei esker eta bideo-deien bidez gure ikasketak aurrera ateratzeko formakuntza jasotzea lortu dugu, eta hau dena tresna digitalei esker izan da.

Statistics

Instances	Annotation	Explanation	Help link	Categories	Value	Points lost	Points gained
3	Aditz laguntzailea	Aditz laguntzailea		Grammar	-2	-6	
1	Ergatiboa	Ergatiboa akatsa		Grammar	-2	-2	
1	HUTSA	HUTSA -- Zer bait falta da		Sense	-2	-2	
3	Koherentzia/kohesioa	Koherentzia eta kohesioan arazoak		Sense	-2	-6	
1	Dekl	Deklinabidea		Grammar	-1	-1	
5	Lexikoa	Hitzegian erderakada edota aukeraketa eskasa		Content Vocabulary	-1	-5	
3	Mugagabea	Mugagabea/mugatua		Grammar	-1	-3	
3	Moldea	Molde desegokia: pertsona, tonua, tratamendua		Grammar Style	-1	-3	
13	Ort	Ortografia		Editing	-1	-13	
2	Sing/Plu	Singular/plural errorea		Grammar	-1	-2	
2	Postposizioa	Postposizioan akatsa		Grammar	-1	-2	
1	Galdegaia	Galdegaia gaizki		Syntax	-1	-1	
Totals						-46	0

Ohar- edo etiketate-prozesu hori askotarikoa da edo izan daiteke (ikus 2. irudia). Akatsak identifikatzen eta markatzen dira, haien gaineko oharrak ere egin daitezke eta ohi-koenak diren ohar horiek automatizatzen bidea ere ematen du *Snippets* izeneko bitartekoa erabilita.

2. irudia: Markin tresnaren etiketatzeko aplikazioa: ikasle-testua zuzentzen.

The screenshot shows the e-Hizpide application window. The main text area displays the document 'BIZITZA DIGITALIZATUA' with several annotations in red and green. The sidebar on the left contains a grid of icons for different annotation types and tools. The bottom of the window shows a search or filter input field.

Aipa dezagun, adibide gisa garatu dugun egitasmo honetan ari ginela, euskarri zuen enpresak uko egin behar izan diola tresna garatzeko lanari zein erabiltzaileei babes eskaizteari. Horrek aurrera begira halako tresnen eskuragarritasuna bermatzeko ahaleginak bultzatu beharra azalarazi duelakoan gaude.

2.3.2. Hasierako prestaketa

Idazlan guztiak kalkulu-orri batean jaso dira. Alde batetik, idazlana dago bere horretan digitalizatuta eta, bestalde, automatikoki esleitu zaion emaitza edo balioa jarri zaio ondoz ondo.

HABEk darabiltzan zuzenketa-irizpideak hartzen ditu aintzat aipatutako ebaluazio automatiko horrek. Bost irizpideren arabera egiten den ebaluazioan bakoitzak eragin jakina hartzen du. Esaterako: 1. Egokitasuna: % 10 || 2. Koherentzia: % 25 || 3. Kohesioa: % 15 || 4. Aberastasuna: % 20 || 5. Zuzentasuna: % 30 || GUZTIRA: % 100.

Horrela, irizpide bakoitzaren barnean, puntuazio bakoitzak (A, B, C, D, E) ere balio desberdina dauka, bai irizpidearen beraren arabera, bai mailaren arabera. Adibidez, egokitasuneko A batek 3 puntu ematen baditu, koherentziakoak 7,5 emango lizkioke C1 mailan. Ikasleak orotara 15 puntu lortu behar ditu azterketa gainditzeko, 30 puntutik.

Balio beza azalpenak jasotako dokumentuaren nolakotasuna deskribatzeko; hala ere, ez da lan honen oinarritzko analisian halako ebaluazio-markarik ezarri. Lagungarri baino ez ditugu aintzat hartu eta, kasuan kasu, etiketatze automatikoan izan zitezkeen desberaketen berri emateko.

Gorago aipatu den gisan, EGAko azterketak ebaluatzeko erabili izan diren etiketa-sortak *Markinera* moldatu ditugu. 47 etiketako sorta nahikotzat jo dugu egin beharreko azterketarako. Horietako batzuk ez dira ebaluazio-prozesuan erabili eta beste zenbait gutxitan. Kontuak kontu, bestelako idazlanetan lagungarri izan daitezkeelakoan ez ezabatzea erabaki da.

2.3.3. Prozedura

Markin tresnaren erabilerarako egileotako batek sorturiko dokumentu honi jarraitu diogu.⁴ Dokumentu hori eskuragarri egiteko Humanitate Digitalak sustatzeko EUDAT datuen azpiegitura europarra erabili dugu, arestian aipatu bezala. Etiketa edo botoiak inportatu eta moldatu ostean, idazlanak *Markin* aplikazioko atal idazgarrira kopiatu eta ebaluatzeari ekin ahal izan diogu.

Beste ezaugarri interesgarri bat *snippet* delakoena da. Etiketatzeaz gain ohiko oharrak ere balia baitaitezke, maiztasun handiz badarabiltzagu. Norberak sor ditzake erabili ahala, baina horiek ere besterengandik eskura daitezke eta aplikazioan txertatu. Horien instalazio-prozedura goiko dokumentuan aurki daiteke.

Testu bakoitza etiketatu ostean ere ebaluazio-oharrak eransteko aukera dago *Markin* aplikazioan eta analisirako baliagarria izan ez arren, benetako egoeran irakasle orok erabiliko eta ikasleek eskertuko luketen moduan ekin diogu ariketa bakoitzaren irizpena osatzeari (ikus hemen adibide bat).⁵

Lan honetan egin ditugunak honako hauek dira:

- Etiketatzea. Testu bakoitza *Markin* tresnaz etiketatu da. Ezer azaltzen hasi aurretik, komeni da emaitzok irakasle bakarraren etiketatze-saioaren ostean aterakoak direla azpimarratzea, hots, saiakera osoago baten atalasean geratzen den lagin gisako ariketa dela.

⁴ <https://b2drop.eudat.eu/s/qFdZt4fet4SmSwF>

⁵ <https://b2drop.eudat.eu/s/7FYzj3KfxFg92tG>

- Corpuseko estatistikak. Idazlan guztietan oharrak ahal bezain egoki jaso ostean, estatistikak edo corpusaren datu interesgarrienak biltzeari ekin diogu. Markinek berak dokumentu bakarrean biltzen ditu nahi bezainbeste idazlanen datuak.
- B2 eta C1 mailen arteko konparaketa. Bildutako informazioa kalkulu-orrien bidez taxutu eta antolatu dugu B2 eta C1 mailen ezberdintasunik ote zegoen antzemateko. Era berean, etiketa esanguratsuenak adieraz dezaketenen araber bereizi ahal izan ditugu.
- Etiketen multzokatzea. Azkenik, etiketak multzotan ere bildu ditugu irizpide aproposak aintzat hartuta. Oraingoan, egitura oinarritzkoenetik konplexuenetarako gradazio bat proposatu dugu. Hots, ortografiatik diskurtsorako bidean diren gainerako osagaiak multzotan uztartu ditugu: ortografia, lexikoa, izena, aditza, esaldia eta diskurtsoa.
- Datuen bistaratzea. Datuak grafikoetan jarri ditugu begi kolpe batez nabarmenago azalduko ditugulakoan. Hurrengo lerroetan emango ditugu lortutako emaitzen berri (ikus 3. azpiatala).

3. LORTUTAKO EMAITZAK ETA ANALISIA

3.1. AURREIKUSPENA (MAILETAKO EZAUGARRIEN ALDERAKETA)

Hizkuntzaren gaitasun-maila neurtzeko erabiltzen diren irizpideetan arakatu ostean, B2 eta C1 mailetako ezaugarriak ondo-ondo jartzeko ahalegina egin dugu idazlanetan auresangarriak zitezkeen etiketak aurkitzeko.

B2 eta C1 mailetan landu behar diren hizkuntza-formen zerrenda eskerga horretatik⁶ irizpide hauek bateratu ditugu hipotesi gisa baliatze aldera, alegia, datuetan ere halako alderaketa egin zitekeelakoan:

- mendeko perpausen aniztasun handiagoa (C1): erlatibozkoak, konparaziozkoak, baldintzazkoak, denborazkoak...
- lokailuen erabilera zehatzagoa, askotarikoagoa (C1),
- hitanoaren erabilera (C1),
- subjuntiboaren orainaldia (B2) // lehenaldia (C1),
- adizki trinkoaren erabilera askotarikoagoa (aspektua, denbora...) (C1),
- zenbait postposizio (C1),
- eta H-ak, a itsatsia, puntuazio-markak, ortografia oro har (C1).

3.2. DATUAK OSOTARA

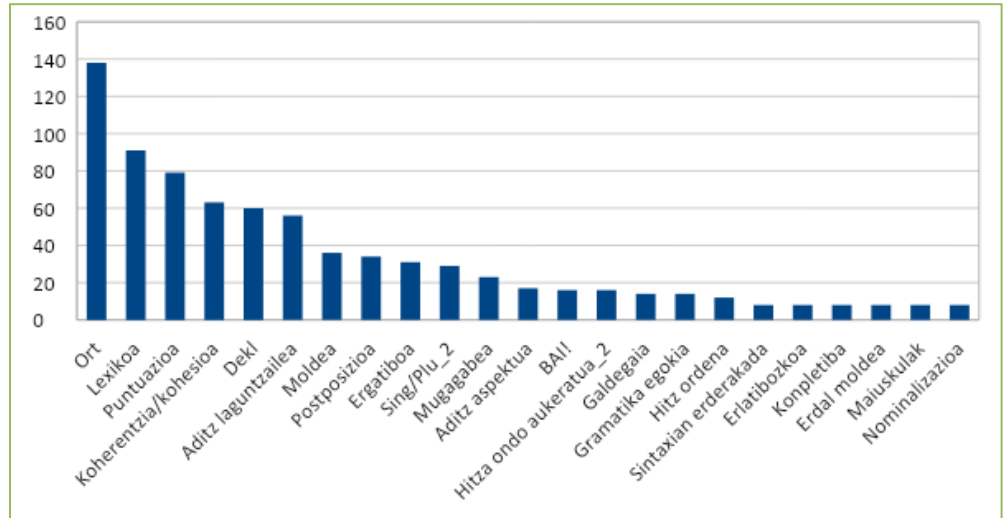
Hauek dira agerpen-kopuruari edo maiztasun hutsari begira corpus osoan gehien ageri diren etiketak (ikus 3. irudia):

Ezarritako etiketen agerpen-kopuruak argi ematen du aditzera zenbait alderdi ohikoa-goak direla aztertutako idazlanetan (40ko maiztasuna gainditzen dutenak): ortografia, lexikoa, puntuazioa, koherentzia/kohesioa, deklinabidea eta aditz laguntzailea. Bestalde, beste maila batean leudeke, 40 baino gutxiago eta 20ko maiztasuna gainditzen dutenak: moldea, postposizioa, ergatiboa, singular/plural komunztadura eta mugagabea. Horrek azterketari gaitasunaren berri ematen duen arren, beti ere, bigarren ol-dotzeta bat behar luke akats bakoitzaren edo zehaztapen bakoitzaren balioa ebaluatzeari dagokionez. Hala ere, aurrera begira eta corpusa deskribatzerakoan baliagarria den irudia delakoan gaude.

Bestalde, goiko zerrendan agerpen adierazgarrienak bakarrik jaso direla ere aipa dezagun; izan ere, analisirako horien azterketa baino ez dugu egin. Hemen zerrenda osoa agerpen-kopuruarekin: <https://b2drop.eudat.eu/s/RrnY3gmtKoXLWwa>.

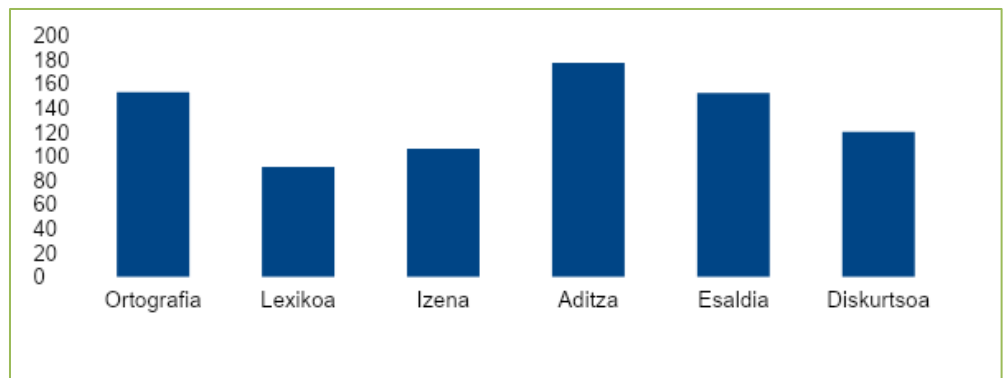
⁶ <https://b2drop.eudat.eu/s/iN4GNMkWjgARR6t>

3. irudia: etiketaturiko akats-zerrenda eta horien maiztasuna Markin tresnarekin.



Analisiari begira eta irudi argiagoa eskaintze aldera, oharrak taldekatzearen onurei begira jarri gara. Irizpidetzat hartu duguna oinarritzko elementuetatik konplexuagoetara doan bidean mailakatzea izan da. Era horretara, ortografia mailako akatsetatik diskurtso mailakoetara doazen markak honela multzokatu ditugu: ortografia, lexikoa, izena, aditza, esaldia eta diskurtsoa. Hara zer banaketa duten bilduma osoa kontuan hartuta (ikus 4. irudia)

4. irudia: Markinekin etiketaturiko akats-zerrenda multzokaturia eta horien maiztasuna.



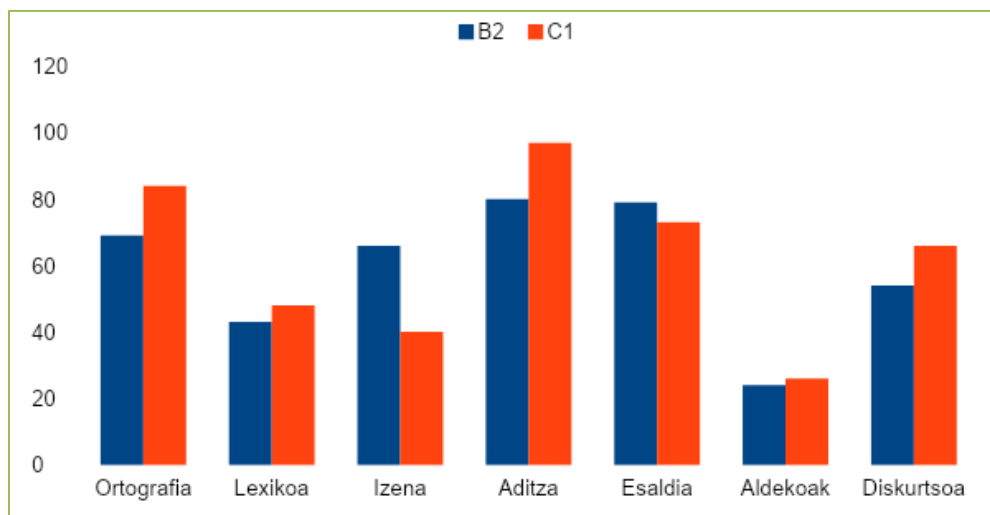
Grafiko honek ikuspegi orekatuagoa eskaintzen du. Orobat, aipagarriak izan daitezkeen beste ideia batzuk ere azalera ditzake. Alegia, badirudi komunztadurak (aditza), sintaxiak (esaldia), eta puntuazioak toki nabarmena hartzen duela hobetu beharreko gaitasunen arloan. Nahiz eta ortografia-akatsak marka bakar gisa gainerako ohar guztien agerpenak aldearekin gainditu, kopuruari bagagozkio.

Hasieran, programazioetako datuak bilduta aurrean zitezkeenarekin uztargarria da B2 eta C1 mailetan diskurtso-mailako konplexutasuna areago adierazteko gai izan behar baitute. Are gehiago, datuek proposatutako ezaugarriekin edota azterketariak erakutsi beharreko gaitasunekin bat egiten dutela esan daiteke, hein batean bada ere.

3.3. MAILA BIAK BEREIZTEKO SAIAKERA

B2 eta C1 hizkuntza-mailaren arabera alderaketa egiteari ere heldu diogu. Ez alferrik, hasierako xedeetarikoa bat mailaren arabera idazlanen edota haietan egindako akatsen nolakoa ikusita haiek egoki deskribatzea eta identifikatzea ere bazelako. Aurreneko ahalleginean multzokatutako datuak hartu ditugu baliagarri izango direlakoan (ikus 5. irudia).

5. irudia: B2 eta C1 mailan etiketaturiko akats-zerrenda eta horien maiztasuna.

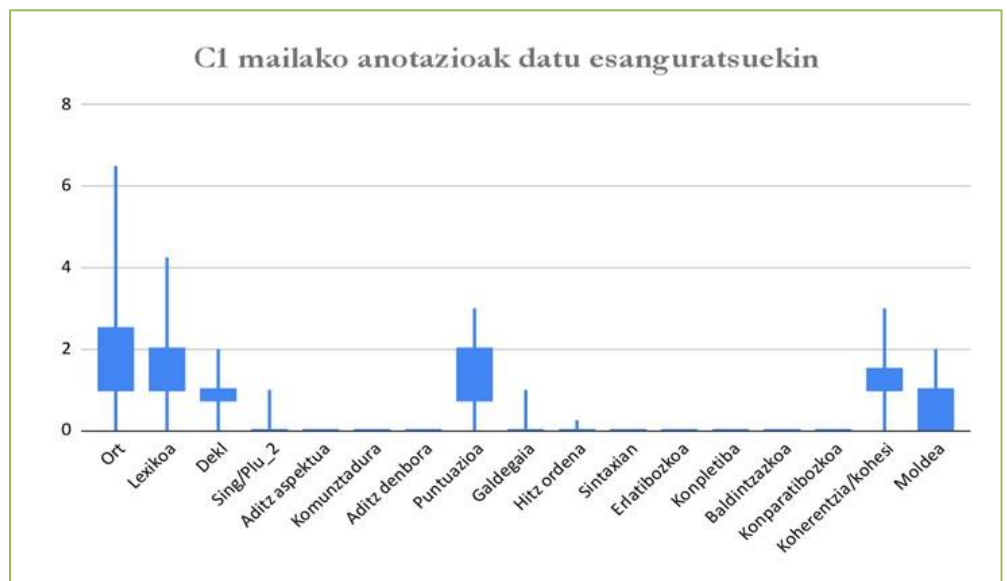
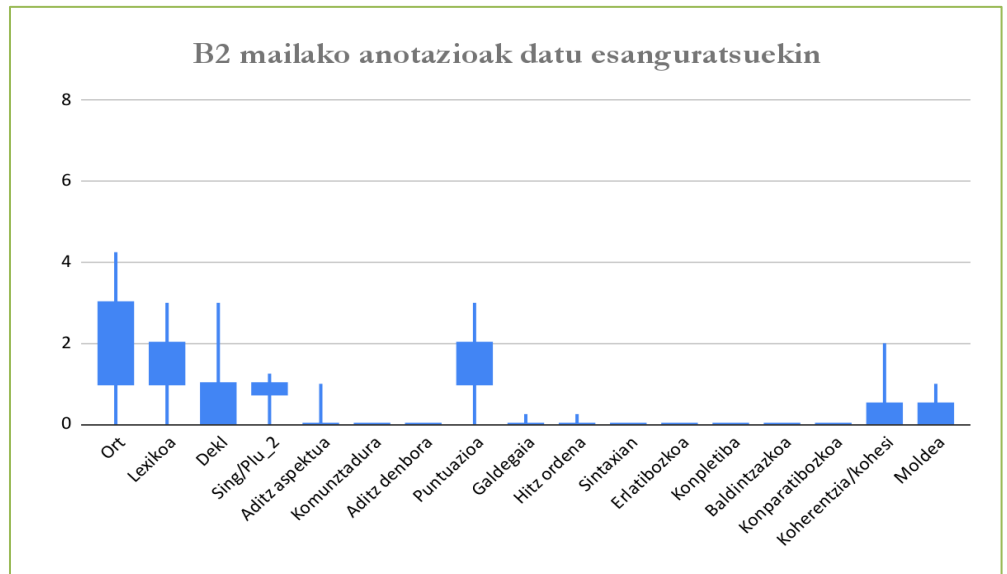


Oro har, C1 mailako testuetan ortografia, lexiko, aditz eta diskurtso mailako marka edo etiketa gehiago esleitu ditu zuzentzaileak. B2ko idazlanetan, berriz, izen eta esaldi mailako akatsak gehixeagotan agertu dira.

Kopuru osoa edo batura kontuan hartuta baino emaitza esanguratsuagoen bila zenbait kalkulu egin ditugu. Horrela, agerpen-kopuru minimoa, maximoa, mediana, lehen pertzentila eta hirugarren pertzentila neurtuta bestelako interpretazioak egiteko modua topatu dugu.

Jarraian eskainitako 6. irudian X ardatzean gora ageri diren kaxa-diagramek akats hori delako mailako testu guztietan agertzen dela adierazten dute (gutxienez behin edo gehiagotan). Kaxaren luzerak, aldiz, kopuru maximoaren eta minimoaren arteko barreiatzea. Beherago, adierazgarrienak iruditu zaizkigun marken datuak ekarri ditugu grafikora, hots, alderatzeak errazago egiten laguntzen dituztenak.

6. irudia: B2 eta C1 mailan etiketaturiko akats-zerrendaren kaxa-diagramak.



B2 zein C1 mailetan agertu diren kaxa bereizita elkarren arteko antzekotasunekin batera zenbait aparteko ezaugarri ere ageri dira. Uztartuta ezarritako kaxa-diagramak aurkeztu ostean azalduko da horien arabera egin dugun interpretazioa. Lehenago ere antzeman izan dioguna azaleratzen da grafiko pare honetan, hots, azterketarako zein interpretaziorako bide ematen duten markak begien bistara ateratzen ditu. Ari gaitetzen, hortaz, ortografia, lexikoa, deklinabidea, singular/plural komunztadura, aspektua, puntuazioa, galdegaia, kohesioa/koherentzia eta moldea etiketen arteko antzekotasunez zein ezberdintasunez.

7. irudiko aurreneko etiketari so (Ort-B2 eta Ort-C1), alegia, orotara agerpen-kopuru handiena duenari begira, gutxieneko agerpena, oro har parekoa da nola B2 mailan hala C1en. B2ko idazlanetan proportzioan akats-kopurua handiagoa da (2,5 vs 3) baina C1eko zenbait idazlani ezarri zaio era honetako akats-kopururik handiena. Lexikoari dagozkion datuetan antzekotasun handia dago bi mailetan, aipatzekotan, akats-kopuru maximoa C1 mailako azterketaren batean neurtu dela esan daiteke.

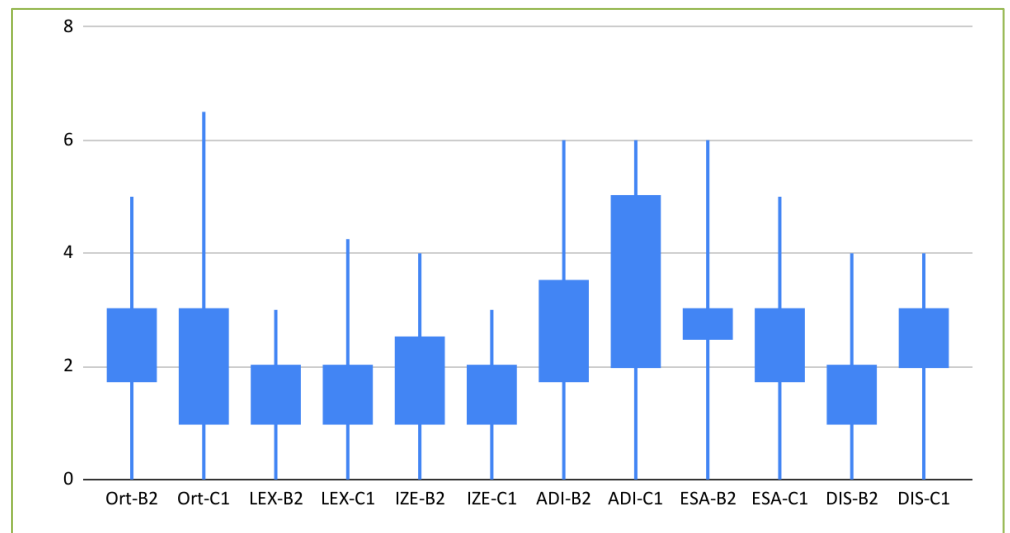
Deklinabidea etiketaren esleipenetan B2 mailan jarritakoen proportzio handiagoa eta C1 mailan gutxieneko agerpen altuxeagoa ikusi dugu. Alderantziko joera nabari da singular/plural (sg/pl) komuntadura kontutan; izan ere, multzo sendoagoa dago B2 mailako agerpenetan. Agerpen gutxiagoko etiketen multzo honetan aditz aspektuarena sar daitekeelakoan gaude, zeinak agerpen-kopuru txikia duen C1 mailan eta handixeagoa B2an aurreikus zitezkeen gisan, bide batez. Izan ere, oinarrizkoagoa den gaitasuntzat jotzen da azken hau.

Puntuazioa ere maiz markatzen den ezaugarria da. Parekotasun handia dago bi mailetan, haatik, C1 mailan barreiatuago dago agerpen-kopuru hori, hau da, agerpen-kopuru minimoa txikiagoa da. Har dezagun tartetxo bat hemen galdegaia etiketapeko akatsak C1en maizago ageri direla esateko.

Azken oldozketa gisa, kohesio/koherentzia eta moldea etiketetan antzemandakoez oharren bat edo beste. Lehenbizikoak gutxieneko agerpen-kopuru handiagoa du proportzioan idazlanez idazlan C1 mailan. Gainera, horietakoren bati ezarri zaio akats-kopuru maximoa ere. Moldea markapeko akatsei dagokionez, C1 mailan gehiagotan ageri direla esan daiteke medianari, 3. pertzentilari eta maximoari begira. Eskatu ere hala egiten zaio maila aurreratuko ikasleari, esan nahi baita, konplexutasun handiagoko testuak egiten saiatzeak ahaleginean hutsegite-kopuru handiagoa ekartzea ez da harrizkekoa.

Hurrengo grafikoa etiketa-multzoak mailaka konparatzeko baliagarria izan daiteke, oinarrizko elementuetatik hasi eta konplexuenerako bidean bi mailotako idazlanak ebaluatzerakoan azaleratu diren ezberdintasunak ageriago daudelakoan (7. irudian).

7. irudia: B2 eta C1 mailan etiketaturiko akats-zerrenda multzoaren kaxa-diagramak.



7. irudian 6. irudian baino gorago daude datu gehienak. Ez da bereziki harrizkekoa akats-multzoez ari baikara. Kopuru osoa 6. irudian aztertu dugu eta hango datuak ez dira gezurtatzen honela bistaratuta, 7. irudian. Haatik, badago maximoen eta minimoen arteko alderatzea aztertzerik. Horrela, ortografian C1 mailako akats gehiago dago baina gutxieneko akats-kopuru handiagoa da B2 mailako lanetan. Gauza bera esan daiteke esaldi mailako oharrei so.

Ez da lan honen xedea gehiegizko zehaztapenetan sartzea. Hala ere, azken gainbegirada bat emanik zenbait joera gaingiroki aipatzerik badelakoan gara.

Gure lagineko C1 mailako testu gutxi dago ortografia-, lexiko-, deklinabide-, puntuazio- eta kohesio-arazorik gabe. B2 mailan ere antzerako deskribapena egin daiteke, haatik, singular/plural komunztadura eta puntuazio aldetikoak erantsi eta bereziki aipatu beharrekoak dira maila honetako lanetan.

Ugaritasunean eta maiztasunean ortografiak, lexikoak, puntuazioak, komunztadurak, aditz denborak eta esaldi mailako akatsek bereizten dituzte C1eko testuak. B2koetan, berriz, deklinabidea, galdegaia, konpletiboak, baldintzakoak eta kohesio- nahiz koherentzia-arazoak ohikoak direla ondoriozta daiteke.

3.4. PARAGRAFOEN ETIKETATZE AUTOMATIKOAZ

Idazlan hauek paragrafoz paragrafo automatikoki etiketatuta jasotzeko modua ere izan dugu. Horretarako, ikasketa gainbegiratu algoritmoak erabili ditugu: *Euskarri Bekto-redun Makina* (EBM edo ingelesez, SVM). Datu-bektoreak lortzeko, testuei ezaugarri linguistiko zenbait erantsita, testu bakoitza ezaugarri-bektoreekin adierazi dira. Sailkatzailak eta hizkuntza-ereduak entrenatzeko hizkuntza-ereduak *Python* programazio-hizkuntzako *Sci-Kit Learn* liburutegia (Pedregosa et al., 2011) erabili dugu eta EBM sailkatzailerako kernel funtzio lineala erabili dugu, erregularizazio parametroa $C=1$.

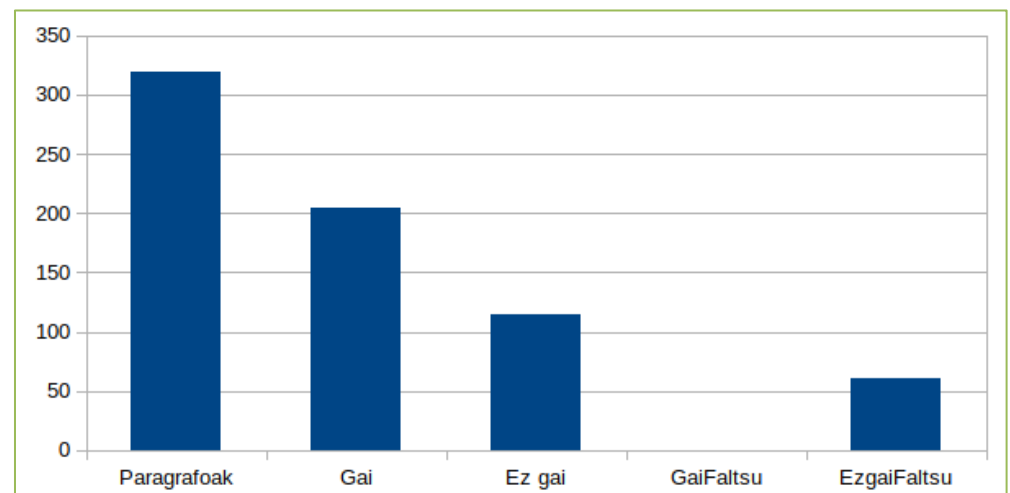
Horrez gain, sailkatzaila automatikoak eskainitako irizpenak ebaluatzeko ohiko irizpi-deak jarraitu ditugu, hots, doitasuna, estaldura eta f -neurria baliatu ditugu.

Tresnak eskainitako ebaluazioa testu osoari zegokion, eskuarki, bat egiten genuen esleitutako emaitza horiekin eta, beraz, ez dugu gai ziren paragrafoen azterketa xehegorik egin. Horrela, doitasunari dagokion emaitzak zehaztasun handiko tresnatzat jotzen du sailkatzaila: $gai / (gai + gai faltsuak) = 1$.

Estaldurari so hobetu beharrekoak badirela sumatu dugu. Izan ere, badira zuzentzat ditugun eta tresnak txartzat jotzen dituen zenbait paragrafo. Hona emaitza: $gai / (gai + ezgai faltsuak) = 0,77$.

Aipatutako bi kontzeptuak biltzen dituen f -neurri estandarra ere atera diegu lagin kokor honetan lortutako emaitzei. Hona hemen zein balio atera zaigun: $2 * (doitasuna * estaldura) / (doitasuna + estaldura) = 0,87$.

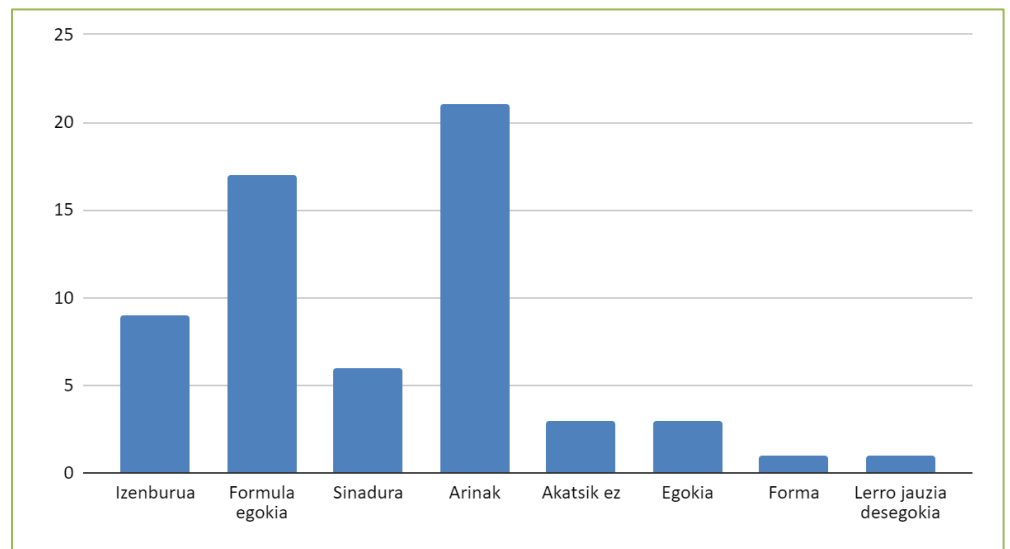
8. irudia: Paragrafoen etiketatze automatikoaren eskuzko ebaluazioa.



Azkenik, mailaz beherako paragrafo gisa esleitutako testu txatalak ebaluatu ditugu gain-giroki. Zutabe bakoitzaren ondoan aldeko edo kontrako irizpen bat ezarri eta *ezgai faltsuak* ekidin ahal izateko etiketa-sorta bat proposatu dugu:

- *Arinak* (ez oso garrantzi handiko akatsa edo akatsak);
- *egokia* (egoki jarri du sistemak maila eta ez dago okerririk);
- *formula egokia* (agur formula batetik bat);
- *izenburua* (artikuluetako lerroburuak);
- *sinadura* (izena edo sinadura bada testua);
- *forma* (akats arraroa edo digitalizatzerakoan akatsen bat ote dagoen).

9. irudia: Paragrafoen etiketatze automatikoaren ez-gai faltsuen ebaluazioa.



9. irudian ikus daitekeen moduan, akats arinak aurkitu dira sailkatzaileak ez-gaitzat hartu dituen paragrafoetan. Horrez gain, sailkatzaileak paragrafo bereziak ere hartu ditu ezgai faltsutzat, horietako gehienak izenburuak, agur formula laburrak (formula egokia) edo sinadurak dira.

Emaitza nagusiekin grafiko hau nahiko adierazgarria iruditu zaigu sailkatzaile automatikoaren gaitasunaren alderdi bat bederen aditzera emateko. *Ezgai faltsuen* atala zehatzago aztertzearen zakuak utzi dugun arren, oro har, nahiko egoki etiketatze gaitasuna dela iruditzen zaigu.

4. ONDORIOAK ETA ETORKIZUNEN LANAK

Euskarri digitalera aldatutako iturriekin analisiak egiteak dauzkan abantailak aldarrikatzeko, hein batean eskuz edo gizakiaren begizko irizpenetan oinarritutako saiakera aurkeztea kontraesankorra izan daiteke. Alabaina, prozesamendu automatikoaren oinarrian gizakiaren irizpideak eta gidaritzak jarri ohi dira, zeina ziber-munduaren adierazgarri izan daitekeen (Arriola eta Iruskieta, 2021).

Etiketak modu deduktiboan kategoriza genitzakeen baina adituen irizpenen gainean ezarri dugu oinarria ahaleginaren neurri xumea aintzat hartuz gero, eginbiderik egoki iruditu zaigulako eta edozein irakasle erreproduzitu dezakeelako.

Ondorio gisa, eskuratutako datuak auresangarri izan daitezkeen arren, interpretazio egokiak azterketa sakonagoa behar duela ikusi dugu. Hein batean bada ere, konplexutasun handiagokoak izan daitezkeen ezaugarrietan hauteman ditugu C1eko ikasleen hobetu beharreko alderdiak eta oinarritzakoagoetan B2koenak. Aditzaren eta esaldiaren

esparruetako akatsen nolakotasuna neurtu behar litzateke baina eskatzen zaien konplexutasun-mailaren adierazgarri ere izan liteke era honetako marken agerpen-kopurua areagotzea C1eko idazlanetan. Hauek dira maiztasun handienarekin esleitu diren markak: *Ortografia, lexikoa, puntuazioa, koherentzia/kohesioa, deklinabidea, aditz laguntzailea, moldea, postposizioa, ergatiboa, sing/plu komunztadura, mugagabea, aditz aspektua, galdegaia, hitz ordena*. Zerrenda horrek bat egiten du hizkuntzen eskoletan ohikoenak diren zuzendu beharreko elementuekin. Emendatu beharreko beste ezaugarri-sorta esaldi mailako akatsena litzateke zeina barreiatuago ageri den, baina multzo gisa kopuru esanguratsua osatzen duena.

B2 eta C1 mailen arteko alderaketa egin nahiko balitz, ortografia, aditz laguntzailea, moldea, ergatiboa, sin/pl komunztadura, mugagabea eta aditz aspektua markek ematen dizkigute begien bistara nabarmenenak diren alderatzeak.

Dagoeneko abian den ekimen zabalago baten hegian edo baztertxoan osatu dugu lana. Testuak automatikoki kategorizatzen duten tresnak asko garatu dira azken urteotan eta lagin koxxor honek haien emaitzak baloratzeko egin daitekeenaren zantzuren bat proposatu nahi zuen. Ildo horri jarraiki, hemen aztertutako testuetako paragrafoak automatikoki nola ebaluatu diren ere begiratu dugu (hurrengo argitalpen batean emango dugu horren berri zehatz eta mehatz) eta haien inguruko zuzenketa-proposamenen bat edo beste iradokitzen ere ahalegindu gara (hemen dokumentu etiketatua).⁷ Hainbat urrats eman daitezke xede horiek erdieste bidean, besteak beste, hezkuntzako eragile ugariak egunero sortzen dituzten testuak eta askotariko datuak biltzeko eta corpus erabilgarri gisa baliatzeko bitartekoak jar daitezke. Jardunbide egokiak proposa genitzake datu horiek ondo edo erabilgarri izan daitezkeen prozedurekin jaso daitezkeen. Bestalde, bildutako testu horiek analizatuko eta aztertuko dituen jendea zein haienezko bitartekoak ere egoki antolatuz gero, lortutako emaitzek zinezko laguntza pedagogikoa eskainiko liekete egunero ikas-irakaskuntza eginkizunetan diharduen banakoari zein sistema osoari.

Bizi dugun mundua eta eraikitzen ari garena arriskuz eta aukeraz beteta dago. Teknologia bizi ikuskera horien oinarriak aldatzen dituzte eta aldatuko dituzte. Nora bultza dezakegun eta nahi dugun pentsatu behar genuke eta ekin kontzientzia osoz. Zer da automatizatu nahi genukeena, errazago, zuzenago edo zorrotzago osa daitekeena egun baino eta zer ez genuke arau zorrotz horien mende jarri nahi? Zer da gizatiarra eta zer makinaren eremukoa? Ba al da hibridatzerik, *ziborg* izatea nahi ote genuke?

Gure lan-arloan, hezkuntzan, eta zehazki hizkuntzen irakaskuntzan badira ikaskuntza-prozesua automatizatzeko ahaleginak:

- *E-learning*a: erritmoa, beharren araberako, bitartekoen konbinazioa, ikasle gehiago aldi berean, azkarrago ikasi?, elkarrekintza elementuak, malgua...
- Bakarka ikasteko metodoak ez al dira irakaslearen egitekoa zalantzan uzten duten ikuspegiak?
- Gero eta gehiago dira ordu gutxiagotan errazago ikas daitekeela iragartzen duten eskaintzak.
- Hainbat esparrutako irakasleok jardun biko ikastaroak eskaintzen ditugu eta bakarka osa daitezkeen ariketaz jositako online ikastaroak proposatzen dizkiegu ikasleei.
- Zer esanik ez erdi jolas eta erdi negozio diren dozenaka aplikazio horiek *Duolingo* edo antzekoak. Edo *Ingura* proiektua (Ertzibengoa, 2021) eta antzekoak.

⁷ <https://b2drop.eudat.eu/s/PdJQ34774JasnCX>

Oraingoz, erakundeek ebaluaziorako pieza nagusitzat gauzkate irakasleok eta ez dirudi hori galtzeko arriskuan dagoenik. Baina datu-bilketa eraginkorrik ez da egiten. Ezabatu egiten dira eta enkriptatu beharrik ez da (hezkuntzan edota osasungintzan). Datu horiek anonimotasunetik azter daitezke, digitalizatu eta interpretatu banakoen nahiz taldeen eskubideak urratu gabe. Ildo horri jarraiki, erakunde pribatuen esku-hartzea, nagusitzea eta datuen ustiaketa arrisku gisa aintzat hartu beharko litzateke eta egiteko horiek komunitatearen esku utzi. Alegia, datuen bilketak zein haien analisiak eta emaitzen ebaluazioak banako ororen eskura behar luke egon, ikerketa-azpiegitura publikoen laguntza baliatuz.

Azterketa eta ariketen emaitzen azterketa eta aurre-ebaluazioa, proposamenak, programazio egokituak, irakas-prozedura aproposak. Horiek, hein batean bada ere, konputazio-tresnen esku uztea aproposa litzateke lanak eraginkorrak izanik gizakiaren kontroltik kanpo geratu gabe.

Azkenik, bestalde, emaitzen araberako erabakiak, programazioetako joera-aldaketak, alderdi osagarriak, gelako jarduna, unean uneko erabakiak edo ohiturak bezalakoak irakaslearen edo hezkuntza taldearen eskumeneko behar lukete eta horretarako bermea ezarri beharko litzateke.

Eskertza

Eskerrak eman nahi dizkiogu HABE erakundeari datuak lagatzearen (HABEren eta UPV/EHUren arteko akordioaren babespean). Ixa Taldeari sailkatzaileen emaitzak eskaintzearen, bereziki Ekain Arrietari.

Bibliografia

- Azpillaga, X. (2022). Euskarazko testuen komunikagaitasun-maila automatikoki sailkatzeko lehendabiziko urratsak. *e-Hizpide*, 99. <https://doi.org/10.54512/YZBR4203>
- Bauman, Z. (2003). *Modernidad líquida*. Fondo de Cultura Económica.
- Arriola, J.M., Goikoetxea, J. eta Iruskieta, M. (2020). Hizkuntza-teknologiak hizkuntzen ikas-irakaskuntzan: zenbat aukera, hainbat erronka. *e-Hizpide*, 95. <https://labur.eus/IOPZ3>
- Arriola, J.M., eta Iruskieta, M. (2021, maiatzaren 20a). Zenbateraino aldatuko dituzte hizkuntza-teknologiaren forma berriek gure komunikazioa eta hizkuntza bera? [Sarrera]. *Irakasbil*. <https://labur.eus/EFXfH>
- Busa, R. A. (2004). Foreword: Perspectives on the Digital Humanities. In S. Schreibman, R. Siemens eta J. Unsworth (ed.), *A Companion to Digital Humanities* (xvi xxiii. or.). Blackwell.
- Camacho, A., & Iruskieta, M. (2021). Euskararen i(ra)kaskuntza-prozesuak: hezkuntza eta hizkuntza teknologiak. *Tantak. EHuko Hezkuntza Aldizkaria*, 32(2), 9–31. <https://doi.org/10.1387/tantak.21654>.
- Chen, X. eta Meurers, D. (2016). CTAP: A web-based tool supporting automatic complexity analysis. In *Proceedings of the workshop on computational linguistics for linguistic complexity at COLING*, 113-119. <https://doi.org/10.17863/CAM.39630>
- Ertzibengoa, A. (2021). Ingura, euskara online ikasteko sistema berria eraikitzen. *e-Hizpide*, 99. <https://doi.org/10.54512/CDLS3475>
- Europako Kontseilua. (2005). *Ikaskuntza, irakaskuntza eta ebaluaziorako Europako Erreferentzia Marko Bateratua*. HABE eta Eusko Jaurlaritzaren, Hezkuntza, Unibertsitatea eta Ikerketa Saila. <https://doi.org/10.54512/EEMB-2005>
- Fyfe, P. (2011). Digital pedagogy unplugged. *Digital Humanities Quarterly*, 5(3). <http://www.digitalhumanities.org/dhq/vol/5/3/000106/000106.html>
- Gonzalez-Dios, I., & Bengoetxea, K. (2021). MultiAzterTest@VaxxStance-IberLEF 2021: Identifying Stances with Language Models and Linguistic Features. *IberLEF@SEPLN* (192-201. or.).
- Isasi, J. eta Iruskieta, M. (2019). Humanitate Digitalak eta haien ikerketaren nondik norakoak. In M. Iruskieta, M. Maritxalar, A. Arroyo eta A. Camacho (ed.), *IKTak eta kompetentzia digitalak hezkuntzan* (307-332. or.). UEU eta UPV/EHU.
- Kirschenbaum, M. G. (2010). What Is Digital Humanities and What's It Doing in English Departments? *ADE Bulletin* 150, 55–61. <https://doi.org/10.1632/ade.150.55>
- Lecarpentier, D., Wittenburg, P., Elbers, W., Michelini, A., Kanso, R., Coveney, P. eta Baxter, R. (2013). EUDAT: a new cross-disciplinary data infrastructure for science. *International Journal of Digital Curation*, 8(1), 279-287. <https://doi.org/10.2218/ijdc.v8i1.260>
- Otegi, A., Imaz, O. Díaz de Ilarraza, A. Iruskieta, M. eta Uria, L. (2017). ANALHITZA: a tool to extract linguistic information from large corpora in Humanities research. *Procesamiento del Lenguaje Natural*, 58, 77-84.