

# GRADU AMAIERAKO LANA

## ITZULPENGINTZA ETA TEKNOLOGIA

### WIKIPEDIAN POSTEDITATZEKO GIDALERROAK

eman ta zabal zazu



Universidad  
del País Vasco

Euskal Herriko  
Unibertsitatea

Amaia Solaun Martínez

Itzulpengintza eta Interpretazioko Gradua

2020/2021 Ikasturtea

**Tutorea:** Elizabete Manterola Agirrezabalaga

**Saila:** Ingeles eta Aleman Filologia eta Itzulpengintza eta Interpretazioa

## 1. LABURPENA

Lan honen helburua postediziorako (PE) gidalerro batzuk sortzea izan da, Wikipedian gaztelaniatik euskarara posteditatzeko. Hori lortzeko, postediziorako gidalerro minimo batzuk sortu ditugu, eta postedizio proba batean testatu ditugu, 7 wikilariren laguntzarekin. Postedizio proban bi talde egin dira, eta horietako bakoitzak testu bat posteditatu behar izan du. Ondoren, posteditatutako testuen akatsak anotatu eta sailkatu dira. Azterketa horretako emaitzetan oinarrituta, hasierako gutxieneko jarraibide horiek hobetzea proposatu da. Halaber, postedizio proban parte hartu dutenei galdetegi bat pasatu zaie, laginaren ezaugarrien berri izateko. Galdetegi horretan informazio demografikoa jasotzeaz gain, euskararen ezagutza maila eta erabileraren inguruko informazioa jaso da, eta baita postedizio eta itzulpen lanetan izandako esperientziaren ingurukoa ere. Faktore horiek behatu dira postedizio emaitzetan eragin zezaketelako.

Lana berritzailea da, itzulpen automatikoak Wikipediaren lan-fluxuan denbora luzea daraman arren, orain arte ez zegoelako postediziorako gidalerrorik, ez euskaraz ez inguruko erdaretan ere. Gainera orain arte euskaraz ia joratu ez den gai batean jartzen du fokua: *crowdsourcing* plataformetako postedizioan. Hortaz, lan honetan postedizio profesionaletik aldendu eta bere beste aplikazio batean zentratu gara. Dena den, postedizio profesionalarekin ezaugarri asko partekatzen dituenetz, eta bien garapena ia aldeberekoa izaten ari denez, lan honen atal teorikoan bai postedizio profesionalaz, eta baita *crowdsourcing* plataformen postedizioaz aritu gara.

Lana hiru bloketan egituratu dugu. Lehenengo blokea postedizioaren alderdi teorikoari eskainiko diogu eta hiru ataletan banatua dago: lehenik, azalduko dugu zertan datzan postedizioa, bigarrenik, ikusiko dugu zer diren *crowdsourcing* plataformak eta zer berezitasun dituen plataforma horietan egindako postedizioak, eta azkenik behatuko dugu zertan den euskal postedizioa. Bigarren blokean, egindako proban zentratuko gara, eta azalduko dugu, alde batetik, zertan oinarritu garen proba diseinatzeko; eta bestetik, zein izan den erabilitako metodologia. Azkenengo zatian probaren emaitzak eta amaieran sortutako gidalerroak aurkeztuko ditugu, eta azken ondorioak eskainiko ditugu.

## AURKIBIDEA

1.	LABURPENA .....	2
2.	SARRERA.....	5
3.	ZER DA POSTEDIZIOA? .....	6
3.1.	Itzulpen automatikoa eta postedizioa .....	6
3.2.	Postedizio motak eta posteditatzeko gidalerroak .....	6
3.3.	Postedizioaren eraginkortasuna.....	7
4.	CROWD POSTEDITING ETA COMMUNITY POSTEDITING .....	9
4.1.	Profil berria: kalitatearen auzia eta ataza desberdinak .....	9
5.	POSTEDIZIOA EUSKAL HERRIAN.....	11
5.1.	Itzultzaile automatikoen eboluzioa .....	11
5.2.	Formakuntza.....	12
5.3.	Postedizioaren erabilera .....	13
5.4.	Wikipedia: IA eta postedizioa.....	14
6.	EUSKAL WIKIPEDIAKO POSTEDIZIOA AZTERTZEN .....	17
6.1.	Oinarri teorikoa .....	17
6.2.	Itzultzaile-automatikoaren aukeraketa .....	19
6.3.	Testu laginen aukeraketa.....	20
6.4.	Pertsona-laginaren diseinua .....	21
6.5.	Gidalerro minimoen sorkuntza.....	22
6.6.	Ebaluazio sistema.....	23
6.7.	Egiaztapen proba.....	27
7.	PROBAREN EMAITZAK.....	28
7.1.	Galdetegiaren emaitzak eta probarako behar izandako denbora.....	28
7.2.	Akats anotazioa: azterketa kuantitatiboa.....	32
7.3.	Anotazioa: azterketa kualitatiboa .....	33
7.4.	Amaierako gidalerroak .....	38
8.	ONDORIOAK.....	40
9.	BIBLIOGRAFIA .....	41
10.	ERANSKINAK .....	44
10.1.	JATORRIZKO TESTUA: OMS .....	44
10.2.	JATORRIZKO TESTUA: GENEROA.....	45
10.3.	ITZULPEN AUTOMATIKOA: OMS .....	46
10.4.	ITZULPEN AUTOMATIKOA: GENEROA .....	47

## **TAULEN AURKIBIDEA**

Taula 1 TAUS eta Mitchellen gidalerroak.....	22
Taula 2 TAUS DQF-MQM Akats kategorien larritasun mailak. ....	24
Taula 3 De Almeida (2013) Postedizio aldaketan sailkapena. ....	25
Taula 4 Akatsen sailkapena larritasunaren arabera.....	26
Taula 5 Erabilitako itzultzaile automatikoak. ....	30
Taula 6 TAUS Itzulpen kalitatearen kalkulua. ....	32
Taula 7 Notazioaren emaitzak. ....	32
Taula 8 Mantendutako akatsak. ....	34
Taula 9 Gehitutako akatsak.....	36
Taula 10 Itzultzaile automatikoak izen bereziekin izandako arazoa. ....	37

## **GRAFIKOEN AURKIBIDEA**

Grafikoa 1 Euskararen erabilera maiztasuna. ....	28
Grafikoa 2 Euskara mailaren ziurtagiria. ....	28
Grafikoa 3 Euskara noiz ikasi dute? .....	29
Grafikoa 4 Euskara ama hizkuntza bada, etxean gurasoak euskaldunak dira?.....	29
Grafikoa 5 Itzulpen esperientzia. ....	30
Grafikoa 6 IAren izandako esperientzia (1etik 10era).....	31
Grafikoa 7 Ezagutu eta erabiltzen dituzten hizkuntza-baliabideak. ....	31

## **IRUDIEN AURKIBIDEA**

Irudia 1 Wikipedian argitaratutako deialdia. ....	21
---	----

## 2. SARRERA

Azken hamarkadan itzulpen automatikoaren (IA) garapenean aurrerapauso oso handiak eman dira; batez ere azpimarragarria da itzultzaile neuronalen sorrera. Itzultzaile horiek emaitza naturalagoak sortzen dituzte, arauetan oinarritutako IA eta IA estatistikoaren aldean. Beraz, IAk eta postedizioak (PE) inoiz baino aukera gehiago eskaintzen dizkigute.

Euskara ere ez da atzean gelditu eta egun hainbat itzultzaile neuronal daude eskuragarri, merkatu pribaturako nahiz publikorako. Horrek ateak ireki dizkio euskal PE profesionalaren garapenari, gero eta gehiago hedatzen ari dena enpresen artean. Hala ere, eraldaketa ez da esparru profesionalera mugatzen; euskaraz hiru itzultzaile neuronal baitaude edonoren eskura. Hortaz, pentsa daiteke itzultzaileak ez direnek ere erabiliko dituztela, eta IAren aplikazio berri eta desberdinak agertuko direla.

Hori dela eta, guk lan honetan erabaki dugu itzulpen profesionaletik apur bat aldentzea eta *crowdsourcing* motako plataformetako PEan murgiltzea, non edukia kolaboratzaile aditu eta ez-adituren elkarlanaz sortzen den eta itzultzaile profesional edo ez-profesionalek parte har dezaketen. Zehazki Wikipedia aztertzea erabaki dugu, plataforma ireki eta eleaniztuna delako, eta neurri handi batean itzulpenaz eta PEaz elikatzen delako. Gainera, euskara hizkuntza txikia izanik, funtsezkoa iruditzen zaigu euskarazko Wikipediarako elikatze-prozesua erraztea, eta uste dugu PEa baliabide interesgarria izan daitekeela horretarako.

Gure ikerlana gidalerroen sorkuntzan zentratu dugu, gaztelaniatik euskararako norabidean. Izan ere, gidalerroak PERako ezinbesteko baliabide bat dira, baina oraindik ere gidalerro publiko gutxi daude. Gainera, Wikipediaren kasu konkretuan ez zegoen gidalerrorik ez euskararako ez beste hizkuntzetarako ere.

Lana hiru bloke nagusitan banatu dugu. Lehenengoan lanaren alderdi teorikoa landuko dugu eta azalduko dugu zertan datzan PEa, zer nolako berezitasunak dituen *crowdsourcing* motatako PEak eta zein den egun PEak Euskal Herrian duen egoera (2, 3, eta 4 atalak). Bigarren blokean, lanaren muinari helduko diogu: diseinatutako probaz arituko gara, eta xehe-xehe azalduko dugu zertan oinarritu garen proba diseinatzeko eta nolako metodologia erabili dugun (5. atalean). Azken atalean, egindako probaren emaitzak aztertu eta azkenean lortutako gidalerroak aurkeztuko ditugu (6 eta 7 atalak).

### **3. ZER DA POSTEDIZIOA?**

#### **3.1. Itzulpen automatikoa eta postedizioa**

Itzulpen automatikoa (IA) hizkuntza batetik bestera itzultzeko prozesua osorik edo zati batean automatizatzean datza (Arnold et al, 1994, 1 or.), edota MemoQ-k bere weborrian adierazten duen bezala: software batek, giza laguntzarik izan gabe, testu bat hizkuntza batetik beste batera itzultzea (MemoQ, d.g).

Hasiera batean IA giza itzulpenaren ordezkapena izatea bilatzen bazen ere, egun onartzen da IAren emaitza bere horretan ez dela nahikoa, eta giza itzultzaile batek aurreitzulpen hori orraztu behar duela ezinbestean. Profesional horren eginkizuna testua posteditatzea da, hau da, itzultzaile automatikoak aurre-itzulitako testua editatzea, zuzentzea edota moldatzea (Allen, 2001).

Postedizioaren (PE) abantaila nagusia da errendimendua eta ekoizpena handitzen direla, betiere output ona eskaintzen duen itzultzaile automatiko bat erabiltzen bada (Collantes et al, 2018). Itzulpen bat egiteko behar den denbora murrizten da, eta, ondorioz, denbora-tarte berdinean eduki gehiago itzuli daiteke. Horrek guztiak ere aukera ematen die enpresei euren etekinak handitzeko eta baita itzulpen epe motzagoak eskaintzeko ere (Collantes et al, 2018).

Dena den, onartzen da posteditatutako IAk ez duela giza itzulpenaren maila bera estilo aldetik (TAUS, 2016, 18 or.), eta hori oso garrantzitsua izan daiteke testu mota batzuetarako, adibidez, publizitate eta literatura testuetarako. Beraz, denbora aldetik PE abantaila bat izan badaiteke ere, testuaren kalitateari begira ez du zertan hala izan.

#### **3.2. Postedizio motak eta posteditatzeko gidalerroak**

Gidalerroak sortzerako orduan kontuan hartu behar da posteditatzeko behar den esfortzua, zeina bi faktoreren menpe dagoen: itzultzaile automatikoaren emaitza gordinaren kalitatea eta zein den markatutako helburua. Azken horren arabera, bi PE mota bereizten dira: postedizio osoa eta postedizio arina (TAUS, 2016, 16 or.).

PE arinak edo *light post-editing*ak amaierako testu ulergarri eta erabilgarri bat sortzea du helburu, nahiz eta ez den ez hizkuntza aldetik, ezta estilo aldetik ere perfektua.

Normalean irakurleak atzeman dezake IA posteditatu bat dela (DePalma, 2013 in Hu eta Cadwell, 2016, 347 or.).

PE osoa edo *full post-editing*ean giza itzulpenaren pareko kalitatea duen testua sortzea bilatzen da. Helburua da giza itzulpenetik ezberdindu ezin den testu bat sortzea, hizkuntza aldetik eta estilo aldetik zuzena dena (DePalma, 2013 in Hu eta Cadwell, 2016, 347 or.). TAUSen arabera publikatzeko moduko testuak sortzea da helburua (2016, 16 or.).

Zehaztutako helburu hauetara iristen laguntzeko posteditoreek askotan gidalerroak baino ez dituzte izaten (Aranberri, 2016, 198 or.), askotan ez baitute posteditatzeko formakuntza espezifikorik jaso. Gainera, oso gidalerro gutxi daude publikoki eskuragarri, gehienak enpresa barnean erabiltzekoak baitira, isilpekoak (Aranberri, 2016, 198 or.).

Badira nolnahi ere publikoki eskuragarri dauden gidalerroak. Microsoftek, besteak beste, 2008. urtean tarte labur batez eskuragarri jarri zituen bereak (Aranberri, 2016, 198 or.). Horrez gain, TAUSek gidalerro orokorrak eskaintzen ditu, gidalerro espezifikoagoak sortzeko oinarri gisa (ikus TAUS, 2016). Euskarari dagokionez, HAEEn IZO taldeak ere eskuragarri jarri ditu euren IArengatik inguruko argibide batzuk (“Itzultzaile juridiko-administratiboa - IZO – IVAP”, 2020).

### **3.3. Postedizioaren eraginkortasuna**

PEa eraginkorra izan dadin, ezinbestean giza itzulpen baterako baino denbora gutxiago inbertitu behar da, PEa erabiltzeko arrazoi nagusia itzulpen-denboraren optimizazioa delako.

Kalitate oneko PEa egiteko behar den denbora bi faktorek baldintzatzen dute: itzultzaile automatikoaren emaitzaren kalitateak; eta posteditorearen trebetasunak.

IArengatik emaitzaren kalitatea hainbat faktoreren mende dago, hala nola jatorrizko testuaren kalitate eta ezaugarriak, itzultzaile automatikoaren nolakotasuna, hizkuntza bikotea, etab. Faktore asko dira eta ondorioz ezin da bermatu emaitza beti PErako egokia denik. Aranberrik adierazten duen bezala, aurreitzulpenaren kalitatea oso txarra baldin bada, gerta daiteke denbora gutxiago behar izatea testua errotik itzultzeko posteditatzeko baino (2016, 198 or.).

Horregatik, hain zuzen ere, posteditorearen trebetasuna oso garrantzitsua da: hasteko, segundo gutxiren buruan ebaluatu beharko duelako ea IAren emaitza posteditagarria den edo ez; eta, bigarrenik, gai izan behar delako PE prozesua ere oso arin egiteko eta soilik ezarritako helburua lortzeko behar-beharrezkoak diren egokitzapenak egiteko (Aranberri, 2016, 198 or.).

### **3.4. Posteditorearen profila**

Lanbide berria denez gero, eztabaida handia dago posteditorearen profilaren inguruan, PEak zuzenketa eta itzulpen lanekin antzekotasunak dituen arren, trebetasun desberdinak eskatzen dituelako.

Itzulpenak ez bezala, PEak aurreitzulitako testu bat du oinarri, eta aurreikusi daitezkeen oztopo edo akatsak bilatu eta moldatu behar dira (Vasconcellos, 1987 in Azou, 2016, 23-24 orr.). Hortaz, zuzentzaile eta posteditoreen lanak bilatu beharreko akatsetan desberdintzen dira: zuzentzaileek giza akatsak bilatzen dituzte (errepikapenak, akats ortotipografikoak, itzuli gabeko segmentuak, etab.), eta posteditoreek, aldiz, IAren akatsak bilatu behar dituzte (McElhaney & Vasconcellos, 1988 in Azou, 2016, 23-24 orr.).

Rico eta Torrejonek adierazten dute PEa itzultzaileen lan berrietako bat dela (2012, 167 or.), eta Kring eta Cobyren (2001) iritzi berekoak dira. Azken bi autore horien ustetan, soilik itzultzaileak du IAren akatsak hautemateko eta emaitzaren kalitatea ebaluatzeko gaitasuna, itzultzaileak oso ondo ezagutzen baititu hala kontzeptuen transferentzia prozesua, nola kontsultarako eskuragarri dauden baliabideen erabilera (Kring eta Koby, 2001 in Rico eta Torrejón, 2012, 167 or.).

Offersgardek, bestalde, adierazten du PERako gaitasun espezifikoak ere behar direla (Offersgard, 2008, 153 or. in Aranberri, 2016, 198 or.), posteditoreak abiadura handiz hartu beharko dituelako erabakiak (Aranberri, 2016, 198 or.). De Almeidak eta O'Brienek adierazten dute posteditore batek ezinbestean 3 trebetasun izan behar dituela: 1) IAren akatsak arin identifikatzeko eta zuzentzeko gaitasuna; 2) ataza abiadura handiz burutzeko gaitasuna; 3) gidalerroak jarraitzeko gaitasuna (De Almeida eta O'Brien, 2010, 2 or. in Aranberri, 2016, 198 or.).



#### **4. CROWD POSTEDITING ETA COMMUNITY POSTEDITING**

*Crowdsourcing* kontzeptua oso zabala da, eta barne hartzen ditu jende askoren kolaborazioaz baliatzen diren metodoak, hau da, ataza zehatzetan laguntzeko deialdi irekiak egiten dituzten metodoak (Geiger, Roseman, Fielt & Schader, 2012, 2 or.).

*Crowdsourcing* ereduak arrakasta izan du itzulpen munduan, eta plataforma askok erabiltzen dute kolaborazio sistema hau beraien edukia beste hizkuntzetan eskaintzeko, besteak beste TED fundazioak, Youtubek eta Wikipediak. Teknologia arloko enpresa handiek ere erabiltzen dute sistema hau beraien web-orrialde edo aplikazioak lokalizatzeko, esate baterako Facebookek 2008an erabili zuen eta Twitterrek 2009az geroztik hainbatetan erabili izan du aplikazioa eta bere eguneratzeak itzultzeko. Normalean horrelako proiektuetan erabiltzaileek jarraibideak jasotzen dituzte, softwarearen erabileraren nahiz espero den emaitzaren inguruan.

Orain doako itzultzaile automatikoak edonoren eskura daude, eta, beraz, ez litzateke harrizkoa izango aurrerantzean itzulpen kolaboratibo horretako prozesuan PEa egitea. Gainera, pentsa daiteke kasu honetan PEak itzulpenak baino potentzial handiagoa izan dezakeela (Tatsumi et al., 2012). Izan ere, askotan aritzen direnak itzultzaile ez-profesionalak dira, eta, hortaz, errazagoa iruditu ahal zaie posteditatzea errotik itzultzea baino; nahiz eta, IAren kalitatearen arabera, kontrakoa gerta daitekeen (Tatsumi et al., 2012).

##### **4.1. Profil berria: kalitatearen auzia eta ataza desberdinak**

*Crowdsourcing* plataformetan kolaboraziorako deialdiak normalean Internet bitartez egiten dira. Erabiltzaileak anonimoak dira eta, horren ondorioz, sortzen den kolaboratzaile multzoa heterogeneoa da, adituak nahiz ez-adituak bil baititzake (Schenk eta Guittard, 2011). Itzulpenaren eta baita PEaren kasuan ere, horrek esan nahi du itzultzaile eta posteditoreen artean profesionalak egon badaitezke ere, edonor izan litekeela, esaterako bi hizkuntza dakizkiten hiztunak. Horregatik, PE mota hau berezia da, eta arlo profesionaletik nabarmen aldentzen da: kasu honetan, PE abiadurak ez du hainbesteko garrantziarik, eta IAren erabileraren arrazoi nagusia jende ez-adituari itzulpen oinarri bat eskaintzea da, prozesua errazteko.

Askotan itzulpen horiek borondatez egiten dira, dirua irabazteko asmorik gabe. Halakoak dira esate baterako *Fansubak* edota guk landuko ditugun Wikipediako artikuluen itzulpenak. Kasu horietan, gainera, ez dago inolako hizkuntza mailako eskaerarik eta edonork parte har dezake. Hortaz, hizkuntza maila aldetik ere itzultzaile multzo heterogeneoa dugu (Mitchell, 2015, 37 or.).

IA edonoren eskura dagoen heinean, komunitate itzulpenean erabiltzaile elebakarrek ere posteditatu dezakete, eta, beraz, hori ere aintzat hartzekoa da gidalerro bat sortzerako orduan. Guk aztertutako hizkuntza norabidean (es-eu), dena dela, PE elebakarrak ez du garrantzia handirik. Izan ere, euskal hiztun gehienak, guztiak ez esatearren, elebidunak dira. Horregatik, lan honetan PE elebakarra alde batera utziko da.

Nolanahi ere, posteditore ez-profesional hauek, itzulpenean prestakuntzarik ez duten arren, testuaren gaiaren inguruko jakintza izan dezakete, eta hori lagungarria edota ezinbestekoa izan daiteke kalitate oneko PE bat lortzeko (Mitchell, 2015, 5 or.). Hortaz, banako bakoitzak gaiaren inguruan duen ezagutzak garrantzia handia izango du komunitate PEan.

Prozesu honetan itzultzaile eta posteditore ez-profesionalek parte hartzen dutenez, zalantzarria izan daiteke lortu daitekeen kalitate maila. Dena den, Mitchellen (2015) aurkikuntzen arabera, badirudi kalitate maila ona lortu daitekeela. Mitchellek Norton enpresako foro bateko PEa aztertu zuen, ingeles-alemana hizkuntza konbinazioan. Bere tesian foroan posteditatzeko gidalerro minimo batzuk sortu zituen, foroko erabiltzaile batzuekin testatu zituen eta ateratako ondorioetan oinarrituta gidalerro sakonago batzuk osatu zituen. Egindako frogan ikusi zuen partehartzaileek batez beste akatsen % 74 zuzendu zituztela, eta, beraz, itzultzaile ez profesionalekin ere PEa erabilgarria izan zitekeela.

## **5. POSTEDIZIOA EUSKAL HERRIAN**

Gure lana testuinguruan kokatze aldera, ezinbestekoa da PEak Euskal Herrian izan duen garapenaz hitz egitea. Horretarako, lehenik eta behin IAren egoera aztertu behar dugu, itzultzaile automatikoaren kalitate gordinak baldintzatzen baitu PEa. Hori atal honen lehengo zatian jorratuko dugu. Ostean, itzultzaileei eskaini zaien PE-formakuntzari buruz hitz egingo dugu; hirugarrenez, PEaren erabilera profesionalari buruz; eta azkenik, guri dagokigun Wikipediaren PEaz.

### **5.1. Itzultzaile automatikoen eboluzioa**

Euskal IAren historia, erdara handiekin alderatuta askoz berriagoa da. Gure kulturara aurrerapenak beranduago iritsi dira, alde batetik, XX. mendean zehar IA hizkuntza handietan zentratu zelako, eta, bestetik, euskarak inguruko hizkuntzekin desberdintasun gramatikal nabarmenak izateaz gain, diglosia egoeran dagoen hizkuntza txikia delako.

Lehen euskarazko itzultzaile automatiko publikoa 2007koa da, Matxin izenekoa. Elhuyarrek eta IXA taldeak garatu zuten, elkarlanean. Erregelen bidezko sistema batean oinarritzen zen, hau da, hiztegi elebidun bat eta arau gramatikalen multzo bat erabiltzen zituen itzultzeko (Aranberri, 2016, 196 or.). Itzultzaile horrek bere azken eguneratzea 2013an jaso zuen (Matxin 2.0), eta egun ez dago eskuragarri itzultzailea zaharkituta gelditu eta baztertu dutelako.

2010ean Google Translaten euskarazko bertsioa aurkeztu zuten. Honakoa itzultzaile estatistiko bat zen. Mota honetako sistemek corpus elebatar batean eta corpus elebidun batean oinarrituta ikasten dute, eta probabilitatean oinarritutako algoritmoak erabiltzen dituzte itzultzeko (Aranberri eta Labaka, 2017, 220 or.). Egun, badirudi, baina ezin dugu konfirmatu benetan hala denik, Google Translatek sare neuronalak erabiltzen dituela. Sistema hori aurrerago azalduko dugu.

2012an Eusko Jaurlaritzak erregeletan oinarritutako beste itzultzaile bat aurkeztu zuen, Lucy Software enpresak garatua. Guk gure PE proba egin genuenean oraindik eskuragarri zegoen, baina gutxira baztertu egin zuten.

Lehen etapa honetan, behatu daitekeenez, erregeletan oinarritutako sistemak garatu ziren batez ere, nahiz eta hizkuntza nagusietan itzultzaile estatistikoak hobesten

ziren (Irurzun, 2018, 13 or.). Euskaraz kalitate oneko itzultzaile estatistikoak garatzeko zailtasun handiak zeuden. Itzultzaile estatistikoek, entrenamendu-datu edo corpus paralelo handi baten faltan, itzulpen literalak egiteko joera dute, eta horrek antzeko hizkuntzetan funtziona dezake, baina euskararen kasuan ez (Aranberri eta Labaka, 2017, 221 or.). Horregatik, ezinbestean corpus handi bat sortu behar zen euskararako eta hori nahiko zaila zen. Gainera, egitura gramatikalen desberdintasunagatik, emaitza hobekak lortzen ziren erregelaren bidezko itzultzaileekin (Aranberri eta Labaka, 2017, 221 or.).

2018an MODELA aurkeztu zuten Elhuyarrek eta Mondragon Linguak. Euskarazko lehen itzultzaile neuronalak izan zen. Itzultzaile hauek, estatistikoek bezala, corpus elebidunak dituzte oinarri, baina azken horiek ez bezala Sare Neuronal Artifizialak erabiltzen dituzte, zeinak garunaren funtzionamenduan inspiratuta dauden. Prozesua bi etapatan gertatzen da: lehenengoan, jatorri testua zenbaki-bektore izatera pasatzen da, eta ostean Sare Neuronalen bitartez zenbaki-bektore horiek xede testuko errepresentazio bihurtzen dira. Itzultzaile Neuronalen abantaila da emaitza naturalak lortzen dituztela eta erabileraz ikasten eta hobetzen doazela (Aranberri, 2017, 222 or.).

Egun MODELA bazterturik gelditu da, eta bere ordean Euskal Herrian garatutako beste 3 itzultzaile automatiko neuronal publiko ditugu eskuragarri: 1) Elia, Elhuyarrek sortua; 2) Itzul, Eusko Jaurlaritzaren ekimenez sortua; eta 3) Batua.eus, Vicomtech enpresak garatua. Hirurak oso segidan aurkeztu ziren 2019an. Interesgarria da zehaztea guk erabiliko dugun itzultzaile automatikoa, Elia, aurretik Itzultzailea.eus deitzen zela. Izena 2021eko martxoaren 22ko eguneratzearekin aldatu zen.

Egun, beraz, 4 itzultzaile automatiko neuronal daude eskuragarri euskaraz: Itzul, Elia, Batua.eus eta Google Translate.

## **5.2. Formakuntza**

Euskarazko PEaren formakuntzari begira, 2015ean UEUK 47 orduko ikastaroa antolatu zuen, non IAren eta PEaren alderdi teoriko eta praktikoa landu ziren. Helburu nagusia zen profesionalek IArekiko zuten mesfidantza uxatzea eta barnetik aztertzea ea baliagarria iruditzen zitzaion. Horretarako, ikastaro honek itzultzaileak gonbidatzen zituen «itzulpen automatikoa erabiltzera, post-edizioa taxuz probatzera, egin beharreko aldaketak finkatzen hastera, produktibitatea neurtzera, norberaren itzulpenen kalitatea

analizatzaera, eta, orokorrean, teknologia honen egoeraz eta baliagarritasunaz hausnartzera» (UEU, 2015). Ikastaro honetan irakasle aritu ziren Nora Aranberri, Karlos del Olmo eta Ane Lopez.

Itzultzaile automatiko neuronalen sorkuntzarekin badirudi, PEarekiko eta IArekiko interesa handitu dela. 2020. urtean, EIZIEk itzulpen automatikoaren inguruko jardunaldi bat egin zuen, baliabide honen hobekuntza eta aplikazioen berri eman eta horren inguruan hausnartzeko. Gainera, urte berean 8 orduko lantegi bat antolatu zuen, non 2015eko aurrekarian bezala, alderdi teoriko eta praktikoak landu ziren. Lantegi honetan irakasle aritu ziren Nora Aranberri eta Uxo Iñurrieta.

Aurten, gainera, UPV/EHUK lehen aldiz PEaren inguruko bi ekitaldi antolatu ditu: alde batetik, tailer bat Itzulpengintza eta Interpretazioko graduako ikasleentzat; eta, bestetik, udako ikastaro bat.

PE tailerra itzulpengintzako irakasleek antolatu eta Nora Aranberrik eman zuen. Tailerrean IARI buruz sarrera orokor bat egin zen eta parte-hartzaileek hainbat PE ariketa egin zituzten, lortutako emaitzak aztertu eta horien inguruan hausnartzeko.

Udako ikastaroa, berriz, itzultzaile profesionaleri zuzendua dago. Zati bat alderdi praktikoari eskainiko zaio, baina bestetik ere itzulpengintzan egon den eraldaketaren inguruan hausnartuko da. Ikastaroan IAren erabilera barnetik aztertuko da: esparru ezberdinetako profesionalen ikuspegia ezagutzera eman eta itzulpengintzan erreferente diren erakundeen lan egiteko modua ezagutzeko aukera egongo da. Zentzu horretan ikastaro berritzailea da, orain arte IAren erabilera praktikoari buruzko informazioa oso eskasa izan delako. Ikastaro horretan irakasle arituko dira: Paula Alvarez, Jon Arbizu, Ander Irizar, Ixiar Iza, Saroi Jauregi, Elizabete Manterola, Celia Rico eta Idoia Santamaría.

### **5.3. Postedizioaren erabilera**

Duela gutxira arte IA lan-fluxuan barneratzea zaila zen, oraindik ere ez zirelako emaitza gordin onak lortzen, erdara askotan gertatzen ez zen bezala. Dena den, itzultzaile automatiko neuronalen garapena aurrerapauso handia izan da euskararen kasuan, itzulpen hobeak eta naturalagoak sortzea lortu delako. Horri esker, PEa aukera erakargarria izan daiteke euskal bezero eta enpresentzako, denbora berdinean gehiago itzuli daitekeelako.

Hortaz, espero daiteke PEa gero eta gehiago zabaltzea Euskal Herriko itzulpengintza merkatuan.

Momentuz ezin dugu datu zehatzik eman erabileraren inguruan, oraindik ez baita ia aztertu, baina bai baieztatu dezakegu IA gero eta gehiago erabiltzen dela itzulpen merkatuan, eta eraldaketa prozesu bat gertatzen ari dela. Besteak beste, badakigu Elhuyarren PEa itzulpen zerbitzuetako lan fluxuan barneratu dutela, eta baita HAEEko Itzulpen Zerbitzu Ofizialean ere.

Euskararen Eragin Ekonomikoaren txostenak (2019) ere eraldaketa prozesu horren berri ematen du. Txosten honetan euskararekin lotutako Gipuzkoako jarduera ekonomikoak aztertzen dira, itzulpengintza barne. Hizkuntza arloko hainbat enpresa eta langileri inkesta bat egin zieten, eta hortik atera zuten ondorioetako bat honakoa da: «itzulpen automatikoarekin lotutako teknologia berrien zabalpena da erronka nagusia, eta horien aurrean, kalitatezko produktua eskaintzea eta tarifak jaisteko presioari aurre egitea» (Siadeco, 2019, 82 or.). Adierazten da ere itzultzaileek tresna lagungarritzat jotzen dutela IA, baina azpimarratzen dutela PEaren zabaltzeak lan-baldintzak okertzea dakarrela: itzulpenarekin alderatuta gutxiago kobratzen dute; eta orokorrean PEan kostua eta epeak lehenesten dira, itzulpen-kalitatearen ordeztuz (Siadeco, 2019, 82 or.).

Dena den, IAren erabilera ez da esparru profesionalera mugatzen. Aurretik esan den bezala, euskal merkatuan 4 itzultzaile automatiko daude eskuragarri dohainik eta publikoki. Espero daiteke, beraz, jendeak IA arlo profesionalarekin loturarik ez duten jarduera eta esparruetan txertatuko duela, eta etorkizunean euskaldun guztiok izango garela neurri batean «posteditore». Guk lan honetan esparru horietako bat aztertuko dugu, non dagoeneko PEa erabiltzen hasia den: Wikipedia.

#### **5.4. Wikipedia: IA eta posteditioa**

Wikipedia Internet bidezko doako entziklopedia aske eta eleanitza da. Irabazi asmorik gabeko Wikimedia Fundazioak sustengatzen du eta *crowdsourcing* plataforma bezala funtzionatzen du, hau da, wikilari deritzen borondatezko kolaboratzaileen lankidetzaren bitartez elikatzen da. Proiektuaren xedea da ahalik eta artikulua gehien eskuragarri izatea ahalik eta hizkuntza gehienetan. Egun Wikipedia 321 hizkuntzatan dago eskuragarri, eta horietako bat euskara da.

Eleaniztasun hori tartean dagoela, Wikipediaren hedapenean itzulpenak garrantzia handia du: kolaboratzaileek maiz, edukia zerotik sortu beharrean, beste hizkuntzetako Wikipediatako artikulua euren hizkuntzetara itzultzen dituzte. Euskararen kasuan, beste hizkuntza txikiagoen kasuan bezala, prozesu hau are garrantzitsuagoa da bi arazoirengatik: lehenik eta behin, inguruko erdaretan baino informazio gutxiago dagoelako eskuragarri; eta, bigarrenik, wikilari kopurua askoz txikiagoa delako. Hortaz, euskaraz eduki berria zerotik sortzeko zailtasunak handiagoak dira.

Itzulpenak hainbesteko pisua duenez Wikipediaren hedapenean, bi tresna sortu dira prozesua errazteko: Content Translation Tool eta Translate gehigarria. Lehenengoak, aukera ematen die wikilariei jatorrizko testua eta sortzen ari diren artikulua alboz albo ikusteko, eta hainbat prozesu automatizatzeko (“Content translation Edukien itzulpena”, d.g). Gainera, hainbat hizkuntzatan hiztegi eta IAko zerbitzuak barneratzen ditu. Bestetik, Translate aurreko tresnaren gehigarria da, eta hari esker beste funtzionalitate batzuk gehitzen zaizkio, besteak beste itzulpen memoriak eta itzultzaile automatikoak erabiltzeko aukera (“Translate”, d.g) . Euskararen kasuan badakigu lehen Yandex IA erabiltzen zutela, gero Matxin eta egun Google API. Etorkizunera begira Elia erabiltzea da asmoa.

Bestetik, ikerketa ekimenak ere sortu izan dira Euskal Wikipediaren inguruan. 2009 eta 2013 urteen artean, IXA eta Euskal Wikilariak elkarlanean aritu ziren bi helbururekin: Matxin sistema hobetzea eta Euskal Wikipedian artikulua berriak sortzea. Itzulpenak egiteko bi teknika erabili ziren: 1) erregeletan oinarritutako itzultzaile automatiko bat eta postedizio estatistikoa. Emaitza horien gainean Omega T erabiliz euskarzko testuak eskuz orraztu ziren, eta egindako zuzenketetan oinarrituta Matxini hobekuntzak egin zitzaizkion. Orraztutako 100 artikulua ostean Euskal Wikipedian argitaratu ziren (“Wikiproiektu:OpenMT2 eta Euskal Wikipedia”, d.g).

Hala ere, momentuz ez da jarraibiderik, ezta gidalerrorik sortu Wikipedian posteditatzeko, nahiz eta jakina den oso lagungarriak direla. Aipatzen diren jarraibide bakarrak programaren erabileraren ingurukoak dira, eta PEari dagozkion gomendio eskasek soilik adierazten dute IAren emaitza wikilariak orraztu beharko duela, eta formatuarekin kontuz ibili behar dela.

Jarraibide falta ez da soilik euskararen arazoa, baizik eta Wikipediaren arazo orokorra, hizkuntza handiei ere eragiten diena. Arazo horretaz jabetuta, gu lan honetan

Wikipedian posteditatzeko gidalerro batzuk sortzen saiatuko gara, erabiltzaileek abiapuntu lagungarri bat izan dezaten. Gure kasuan, gaztelaniatik euskararako norabidea landuko dugu.



## **6. EUSKAL WIKIPEDIAKO POSTEDIZIOA AZTERTZEN**

GRAL honen helburu nagusia Wikipedian gaztelaniatik euskarara posteditatzeko gidalerro batzuk sortzea da.

Helburu hori lortzeko gaztelaniatik euskararako PE proba bat diseinatu dugu: lehenik, gidalerro minimo batzuk sortu ditugu; bigarrenik, gidalerro horiek, 7 wikilariren laguntzaz, PE proba baten bidez testatu ditugu; eta, azkenik, proba horretatik ateratako emaitzak aztertu eta horien arabera gidalerro minimoei hobekuntzak egin dizkiegu. PEko emaitzak ebaluatzeko akatsen anotazioa erabili dugu, hau da, akatsak identifikatu eta sailkatu ditugu.

Bestetik, partehartzaileek galdetegi bat bete dute, euren profilaren berri izateko, eta aztertzeko ea hiru faktore hauek eraginik izan dezaketen PE emaitzan: 1) euskararen ezagutza maila; 2) itzulpen esperientzia; 3) PE esperientzia.

Atal honetan, azalduko dugu lehenik eta behin zertan oinarritu garen probaren diseinurako. Gerora, aukeratutako metodologia eta probaren diseinua sakonean aztertuko dugu, eta ondorengo galderak argitzen saiatuko gara: 1) zergatik aukeratu dugun Elia itzultzailea eta ez beste bat; 2) nola aukeratu diren erabilitako testuak; 3) nola osatu den partehartzaileen lagina; 4) nola osatu diren probarako behin-behineko gidalerroak; eta 5) nola ebaluatu diren postedizio emaitzak. Horrez gain, proba egin aurretik egindako egiaztapen probaren eta bertatik lortutako emaitzen berri emango da.

### **6.1. Oinarri teorikoa**

Proba hau diseinatzeko, lehenik ereduak bilatzen hasi ginen, ikusteko ea aurretik antzekorik egin zen. Eredu horien artean Mitchell-en doktoretza tesia aurkitu genuen (2015) eta hori da guk oinarri gisa erabili izan duguna gure metodologia garatzeko, guk egin nahi genuen lanarekin antzekotasun gehien zituen ikerlana zelako.

Michellek bere tesian lehenengo aldiz komunitate postediziorako gidalerro batzuk sortu zituen, zehazki Norton enpresako produktuen erabiltzaile-foroko ekarpenak ingelesetik alemanera itzultzeko. Helburu hori lortzeko, PE proba bat sortu zuen, guk hemen planteatu dugunaren antzekoa.

Michellek, hasteko, egiaztapen proba bat egin zuen, eta konturatu zen aurre-edizio fase bat beharrezkoa zela. Ostean, benetako proba egin zuen. 18 partaidek parte hartu zuten proba horretan, bik izan ezik denek alemana zuten ama hizkuntza. Lehenik eta behin, galdetegi bat bete zuten, non informazio demografikoa jaso zen eta IAz eta PEaz zuten iritziaren inguruan galdetu zitzairen. Gero, partaide bakoitzak 12 testu posteditatu zituen, Michellek sortutako gidalerroen laguntzaz. PEko emaitzen ebaluazio sistemarako, Michellek hiru ebaluazio sistema erabili zituen eta haien arteko emaitzak alderatu zituen: 1) akats anotazioa; 2) eremuko-espezialistaren ebaluazioa; eta 3) komunitateko kideek egindako ebaluazioa. Akats anotazioari dagokionez, aipatzekoa da IAren emaitza eta PEaren emaitza anotatu eta biak alderatu zituela.

Michellek proposatutako gidalerroak sortzeko metodologia guretzako oso interesgarria zen hainbat arrazoiengatik. Alde batetik, guk gai berdina jorratuko genuen: komunitate PE, non posteditoreek ez duten zertan itzultzaile profesionalak izan. Bestetik, ikuspuntu linguistiko batetik ere interesgarria zen. Michellek hitz ordena eta joskera desberdina duten bi hizkuntzen arteko PEa ikertu zuen, ingelesa eta alemanaren artekoa alegia, eta hain justu guk proposatzen genuen hizkuntza konbinazioan ere arazo berdina genuen.

Dena den, bi proiektuen artean alde nabarmenak zeuden. Lehenik, Michellek itzultzaile estatistiko bat erabili zuen, eta guk, aldiz, neuronal bat erabili dugu, beraz, posible da gure aurreitzulpenen kalitatea hobea izatea. Bigarrenik, bere tesian ez bezala, guk ez genuen PE elebakarra aztertzeko beharrik. Auresuposatu genuen euskarazko wikilari gehienak Hego Euskal Herrikoak izango zirela eta, beraz, jatorrizko hizkuntza, gaztelera alegia, bazekitela maila altuan, askorentzat ama hizkuntza delako eta bestela ere gaztelerekin izaten den harremana handia delako. Hirugarrenik, Mitchellen esperimentuan, PEaren emaitzak jatorrizko hizkuntzaren, hots, ingelesaren ezagutzaren mende zeuden; gure kasuan, aldiz, xede hizkuntzaren, hau da, euskararen mende zeuden. Kontuan hartu behar da euskara, alemana ez bezala, hizkuntza txikia dela, diglosia egoeran dagoena, eta hori ere oztopo bat dela hiztunentzako.

Bestalde, lan honen luzeraren mugak direla eta, proba asko sinplifikatu behar izan dugu Michellenarekin alderatuta. Proba hiru fasetara mugatu dugu: 1) partehartzaileen profila osatzeko galdetegi bat egitea; 2) partehartzaileek testu bat posteditatzea guk

sortutako gidalerroen laguntzaz; eta 3) aurreitzulpenaren eta postedizioen akats anotazioa egitea eta biak alderatzea.

## **6.2. Itzultzaile-automatikoaren aukeraketa**

Proiektuaren nondik norakoak diseinatzen hasi aurretik, otsailaren erdialdean harremanetan jarri ginen bai Wikipediako bi kolaboratzaileekin, Aiora Mujikarekin eta Galder Gonzalezekin, eta baita Elhuyarreko Itziar Cortésekin.

Haien eskutik jakin genuen elkarlanean zebiltzala Elia (Elhuyarren Itzultzaile Automatikoa) Wikipediako testu editorean txertatzeko. Hori dela eta, erabaki genuen beraiek hasitako ibilbideari jarraitzea eta Elia erabiltzea lan-tresna gisa, sortutako gidalerroak gerora erabilgarritasun erreal bat izan zezaten.

Testu mota zehatz bat posteditatzeko gidalerroak sortzerako orduan, normalean hainbat itzultzaileekin egiten dira itzulpen probak, eta ostera itzulpen gordin horien kalitatea alderatzen da. Azterketa horretatik ateratzen diren emaitzetatik kalitate hoberena lortzen duen itzultzailea aukeratzen da, aurretik esan den bezala PEaren eraginkortasuna lehen itzulpen gordin horren menpekoa delako. Gure kasuan, aldiz, dagoeneko bagenekienez Elia izango zela Wikipedian erabiliko zen tresna, urrats hori omititu genuen eta ez genuen Eliaren itzulpen kalitatea beste itzultzaile batzuenarekin alderatu.

Bestetik, aipatu behar dugu, gure proban Wikipediako testu-editoretik kanpo lan egin dugula, nahiz eta etorkizunean hori izango den wikilarien lan tresna. Hasiera batean gure asmoa zen bertan egitea proba, betiere Elia Wikipedian integratzen bazen guk proba hasi aurretik. Azken finean, askoz egokiagoa da proba erabiltzaileen ohiko lan-tresnan egitea, egoera errealagoa delako eta sortu daitezkeen erabilera arazoak hauteman daitezkeelako. Dena den hori ez zen posible izan.

Proba diseinatzen hasi ginenean, Elhuyar eta Wikipediako kolaboratzaileek adierazi ziguten Wikipediako testu-editorean oraindik ezin zela Elia erabili, eta ez zekitela noizko izango zuten prest. Une horretan apur batez itxarotea erabaki genuen. Hala ere, martxoaren hasieran jakin genuen oraindik ere gai legalak ebatzi behar zituztela itzultzailea barneratzeko, eta, ondorioz, testu-editoretik kanpo lan egitea erabaki genuen.

### 6.3. Testu laginen aukeraketa

Testu laginak aukeratzeko orduan 5 faktore hartu genituen kontuan:

- 1) Testuak Wikipediako artikuluen erakusgarri izatea.
- 2) Euskarazko artikuluen kalitate maila.
- 3) Testuak euskaratu gabe egotea.
- 4) Testuen espezializazio maila.
- 5) Laginen luzera.

Lehenik eta behin, ezinbestekoa zen aukeratutako testuak Wikipediako artikuluen erakusgarri izatea, eta, hortaz, aukerarik onena testu errealak aukeratzea zen, hau da, gaztelerazko Wikipediako artikuluetatik ateratako testuak. Azpimarratzekoa da euskaratu gabeko atalak izan direla aukeratutako guztiak. Halaber, aintzat hartu da euskarazko artikuluen kalitatea: ORES kalitate neurgailuan puntuazio baxuagoa zuten artikuluen hobetsi dira, horiek baitira euskaraz eduki falta duten artikuluen.

Bestetik, erabaki genuen aukeratutako testuen gaiak ere jakintza orokorrekoak izan behar zirela, lagineko partehartzaileak jakintza alor desberdinekoak izan zitezkeelako. Testu oso espezializatu bat aukeratzekotan, gaiaren inguruan zuten jakintzak PEaren kalitatea baldintza zezakeen, eta, kasu horretan emaitzak ez ziren guretzako erabilgarriak izango. Gainera, auresuposatu genuen wikilari gehienek testu oso espezializatuak jorratzekotan, euren alorretik gertukoak zirenak aukeratuko zituzketela.

Dena den, aldi berean bermatu behar genuen testuak ez zirela orokorrekiak, eta terminologia ere aztertu zitekeela. Kontuan hartu behar da Wikipediako artikuluen gehienek terminologia nahiko izan ohi dutela eta, beraz, hau aztergai izatea ezinbestekoa zen.

Azkenean, gaiari dagokionez, erabaki genuen egunerokotasun handiko gai sozialak jorratzen zituzten testuak aukeratzea, nolabait eskuragarriak, ulergarriak eta interesgarriak izan zitezkeelako edonorentzat, baina aldi berean izen berezi eta terminologiaren agerpena ere bermatzen zutelako.

Testuen luze-laburra ere zehaztu behar zen, proba gehiegi ez luzatzearen. Hori dela eta, erabaki genuen artikuluen azpiatalen bat aukeratzea lagin gisa, artikuluen oso bat

baino laburragoa delako. Azpiatal bat aukeratzeak, gainera, badu beste abantaila bat: ideia edo azpiadar bat jorratzen duten heinean, neurri batean testu independente bezala funtziona dezaketela, zentzu osoa dutelako. Azpiatalek ere luzera desberdina izan zezaketenez, 200-300 hitz arteko muga ezarri genuen testuak hautatzerakoan. Erabilitako bi testuak eta baita haien itzulpen automatikoa ere lan honen eranskinean jasota daude.

#### 6.4. Pertsona-laginarekin diseinua

Proba egingo zuten pertsona lagina sortzeko, erabaki genuen 8 pertsonarekin egingo genuela proba, launa testu bakoitza posteditatzeko. Pentsatu genuen tamaina horretako lagina egokia izan zitekeela dimentsio hauetako lan baterako, gidalerroak doitzeko adina informazio eskura genezakeelako, baina gehiegi luzatu gabe.

Gidalerroa wikilariarentzako egina izango zenez, erabaki genuen horiek izan behar zutela lehenetsi beharreko hautagaiak. Galder Gonzalezen laguntzaz, Euskal Wikilarien Kultur Elkarteko sortzaileetako bat, deialdi bat egin genuen wikilarien artean. Wikipediako Txokoa atalean mezu bat publikatu genuen eta Galderrek zabaltzen lagundu zigun.

*Irudia 1 Wikipedian argitaratutako deialdia.*

### Wikipedian posteditatzeko gidalerroak

Iruzkina bat • duela 20 egun

**Amaia Solaun**

Kaixo!

Amaia Solaun naiz, Itzulpengintza eta Interpretazioko 4. mailako ikaslea. Aurten grala egiten ari naiz posteditazioaren inguruan, hau da, itzulpen automatiko bat zuzentzearen eta editatzearen inguruan.

Galderako Wikipedian posteditatzeko gidalerro batzuk sortzen ari naiz. Asmoa da, itzulpen automatikoaz baliatuta, Wikipediako gaztelarazko testuak euskaratzeko prozesua erraztea.

Dagoeneko oinarrizko gidalerro batzuk sortu ditugu, baina horiek hobetzeko eta gidalerro zehatzagoak sortzeko, zuen laguntza beharko genuke. 8 bolondres beharko genituzke posteditazio froga txiki bat egiteko. Asmoa litzateke, posible izatekotan, datorren astean bertan egitea, hilaren 14-18 artean.

Partehartzaileei korreoz 3 dokumentu bidaliko genizkieke: jatorrizko gaztelarazko testua, itzultzaile automatikoaren itzulpena eta sortu ditugun oinarrizko gidalerroak. Ostean, gidalerroen gomendioei jarraituta, itzulpena posteditatu beharko dute, hau da, zuzenketak egin.

Probarako gehienez ere ordubete kalkulatu dugu. Orduetgia malgua izango da, partehartzaileek ahal edo nahi duten unean egin ahalko dute proba.

Interesik baduzue edo galderarik baduzue jarri zaitezte harremanetan nirekin: [asolaun005@ikasle.ehu.eus](mailto:asolaun005@ikasle.ehu.eus)

Zuen laguntza biziki eskertuko genizueke!

Amaia Solaun

Deialdia arrakastatsua izan zen eta aste eta erdiren buruan gure helburua bete genuen: 8 pertsonak idatzi ziguten esanez prest zeudela parte hartzeko. Hala ere, proba egiterakoan batek huts egin zigun, eta azkenean 7 izan ziren partehartzaileak.

Partaide bakoitzari emailez galdetegia eta beste hiru dokumentu bidali zitzaizkion: gidalerroak, jatorrizko bertsioa eta IAren emaitza. Ostera, eskatu zitzaien bueltan galdetegia eta posteditatutako dokumentua bidaltzea, eta baita adierazteko zenbat denbora behar izan zuten PEa egiteko.

### 6.5. Gidalerro minimoen sorkuntza

Kontuan hartuta itzultzaileak ez zutela zertan profesionalak izan, erabaki genuen PE-arina ezartzea gidalerroen helburu gisa. Horretarako, TAUS-en (2016) PE arinerako gidalerroak erabili genituen oinarri gisa eta baita Mitchellek (2015) tesian proposatutako oinarrizko gidalerroak ere.

*Taula 1 TAUS eta Mitchellen gidalerroak.*

TAUS (2016)	MITCHELL (2015)
<ul style="list-style-type: none"> <li>• Se intenta obtener una traducción correcta desde el punto de vista semántico.</li> <li>• Se asegura que no se haya añadido ni omitido ninguna información accidentalmente.</li> <li>• Se modifica todo contenido ofensivo, inadecuado o no aceptable culturalmente.</li> <li>• Se utiliza tanto como sea posible del resultado “en bruto” de la traducción automática.</li> <li>• Se aplican las reglas básicas con respecto a la ortografía.</li> </ul>	<ul style="list-style-type: none"> <li>• Edit the text to make it more fluent and clearer based on your interpretation.</li> <li>• Try to correct phrasal ordering and spelling, for example, if they make the text hard or impossible to understand.</li> <li>• Use words, phrases or punctuation as they are, if they are acceptable.</li> <li>• If you are working with reference to the original text, make sure that no information has been added or deleted.</li> </ul>

<ul style="list-style-type: none"> <li>• No hay que implementar correcciones de naturaleza estilística.</li> <li>• No hay que reestructurar oraciones solo para mejorar el flujo natural del texto.</li> </ul>	
--	--

Guk gure gidalerroetan bi gidalerro hauetan agertzen ez den beste gomendio bat barneratzea erabaki genuen (7. irizpidea). Pentsatu genuen PEaren inguruan ezagutza maila desberdinak egon ahal zirenez, komenigarria zela oinarritzko printzipio hau espezifikatzea: itzultzaile automatikoaren proposamena baztergarria baldin bada, hobe dela errotik itzultzea. Azkenean, honako gidalerro zerrenda hau osatu eta erabili izan dugu lehen probarako:

1. Bermatu ez dela informazioa gehitu edo ezabatu.
2. Bermatu itzulpenak jatorrizkoaren esanahia biltzen duela eta ez dagoela kontraesanik.
3. Berrordenatu esaldia beharrezkoa bada eta zuzendu gramatika akatsak. Testuak ulergarria izan behar du.
5. Ortografia akatsak zuzendu.
6. Ahal den neurrian, mantendu itzultzaile automatikoak proposatutako hitzak, esaldiaren ordena eta puntuazioa.
7. Itzulpen automatikoak sortutako esaldia ez bada batere erabilgarria, errotik itzuli.

## 6.6. Ebaluazio sistema

Itzultzaile automatiko baten kalitatea neurtzeko batez ere hiru metodo erabiltzen dira: ebaluazio automatikoa, giza ebaluazioa eta erabilgarritasun ebaluazioa (Mitchell, 2015, 42 or.). Proba honetarako bigarrena burutzea erabaki da, giza ebaluazioa, akats anotazioaren bitartez.

Akats anotazioa, hots, testu batean dauden akatsak hautematean eta sailkatzean datza. Sistema honek ahalbidetzen du itzultzaile automatiko batek sortutako emaitzaren

kalitatea neurtzea, eta baita emaitzak hobetzeko estrategiak diseinatzea ere (Comparin eta Mendes, 2017, 1 or.).

Akats anotazioak, gainera, ahalbidetzen du IA eta PEaren emaitza alderatzea (ikusirik Comparin eta Mendes, 2017; Mitchel, 2015), baldin eta bi testuak anotatzen badira. Metodo honekin jakin daiteke zenbat akats zuzentzen diren eta baita zein akats ez diren zuzendu (Comparin eta Mendes, 2017, 2 or.). Informazio hau oso egokia da gidalerro erabilgarri batzuk sortzerako orduan.

Akats anotazioan oinarritutako ebaluazio bat planteatzerako orduan, ezinbestean errubrika bat osatu behar da akatsak sailkatzeko. Gomendatzen da egindako aldaketa indibidual bakoitzari kategoria bat esleitzea. Kategoria eta azpikategoria kopurua aldakorra izan daiteke, gomendatzen baita errubrika kasuan kasuko beharretara egokitzea (Solokiev, 2017).

Ohikoa da ere, nahiz eta ezinbestekoa ez izan, aldaketa bakoitzari larritasun maila bat ematea. TAUSEk 4 erabiltzea gomendatzen du: kritikoa, handia, txikia eta neutrala (TAUS, d.g). «Kudos» izeneko beste kategoria bat ere proposatzen du, feedback positiboa emateko. Kudos kategoriak adierazten du posteditoreak oso lan ona egin duela eta hitzak oso ondo aukeratu dituela (Solokiev, 2017).

Taula 2 TAUS DQF-MQM Akats kategorien larritasun mailak.

#	Definition	Description
1	<b>Critical</b>	Errors that may carry health, safety, legal or financial implications, violate geopolitical usage guidelines, damage the organization's reputation, cause the application to crash or negatively modify/misrepresent the functionality of a product or service, or which could be seen as offensive.
2	<b>Major</b>	Errors that may confuse or mislead the user or hinder proper use of the product/service due to significant change in meaning or because errors appear in a visible or important part of the content.
3	<b>Minor</b>	Errors that don't lead to loss of meaning and wouldn't confuse or mislead the user but would be noticed, would decrease stylistic quality, fluency or clarity, or would make the content less appealing.
4	<b>Neutral</b>	Used to log additional information, problems or changes to be made that don't count as errors, e.g. they reflect a reviewer's choice or preferred style, they are repeated errors or instruction/glossary changes not yet implemented, a change to be made that the translator is not aware of.
5	<b>Kudos</b>	Used to praise for exceptional achievement.

Gure kasuan erabaki genuen Giselle de Almeidaren (2013, 95 or.) akats sailkapena errubrikarako oinarri gisa erabiltzea, baina euskarako egokitzapen gramatikalak eginez. DQF-MQM akats tipologia (TAUS, d.g) baino konpaktuagoa zen orokorrean, baina gramatika atalean TAUSEn errubrika baino zehatzagoa zen, eta hori interesgarria izan



zitekeen guretzat. Aurreuposatu genezakeen zuzendu beharreko akatsetako asko gramatikalak izango zirela, euskarak eta gaztelaniak arau gramatikal desberdinak dituztelako. Gainera, errubrika egokia da erabiltzaileek sortutako edukiaren PEa ebaluatzerako orduan (Mitchell, 2015, 94 or.).

Taula 3 De Almeida (2013) Postedizio aldaketen sailkapena.

**Table 3.6: Typology for classifying post-editing changes**

Master categories			
Essential changes	Preferential changes	Essential changes not implemented	Introduced errors
Subclassification for the four master categories			
Main categories		Subcategories (if applicable)	
Accuracy (completeness)		Extra information in MT output Information missing from MT output Untranslated text	
Consistency		N/A	
Country		Decimal points Quotation marks Currency symbol Date/time format	
Format		N/A	
Language		Adjectives Adverbs Capitalisation Conjunctions Determiners Gender Nouns Number Phrasal ordering Prepositions Pronouns Punctuation Spelling Verb tense	
Mistranslation		N/A	
Style		N/A	
Lexical Choice		N/A	

**Colour coding:**

Black - categories from the LISA QA Model  
 Blue - categories from the GALE PE Guidelines  
 Magenta - categories devised by the researcher

Akatsen larritasun mailari dagokionez, TAUSen MQM eta Comparin eta Mendesen larritasun sailkapenetik abiatuta (2017, 3 or.) gure eskema propioa sortu dugu. Alderik nabarmenena «kudos» bonifikazio sisteman dago; izan ere, erabaki da jatorrizkoaren edukia argitzen edo hobetzen laguntzen duten aldaketetarako erabiltzea.

Horrelako bonifikazio batek ez luke eremu profesional batean funtzionatuko, baina Wikipedia bezalako kolaborazio-plataforma batean bai. Wikipediako edukia mota

askotako erabiltzaileek sortua da, eta, hortaz, espero daiteke akatsen bat edota kalitate eskaseko segmenturen bat egotea, baita jatorrizko testuan ere. Horregatik aproposa iruditu zaigu bonifikazioa honetarako erabiltzea, ziurrenik wikilariak posteditatzean halako aldaketak ere egingo dituztelako, artikulua edukiaren kalitatea eta ulergarritasuna bermatzeko. Dena den, halakoak ez zuzentzeak ez du penalizaziorik izango, lehen esan bezala ulertzen baita PE jardunetik haratago doala, batik bat akatsa jatorrizko testuan dagoelako eta ez hainbeste IAn.

Taula 4 Akatsen sailkapena larritasunaren arabera.

AKATS MOTA	DEFINIZIOA	PENALIZAZIOA/ BONIFIKAZIOA
Arina	Ez dute informazioa aldatzen edo kaltetzen, akats hauekin testuak ulergarria izaten jarraitzen du. Hala ere, testuaren jariakortasunari kalte egiten diote.	-0.5
Larria	Xede testua nahasgarria edo anbiguoak bihurtzen dute, nahiz eta testuak jatorrizkoaren esanahia mantentzen duen.	-1
Kritikoa	Testua ulergaitza bihurtzen dute edota jatorrizkoaren esanahia aldatzen dute. Jatorrizko testuarekin kontraesanak daude.	-3
“Kudos”	Jatorrizkoan eduki edo esanahi aldetik okerrak edota anbiguoak diren esaldiak xede testuan zuzendu dira. Jatorrizkoan gaizki adierazitako ideiak modu ulergarri eta zuzen batean adierazi dira.	+0.5

Testuen akats anotazioak ezin izan dira eranskinean jaso azkenean hartu duten formatuagatik ez direlako bertan txertatzeko egokiak. Hala ere, lanaren azken herenean lortutako emaitzak aztertuko dira.

## **6.7. Egiaztapen proba**

Benetako proba egin aurretik, egiaztapen proba bat egin dugu, ataza egiteko beharko zen denboraren estimazio zehatzago bat egiteko.

Gure inguruko bi pertsonari eskatu diegu, gidalerroez baliatuta, hautatutako testuak posteditatzeko eta erabilitako denbora neurtzeko. Bakoitzak testu bana posteditatu du. Aukeratutako pertsonak itzulpengintzako ikasle bat eta kazetaritzan graduatutako pertsona bat izan dira. Biek euskara maila altua dute (C1-C2).

Behar izan zuten denboran nahiko aldea egon da, batak 45 minutu inguru behar izan zituen eta besteak 15 minutu. Hainbesteko aldea egotearen atzean hainbat arrazoi egon daitezke: atazari aurre egiteko prestakuntza maila, ataza burutzeko interesa, testuak planteaturiko zailtasunak, etab. Oso zaila da zehaztea ze faktorek eragin duen gehien behar izandako denboraren aldean, baina edozein kasutan hori gure helburutik kanpo dago.

Proba honek ahalbidetu digu denbora estimazio bat egitea, eta kalkulatzeko partehartzaileek gehienez ere ordubete inguru beharko zutelako. Estimazioa egiterakoan, gehien tardatu zuen partehartzailearen denbora tarte apur bat handitu dugu, 15 minutu gehitu ditugu. Erabaki hau hartu genuen itzulpenean esperientzia gutxien zutenek motelago posteditatuko zutelakoan.

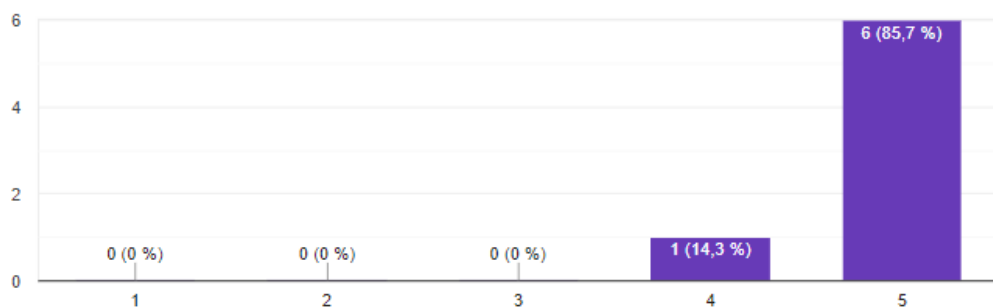
## 7. PROBAREN EMAITZAK

### 7.1. Galdetegiaren emaitzak eta probarako behar izandako denbora

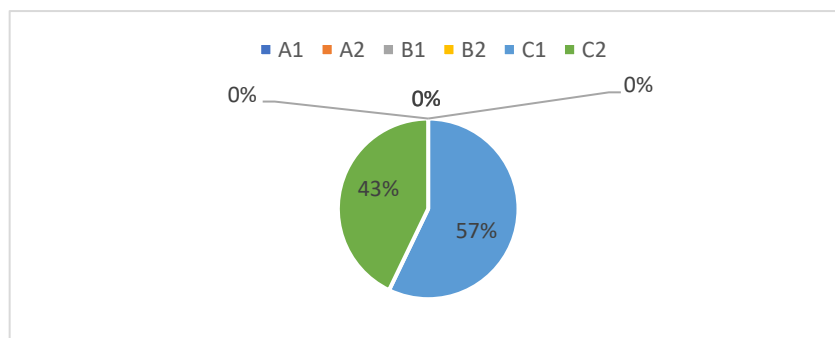
Proban parte hartu zuten 7 partehartzaileak gizonak izan ziren, 35 eta 62 urte artekoak. Gehienak aspalditik dira euskarazko Wikipediako kolaboratzaileak: bostek adierazi zuten 12-15 urte bitartean zeramatzatela, eta beste batek 8 urte zeramatzala. Bakar batek adierazi zuen 3 hilabete baino ez zeramatzala kolaboratzen.

Euskararen aldetik ere orokorrean profila nahiko antzekoa da. Partehartzaile guztiek maiztasun handiarekin erabiltzen dute euskara (1. grafikoa) eta, gainera, euskararen ezagutza aitortzen duen ziurtagiriren bat daukate: lauk (% 57, 14) Europako hizkuntzen Erreferentzia Markoaren C1 mailako ziurtagiriren bat dute; eta, gainerako hirurek (% 42, 85) C2 mailakoa. Beraz, ezagutza maila altua dute denek.

Grafikoa 1 Euskararen erabilera maiztasuna.

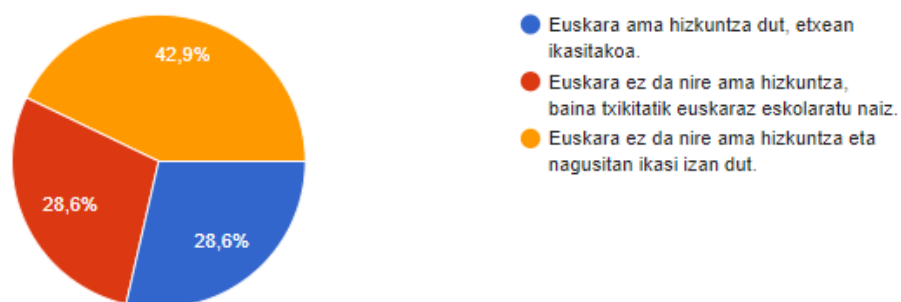


Grafikoa 2 Euskara mailaren ziurtagiria.

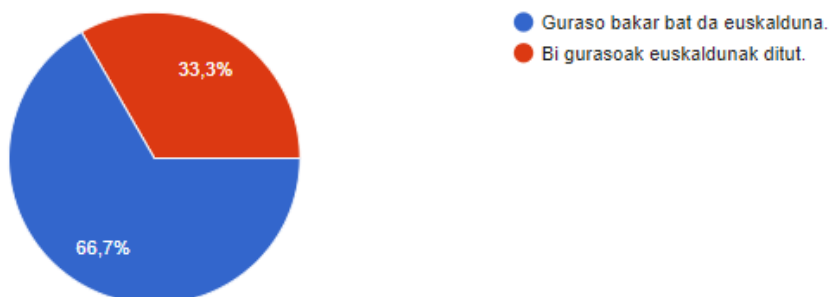


Euskarari dagokionez, alderik handiena euskararen ikasketetan zegoen: 3k etxean ikasia zuten, beste 1ek eskolan eta azkenengo 3rek nagusitan ikasi zuten euskara. Gainera, euskara etxean jaso zutenen artean, 2k guraso bakarra zuten euskalduna eta beste batek biak.

Grafikoa 3 Euskara noiz ikasi dute?

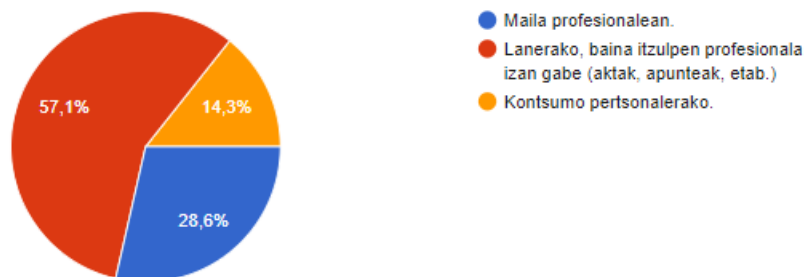


Grafikoa 4 Euskara ama hizkuntza bada, etxean gurasoak euskaldunak dira?



Esperientziaren aldetik, guztiek zuten era bateko edo besteko esperientzia itzulpenean nahiz PEan. Desberdintasun nagusienak helburuan zeuden. Gehienek (% 57,1) adierazi zuten lanerako itzultzen zutela, baina maila profesionalean izan gabe. Beste batek (% 14,3) adierazi zuen kontsumo pertsonalerako itzultzen zuela eta azkeneko biek itzulpena jardun profesional bezala zuten (% 28,6).

Grafikoa 5 Itzulpen esperientzia.



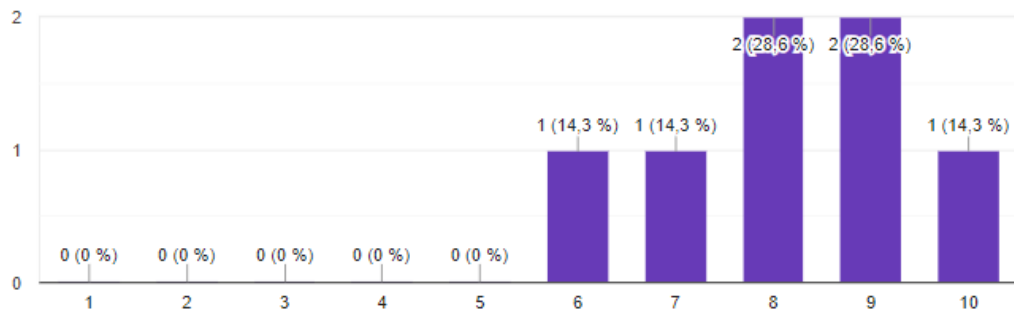
Guztiek adierazi zuten inoiz Wikipedian itzuli eta posteditatu zutela. Gehienek ere adierazi zuten bestelako testuak ere itzuli zituztela, eta testu tipologia aldetik elkarrekiko oso desberdinak diren itzulgaiak aipatu zituzten: esku-orriak, webguneak, liburuen pasarteak, literatura, lanerako dokumentuak eta software libreak. PEaren kasuan, aldiz, testu tipologia mugatua izan da eta, batez ere lanerako dokumentuak itzultzeko erabiltzen dutela adierazi zuten; nolana ere, bakar batek adierazi zuen literatur testu txikiak itzultzeko erabilia zuela.

PERako erabili izan dituzten itzultzaileak ere anitzak izan ziren. Beheko taulan, aipatutako itzultzaileak adierazten dira eta baita zenbaterik aipatu zuten itzultzaile hori. Ikusten da orokorrean gehien erabiltzen dituzten itzultzaile automatikoak Euskal Herrian garatutako itzultzaile neuronalak direla (Elia, Batua eta Itzuli) eta horien artean gehienek aipatu zutena Itzuli izan dela. Partehartzaile gehienek, gainera, adierazi zuten tresna horiek erabiltzean esperientzia positiboa izan dutela.

Taula 5 Erabilitako itzultzaile automatikoak.

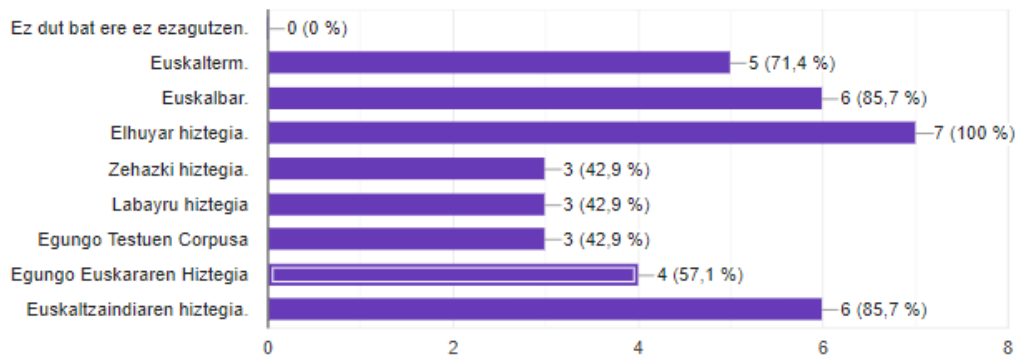
Itzuli	IVAP	Elia	Batua	Google Translate	Beste batzuk (Deepl, Modela, OpenTrad, Matxin)
4	1	2	3	1	1

Grafikoa 6 IAren izandako esperientzia (1etik 10era).



Azkenik, gehienek adierazi zuten itzulpenerako eskuragarri dauden baliabideak (hiztegi elebidunak, corpusak, banku terminologikoak, etab.) ezagutu eta erabiltzen zituztela. Zerrenda batetik ezagutu eta erabiltzen zituzten baliabideak markatzeko eskatu zitzaien, eta gehienek markatu zituztenak Euskalterm, Euskalbar, Elhuyar hiztegia eta Euskaltzaindiaren hiztegia izan dira.

Grafikoa 7 Ezagutu eta erabiltzen dituzten hizkuntza-baliabideak.



## 7.2. Akats anotazioa: azterketa kuantitatiboa

IAk sortutako bertsioa eta posteditatutako bertsioko akatsak anotatu eta alderatu ditugu. Akatsak identifikatu eta ostean sailkatu ditugu, 6.6. azaldutako errubrikei jarraituz. Bestetik, kalitatearen ehuneko (IQ) ateratzeko, TAUSEk (2017) proposatutako formulak erabili dira, eta kalkulua Excel batean egin da.

Taula 6 TAUS Itzulpen kalitatearen kalkulua.

1	Kategoria bakoitzeko penalizazioa = akats arina*akats arin kopurua + akats larri*akats larri kopurua+akats kritiko*akats kritiko kopurua
2	Penalizazio totala=SUM (kategoria bakoitzeko penalizazioa)
3	Penalizazioak hitzeko (%) = Penalizazio totala/Jatorrizkoaren hitz kopurua.
4	Itzulpen kalitatea (IQ)= % 100 – Penalizazioak hitzeko (%)

Taula 7 Notazioaren emaitzak.

	Zehaztasuna		Hizkuntza		Formatua		Totala		
	Akats kop.	Penalizazioa	Akats Kop.	Pen.	Akats Kop.	Penalizazioa	Akats Kop.	Penalizazioa	IQ (%)
IA	7	-8,5	2	-1,5	0	0	9	-10	94,69
A1	1	-0,5	2	-1,5	0	0	4	-2	99,03
A2	1	-0,5	2	-1	0	0	3	-1,5	99,28
A3	2	-1	2	-1,5	0	0	4	-2,5	98,8
A4	2	-1,5	4	-3	0	0	6	-4,5	97,83
IA	3	-4,5	4	-5,5	1	-0,5	9	-10,5	95,52
B1	2	-1	1	-0,5	1	-0,5	4	-2	99,14
B2	1	-0,5	3	-3	1	-0,5	5	-4	98,3
B3	1	-0,5	2	-1,5	1	-0,5	4	-2,5	98,94

Taula honetan jasotzen dira IAk eta posteditore bakoitzak egindako akats kopurua (mantendutakoak nahiz posteditoreak gehitutakoak) eta kategoria bakoitzeko penalizazioa. Azken zutabeetan datu totalak biltzen dira: akats kopuru eta penalizazio totala eta itzulpen kalitatea ehunekoetan emanda.

Lehenik eta behin, aipatu behar dugu IAren bertsioek bi kasuetan kalitate puntuazio altua lortu dutela (IA Q), eta kopuru aldetik ere parekoak direla. Dena den, 5. taulan ikusi daitekeen moduan, bietan akats mota desberdinak egon dira: A testuko (OMS-en ingurukoa) akats gehienak zehaztasun akatsak dira (omisioak, kontraesanak,



etab.); B testukoak, aldiz, gehienak hizkuntza motakoak. Horrez gain, A testuan ez dugu formatu akatsik (kakotxak, kurtsibak, etab.) aurkitu, B testuan ez bezala.

Halaber, PE emaitza onak lortu dira, kasu txarrean ere jatorrizkoa 3 puntu hobetzea lortu da. Hala ere, kontuan izan behar da PE arineko ikuspuntu batetik zuzendu dela; eta, beraz, estilo kontuak alde batera utzi direla. Horiek kontuan hartuz gero puntuazioa baxuagoa litzateke, bai IArena eta bai posteditoreena.

Akats kopuru totala ere kasu guztietan gutxitu da eta, gainera, penalizazio handiena zuen kategorian IAk baino emaitza hobea lortu da. Horrek adierazten du itzultzaileak gai direla hautemateko zeintzuk diren akats larrienak eta soluzio egoki bat proposatzeko. Kategoria guztietan emaitzak hobetu edo bestela IAren emaitza bera mantendu da, salbuespen bat kenduta: A4 posteditorearena. A4 posteditoreak hizkuntza kategorian IAk baino penalizazio altuagoa du eta baita akats kopuru handiagoa ere. Kasu konkretu horretan ikusi dugu IAren akats larrienak identifikatu direla, baina beti ez direla behar bezala ebatzi, eta, ondorioz, zenbait kasutan zehaztasun akatsen ordean hizkuntza akatsak gehitu dira.

Posteditoreen profilari dagokionez, ez da alde handirik ikusi posteditoreen artean amaierako testuaren kalitatearen aldetik, ez behintzat PE arineko ikuspuntutik begiratuta. Dena den, bai ikusi dugu desberdin jokatu dutela posteditatzean: itzulpen profesionalean esperientzia zutenek eta hizkuntza maila altuagoa zutenek aldaketa gehiago egin dituzte, eta maizago proposatu dituzte estiloari eragiten dioten aldaketak, zeinak lotuago dauden PE osoko helburuekin.

### **7.3. Anotazioa: azterketa kualitatiboa**

Atal honetan aztertuko da zer nolako akatsak egin diren ikuspuntu kualitatibo batetik. Behatuko da nolakoak izan diren: 1) mantendutako akatsak eta 2) gehitutako akatsak. Horrez gain, itzultzaile automatikoak duen ahulezia bat ere aztertuko dugu: izen berezien itzulpena. Aurkitutako akats motak argiago azaltzeko, testu lagin desberdinetako adibideak aukeratu eta tauletan jaso dira. Taulen ezkerreko zutabearen adibidearen zenbakia jaso da, ondoren testuan horri egingo baitzaio erreferentzia bertan behatutako akatsa azaltzeko.

Lehenik eta behin, mantendu diren akatsetatik hasiko gara. Batzuetan mantendu diren akatsak oso arinak izan dira, eduki eta ulermenari kalterik egiten ez diotenak. Esate baterako, artikulua indeterminatu baten ordean artikulua determinatu bat jarri da [1. adibidea]. Horrelako akatsak ez dira zenbait ez IA eta PE emaitzetan, ez baitzioten kalterik egiten testuaren ulermenari.

Beste batzuetan ortotipografia kontuak izan dira, esaterako hainbat kasutan kakotxak berrezarri gabe utzi dira [2. adibidea]. Ikusi dugu IAk kakotxa latindarrak ezabatu dituela, eta ingelesak aldiz ez. Hainbat saiakera egin ditugu Elia itzultzailearekin eta ikusi dugu beti egiten duela akats hau. Kakotxak akats arin bezala sailkatu dira, behatutako kasuetan edukiari kalte handirik eragiten ez ziolako, baina beste kasu batzuetarako garrantzitsuagoak izan daitezke, esaterako, aipuen kasuan.

Taula 8 Mantendutako akatsak.

	<b>Jatorrizkoa</b>	<b>PE</b>	<b>Azalpena</b>
<b>1</b>	Para Millet, el sexo es biológico y el género <b>un constructo social</b> .	Milleten aburuz, sexua biologikoa da eta generoa <b>konstruktua soziala</b> .	“Un” artikulua “-a” artikulua determinatuaren bidez itzuli da.
<b>2</b>	El uso del término «género» <b>también</b> se generalizó una década más tarde en el mundo feminista español, [...].	<b>Genero</b> terminoaren erabilera <b>ere</b> hamarkada bat geroago orokortu zen Espainiako mundu feministan, [...].	Kakotxa latindarrak ezabatu dira.
<b>3</b>	[...] considera que <b>los factores culturales inciden en una socialización universal que moldea los comportamientos de los hombres y las mujeres</b> .	[...] uste du <b>kultura-faktoreek eragina dutela gizonen eta emakumeen portaerak moldatzen dituen sozializazio unibertsean</b> .	Erlatibozkoa oso atzean gelditu da eta atzerakarga arazo bat dago.

4	Las funciones principales del Consejo son las de hacer efectivas las decisiones y las políticas de la Asamblea, <b>aconsejarla y facilitar su trabajo.</b>	Kontseiluaren eginkizun nagusiak Batzarraren erabakiak eta politikak gauzatzea, <b>aholkatzea eta lana erraztea</b> da.	Sintaxiagatik, aholkatzeak eta lana errazteak euskarazkoan ez dio Batzarra berari erreferentzia egiten, baizik eta batzarraren erabaki eta politikei.
5	Los Estados miembros de la Organización Mundial de la Salud designan sus delegaciones <b>a la Asamblea</b> Mundial de la Salud, la cual se reúne <b>generalmente</b> en mayo de cada año, [...].	Osasunaren Mundu Erakundeko kide diren estatuek Osasunaren Mundu <b>Batzarrera</b> izendatzen dituzte beren ordezkariak. Batzar hori urtero egiten da, maiatzean, [...].	Deklinabiderik egokiena “Batzarrerako” izango litzateke. “Generalmente” ezabatu da.
6	Algunas entidades pueden también tener estatuto de observador, <b>como lo es la Santa Sede.</b>	Erakunde batzuek behatzaile-estatutua ere izan dezakete, <b>Egoitza Santua</b> , esaterako.	Ergatibo marka falta du.

Beste hainbat kasutan behatu dugu akats batzuek mantentzearen atzean euskararen normalizazio eta ezagutza falta egon daitezkeela. Kasu konkretu batean ikusi dugu batek baino gehiagok arazoak izan dituela ‘ere’ partikula posizio okerrean zegoela hautemateko eta zegokion lekuan jartzeko [2. adibidea]. Gauza jakina da ‘ere’ partikularen posizioak euskaldun askori arazoak ematen dizkiola, baita euskara maila aurreratua dutenen artean ere, beraz, ez da harritzekoa akats hori mantendu izana. Behatu dugu ere beste batzuek zailtasunak izan dituztela atzerakarga arazoak hautemateko [3. adibidea], deklinabide egokiena aukeratzeko [5. adibidea], IAK esaldiaren subjektua aldatu duela hautemateko [4. adibidea], ergatibo akatsak hautemateko [6. adibidea] eta omisio txiki batzuek identifikatzeko [5. adibidea].

Taula 9 Gehitutako akatsak.

	IA	PE	Azalpena
7	Kontseiluaren eginkizun nagusiak Batzarraren erabakiak eta politikak gauzatzea, aholkatzea eta lana erraztea <b>dira.</b>	Kontseiluaren eginkizun nagusiak Batzarraren erabakiak eta politikak gauzatzea, Batzarrari aholkuak ematea eta haren lana erraztea <b>da.</b>	“Batzarrari” gehitu da subjektu aldaketa zuzentzeko, baina aditz komunztadura hautsi da.
8	OME 194 estatu kidek osatzen dute, <b>NBEko estatu kide guztiak barne, Liechtenstein, Amerikako Estatu Batuak eta NBEko kide ez diren bi lurralde: Niue eta Cook Uharteak.</b>	Munduko Osasun Erakundea (edo OME) 194 estatu kidek osatzen dute, <b>NBEko estatu kide guztiak barne hauek ez ezik: Liechtenstein eta Amerikako Estatu Batuak eta NBEko kide ez diren bi lurralde: Niue eta Cook Uharteak.</b>	IAN jatorrizkoaren “excepto” ezabatu da. Posteditorea akatsa zuzentzen saiatu da, baina azkenean birritan erabili ditu bi puntuak esaldi berean, eta hori euskaraz ez da zuzena.
9	Kide behatzaile izateko <b>proposatu du</b> Taiwanek, [...].	Kide behatzaile izateko <b>proposatu da</b> Taiwan [...]	Jatorrizkoaren bihurkaritzaren zentzua galdu da IAN, eta PEan jatorrizkoaren kalko okerra egin da.
10	Genero terminoa modu argiengan eta zehatzenean definitu duen <b>egiletzat jotzen da</b> [...].	Genero terminoa modu argi eta zehatzenean definitu duen <b>egileetako bat jotzen da</b> [...].	Jatorrizkoak dio “como <u>una</u> de las autoras” eta xehetasun hori IAN galdu da. PEan informazio hori

		berreskuratzen saiatu dira, baina akats bat gehitu dute: subjektu eta aditzak ez dute komunztadura zuzen egiten.
--	--	--

Behatu ditugun gehitutako akats batzuek arreta galtzearen ondorioa dirudite. Batzuetan badirudi arreta akats handiago bat zuzentzean jarri dela, eta artean akats txikiren bat gehitu dela, batez ere gramatika mailan [7. adibidea] . Orokorrean tankera horretakoak akats arinak dira eta eduki eta ulermenari ia ez diote eragiten. Gehitutako beste akats batzuek, berriz, agerian uzten dute posteditoreak ez duela ondo jakin nola ebatzi IAren akatsa. Batzuetan saiatu dira mantentzen IAk proposaturiko hitzak, baina zuzendu nahian beste mota batzuetako akatsak gehitu dituzte, batik bat gramatikalak [8., 9. eta 10. adibideak]. Kasu horietan jatorrizkoa zertxobait hobetzea eta argitzea lortu da, baina oraindik ere gelditu diren akatsak batzuetan larriak izan dira.

*Taula 10 Itzultzaile automatikoak izen bereziekin izandako arazoa.*

	<b>JATORRIZKOA</b>	<b>IA</b>
<b>11</b>	[...] en su libro <b>Política sexual</b> [...]	[...] bere <b>Política sexual</b> liburuan [...]
<b>12</b>	[...] actualmente, si son aprobados por mayoría de <b>la asamblea Puerto Rico y Tokelau se convertirán en miembros asociados.</b>	[...] gaur egun, kide horiek Elau kide bihurtuko dira, gehiengoarekin.

Halaber, konprobatu dugu izen bereziak gatazkatsuak direla itzultzaile automatikoarentzako. Kasu batzuetan ondo identifikatu eta itzuli ditu, baina beste batzuetan ez. Esaldi luzeetan bereziki galdu da eta oso esaldi aldrebestua sortu du [11. adibidea]. B testuan obra baten izenarekin beste arazo bat ere hauteman dugu [12. adibidea]. IAk izena jatorrizkoan zetorren bezala mantendu du, hau da, gazteleraz. Posteditore gehienek ere hala jokatu dute, baina bakar batek euskaratu du. Orduan, guri posteditoreei sortu zaien galdera bera sortu zaigu: zelan jokatu halako kasuetan?

Printzipioz guk proposatutako PE arinaren helburuari begira, bi soluzioak dira baliozkoak. Kontua da euskarazko Wikipedian nolabaiteko barne koherentzia bat izatea komeniko litzatekeela, eta horretarako estilo liburu txiki bat osatzea izango litzateke egokiena, batez ere horrelako arazoak maiz agertuko direlako. Dena den, zeregin hori lan honen helburuetatik haratago doa.

#### **7.4. Amaierako gidalerroak**

Guztira 7 PE lagin eta 2 IA lagin baino ez ditugu erabili lanaren dimentsioak direla eta, hortaz, gidalerroei egindako hobekuntzak minimoak dira. Gidalerro oso eta sakon batzuk sortzeko lagin gehiago aztertzea komeniko litzateke. Dena den, hemen egindako proposamenak oinarri bat izan daitezke etorkizunean gidalerro hauek hobetzen etaosatzen joateko.

IAk egindako akatsetan eta aurreko atalean aztertutako PE akatsetan oinarrituta, honako gidalerro definitiboa proposatzen dugu:

**1. Bermatu ez dela informazioa gehitu edo ezabatu.**

**2. Bermatu itzulpenak jatorrizkoaren esanahia biltzen duela eta ez dagoela kontraesanik.**

**3. Berrordenatu esaldia beharrezkoa bada eta zuzendu gramatika akatsak.**

- Kontuz ‘ere’ partikularekin, itzultzaileak batzuetan ez du dagokion tokian jartzen. Gogoratu beti gehitu nahi dugun elementu horren atzean doala.
- Batzuetan itzultzaileak atzerakarga arazoak ditu, esaldi luzeetan aditza oso atzean uzten du. Horrelakoetan posible bada aditza aurreratu, esaldiaren esanahia argitzeko.

**4. Ortografia eta ortotipografia akatsak zuzendu.**

- IA tresnak kakotxa ingelesak mantentzen ditu, baina latindarrak ez, beraz, kontuz.

**5. Izen propioak ondo itzuli dituela egiaztatu.**

- Batzuetan izen bereziak itzultzerakoan arazoak izan ditzake. Gerta daiteke izena omititzea edota izen berezi bat egonagatik esaldia itzultzerakoan galtzea.

**6. Ahal den neurrian, mantendu itzultzaile automatikoak proposatutako hitzak, esaldiaren ordena eta puntuazioa.**

**7. Itzulpen automatikoak sortutako esaldia ez bada batere erabilgarria, errotik itzuli.**

## 8. ONDORIOAK

Lan honetan gaztelaniatik euskarara posteditatzeko gidalerro batzuk osatzen saiatu gara. Izan ere, Wikipedian duela hainbat urtetik ona IA erabiltzen bada ere, oraindik ez zeuden PE gidalerro batzuk eskuragarri. Zentzu horretan gure lana berritzailea da.

Dena den, gidalerroen sorkuntzan murgildu aurretik, euskal PEaren aurkezpen eta azalpen labur bat eskaini dugu, bai ikuspuntu profesionaletik, eta baita guri dagokigun *crowdsourcing* plataformen PEaren ikuspegitik ere. Azkenean, bi alor desberdin badira ere amankomunean ezaugarri asko dituzte, eta, gainera, ikusten ari garenez PEaren iraultza ia aldi berean gertatzen ari da bi testuinguruetan.

Lan honen mamiari heldu diogu gero, gidalerroak sortzeari. Gidalerro baliagarriak sortzeko behar diren urratsak eman ditugu. Horretarako, testuinguru konkretu bat hautatu dugu, gidalerro erabilgarri eta espezifikoa sortzeko aukera izan dezagun, eta PE proba bat egin dugu, non partehartzaileak gidalerroen erabiltzaile potentzial errealek izan diren.

Bestetik, azpimarratu nahi ditugu egindako proban lortutako emaitzak. Ikusi dugu itzultzaile ez-profesionalak ere gai izan direla IAren testuak orrazteko eta kalitate onargarri bateko emaitzak sortzeko, betiere PE arineko ikuspuntu batetik, zeinetan onartzen den akats guztiak zuzendurik ez egotea. Kasu guztietan IAren kalitate emaitza hobetu da eta kasurik txarrean 3 puntuko hobekuntza lortu da.

Horrez gain, ikusi dugu euskararen ezagutzak eta esperientziak ez dutela hainbesteko eraginik amaierako testuaren kalitatean, edo ez behintzat PE arineko helburu bat jarrita. Hala ere, badirudi itzulpen- eta posteditzio-esperientziak PEa egiterako moduan eragin dezakeela. Behatu dugu esperientzia handiagoa dutenek orokorrean aldaketa gehiago egiteko joera dutela, eta gainera estilo aldaketak maizago proposatzen dituztela. Dena den, lan honen dimentsioak txikiegiak dira benetan hori guztia egiaztatzeko, konklusio argiako ateratzeko proba handiago bat egin beharko genuke.

Amaitzeko, gogoratu nahiko genuke guk Wikipediako testu editoretik kanpo lan egin behar izan dugula, eta ezin izan ditugula gidalerroak dagozkien testuinguruan testatu. Hortaz, behin Elia Wikipedian integratuta dagoenean, interesgarria litzakete horrelako beste PE proba bat egitea gidalerroei dagozkien hobekuntzak egiteko.



## 9. BIBLIOGRAFIA

- Allen, J. (2001). Post-editing or no post-editing. *International Journal For Language And Documentation*, (8), 41-42. 2021eko otsailaren 15a, URL: <http://www.mt-archive.info/jnl/IJLangDoc-2001-Allen-1.pdf>
- Aranberri, N. (2016). Ba al dago lekurik euskararako postedizioarentzat?. *SENEZ Aldizkaria*, (47), 195-203.
- Aranberri, N., eta Labaka, G. (2017). Euskarazko Itzulpen Automatikoa (Atal berezia: Besterena nuen neuregana). *SENEZ Aldizkaria*, (48), 219-228.
- Arnold, D., Balkan, L., Meijer, S., Humphreys, R., eta Sadler, L. (1994). *Machine translation: an Introductory Guide* (1 or.). URL: [https://www.researchgate.net/publication/233967760\\_Machine\\_Translation\\_an\\_Introductory\\_Guide](https://www.researchgate.net/publication/233967760_Machine_Translation_an_Introductory_Guide)
- Azou, K. (2016). *La post-edición. Un análisis crítico de unas post-ediciones del crowd, de un hablante nativo y de una estudiante de traducción*. [MAL]. Universiteit Gent. 23-24 orr. <https://lib.ugent.be/catalog/rug01:002271677>
- Collantes, C., Mallo, J., Parra, C., Quiñones, H., eta Serrano, R. (2018). Pásate al laso oscuro: ventajas de la traducción automática para el traductor profesional. *La Linterna Del Traductor*, (17). URL: <http://www.lalinternadeltraductor.org/n17/ventajas-traduccion-automatica.html#:~:text=Carlos%20Collantes%20Fraile%20es%20traductor,pri vado%20como%20en%20entornos%20institucionales>
- Comparin, L., eta Mendes, S. (2017). Using error annotation to evaluate machine translation and human postediting in a business environment. *EAMT 2017 konferentzia*. Praga. 2021eko apirilaren 12a, URL: [https://ufal.mff.cuni.cz/eamt2017/user-project-product-papers/papers/user/EAMT2017\\_paper\\_76.pdf](https://ufal.mff.cuni.cz/eamt2017/user-project-product-papers/papers/user/EAMT2017_paper_76.pdf)
- Content translation (Edukien itzulpena) – MediaWiki (d.g.). In *Mediawiki.org*. 2021eko maiatzaren 10a, URL: [https://www.mediawiki.org/wiki/Content\\_translation/eu](https://www.mediawiki.org/wiki/Content_translation/eu)
- De Almeida, G. (2013). *Translating the post-editor: an investigation of post-editing changes and correlations with professional experience across two Romance*

- languages* [Doktoretza tesia]. Dublin City University. 95 or. URL: [http://doras.dcu.ie/17732/1/THESIS\\_G\\_de\\_Almeida.pdf](http://doras.dcu.ie/17732/1/THESIS_G_de_Almeida.pdf)
- Extension:Translate - MediaWiki. In *Mediawiki.org*. 2021eko maiatzaren 6a, URL: <https://www.mediawiki.org/wiki/Extension:Translate>
- Geiger, D., Roseman, M., Fielt, E., eta Schader, M. (2012). Crowdsourcing Information Systems – Definition, Typology, and Design. *ICIS Konferentzia*, Orlando. 2 or. URL: [https://www.researchgate.net/publication/258031893\\_Crowdsourcing\\_information\\_systems\\_-\\_definition\\_typology\\_and\\_design](https://www.researchgate.net/publication/258031893_Crowdsourcing_information_systems_-_definition_typology_and_design)
- Hu, K., eta Cadwell, P. (2016). A Comparative Study of Post-editing Guidelines. *EAMT 2016 konferentzia*. Riga. 347 or. DOI: [10.13140/RG.2.1.2253.1446](https://doi.org/10.13140/RG.2.1.2253.1446)
- Itzultzaile juridiko-administratiboa - IZO - IVAP. (2020). In *Ivap.euskadi.eus*. 2021eko apirilaren 10a, URL: <https://www.ivap.euskadi.eus/itzultzaile-automatiko-neuronala/z16-a2izo/eu/>
- Irurzun, I. (2018). *Euskarazko postedizioa: aurreiritziak, aplikazioa eta gogoetak itzulpengintza eta interpretazioko ikasleen artean* [GRAL]. URL: <http://hdl.handle.net/10810/30073>
- Mitchell, L. (2015). *Community Post-Editing of Machine-Translated User-Generated Content* [Doktoregotza Tesia]. Dublin City University. URL: [http://doras.dcu.ie/20463/mod1/PhDThesis\\_LindaMitchell\\_Final.pdf](http://doras.dcu.ie/20463/mod1/PhDThesis_LindaMitchell_Final.pdf)
- Rico, C., eta Torrejón, E. (2012). Skills and Profile of the New Role of the Translator as MT Post-editor. *Revista tradumàtica: tecnologies de la traduciò*, (10), 166-178. 2021eko apirilaren 12a, URL: [https://ddd.uab.cat/pub/tradumatica/tradumatica\\_a2012n10/tradumatica\\_a2012n10p166.pdf](https://ddd.uab.cat/pub/tradumatica/tradumatica_a2012n10/tradumatica_a2012n10p166.pdf)
- Siadeco. (2019). *Euskararen eragin ekonomikoa gipuzkoan* [PDF] (82 or.). 2021eko maiatzaren 10a, URL: <https://www.langune.eus/azken-txostena.pdf>
- Schenk, E. eta Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics & Management*, (7), 93-107. <https://doi.org/10.3917/jie.007.0093>

- Soloviev, K. (2017). *Measuring Content Quality with Error Typology: Step by Step Guide*. Blog.taus.net. 2021eko apirilaren 12a, URL: <https://blog.taus.net/measuring-content-quality-with-error-typology-step-by-step-guide>
- TAUS. (2016). *Mt post-editing guidelines*. Amsterdam: TAUS Signature Editions.
- TAUS (D.g.). *DQF-MQM error typology*. Taus.net. 2021eko apirilaren 12a, URL: <https://www.taus.net/qt21-project#harmonized-error-typology>
- Tatsumi, M., Aikawa, T., Yamamoto, K., eta Isahara, H. (2012). How good is crowd post-editing? Its potential and limitations. In *AMTA-2012: Workshop on post-editing technology and practice*. San Diego. URL: <http://www.mt-archive.info/AMTA-2012-Tatsumi.pdf>
- UEU. (2015). *Itzultzaile izatetik, posteditore izatera :: ikasi :: Udako Euskal Unibertsitatea*. Ueu.eus. 2021eko apirilaren 12a, URL: <http://www.ueu.eus/ikasi/udako-ikastaroak>
- MemoQ. (d.g.) *What is Machine Translation?* Translation Software - MemoQ. 2021eko maiatzaren 6 a, URL: <https://www.memoq.com/tools/what-is-machine-translation>
- Wikiproiektu: OpenMT2 eta Euskal Wikipedia. In *Eu.wikipedia.org*. 2021eko maiatzaren 4a, URL: [https://eu.wikipedia.org/wiki/Wikiproiektu:OpenMT2\\_eta\\_Euskal\\_Wikipedia](https://eu.wikipedia.org/wiki/Wikiproiektu:OpenMT2_eta_Euskal_Wikipedia)

## 10. ERANSKINAK

### 10.1. JATORRIZKO TESTUA: OMS

Artikuluua: Organización Mundial de la Salud

Atala: Estructura

Los Estados miembros de la Organización Mundial de la Salud designan sus delegaciones a la Asamblea Mundial de la Salud, la cual se reúne generalmente en mayo de cada año, y tiene la capacidad de definir las políticas financieras de la organización, revisa y aprueba el presupuesto por programas. La Asamblea elige a 34 miembros, técnicos en el campo de la salud, para un mandato de tres años, y que forman el Consejo Ejecutivo. Las funciones principales del Consejo son las de hacer efectivas las decisiones y las políticas de la Asamblea, aconsejarla y facilitar su trabajo.

La OMS está constituida por 194 Estados Miembros, incluyendo todos los Estados Miembros de la [ONU](#), excepto [Liechtenstein](#), los [Estados Unidos de América](#) y dos territorios no miembros de la [ONU](#): [Niue](#) y las [Islas Cook](#), los cuales funcionan bajo el estatuto de asociados (con acceso a la información completa pero con participación y derecho a voto limitados), actualmente, si son aprobados por mayoría de la asamblea [Puerto Rico](#) y Tokelau se convertirán en miembros asociados. Algunas entidades pueden también tener estatuto de observador, como lo es la [Santa Sede](#). [Taiwán](#) se propone como miembro observador, contando con la oposición de [China](#) que lo considera como parte de su territorio.

URL: [https://es.wikipedia.org/wiki/Organizaci%C3%B3n\\_Mundial\\_de\\_la\\_Salud](https://es.wikipedia.org/wiki/Organizaci%C3%B3n_Mundial_de_la_Salud)

## 10.2. JATORRIZKO TESTUA: GENEROA

Artikulua: Género (ciencias sociales)

Azpiatala: El feminismo y los estudios de género

El auge de los estudios [feministas](#) en la [década de 1970](#)<sup>41</sup> comenzó a impulsar en el mundo anglosajón<sup>44</sup> el término *gender* desde un punto de vista específico: las diferencias sociales y culturales, en oposición a las biológicas existentes entre hombres y mujeres; aquí se pueden encontrar los trabajos de [Kate Millett](#), [Germaine Greer](#), [Ann Oakley](#) y [Nancy Chodorow](#), entre otras.<sup>45</sup> El uso del término «género» también se generalizó una década más tarde en el mundo feminista español, especialmente con un sentido social y cultural.<sup>46</sup>

[Kate Millett](#), fue una de las feministas radicales de la década del 70, quien en su libro [Política sexual](#) realiza un análisis de la sexualidad, la política y el género desde el punto de vista biológico, sociológico, antropológico, económico, mitológico y literario entre otros. Para Millet, el sexo es biológico y el género un constructo social. A la hembra humana se la educa para que sea "femenina" y al macho humano para que sea "masculino". Es considerada como una de las autoras que ha definido de la forma más clara y precisa el término género como "aquel constructo social que 'se hace' sobre un sexo determinado naturalizando cualidades y aspectos que se creen vinculados necesariamente a dicho sexo". Para Millett el género es el resultado de una construcción cultural y considera que los factores culturales inciden en una socialización universal que moldea los comportamientos de los hombres y las mujeres.<sup>47</sup>

URL: [https://es.wikipedia.org/wiki/G%C3%A9nero\\_\(ciencias\\_sociales\)](https://es.wikipedia.org/wiki/G%C3%A9nero_(ciencias_sociales))

### **10.3. ITZULPEN AUTOMATIKOA: OMS**

Artikuluua: Osasunaren Mundu Erakundea

Atala: Egitura

Osasunaren Mundu Erakundeko kide diren estatuek Osasunaren Mundu Batzarrera izendatzen dituzte beren ordezkariak. Batzar hori urtero egiten da, maiatzean, eta erakundearen finantza-politikak zehazteko gaitasuna du, programen araberako aurrekontua berrikusi eta onartzen du. Batzarrak 34 kide hautatzen ditu, osasunaren arloko teknikariak, hiru urteko agintaldirako, eta Kontseilu Betearazlea osatzen dute. Kontseiluaren eginkizun nagusiak Batzarraren erabakiak eta politikak gauzatzea, aholkatzea eta lana erraztea dira.

OME 194 estatu kidek osatzen dute, NBEko estatu kide guztiak barne, Liechtenstein, Amerikako Estatu Batuak eta NBEko kide ez diren bi lurralde: Niue eta Cook Uharteak. Horiek elkartekideen estatutuarekin funtzionatzen dute (informazio osoa eskuratzeko aukera dute, baina partaidetza eta boto-eskubidea mugatua dute), eta, gaur egun, kide horiek Elau kide bihurtuko dira, gehiengoarekin. Erakunde batzuek behatzaile-estatutua ere izan dezakete, Egoitza Santua, esaterako. Kide behatzaile izateko proposatu du Taiwanek, eta Txinak aurka egin dio, bere lurraldearen zati dela uste baitu.

#### **10.4. ITZULPEN AUTOMATIKOA: GENEROA**

Artikuluua: Generoa (gizarte-zientziak)

Azpiatala: Feminismoa eta genero-ikasketak

1970eko hamarkadako azterketa feministen gorakada gender terminoa sustatzen hasi zen mundu anglosaxoian, ikuspuntu espezifiko batetik: desberdintasun sozialak eta kulturalak, gizonen eta emakumeen arteko biologikoen aurka; hemen aurki daitezke, besteak beste, Kate Millett, Germaine Greer, Ann Oakley eta Nancy Chodorow. Genero terminoaren erabilera ere hamarkada bat geroago orokortu zen Espainiako mundu feministan, bereziki zentzu sozial eta kulturalarekin.

Kate Millett 1970eko hamarkadako feminista erradikaletako bat izan zen, eta bere *Política sexual* liburuan sexualitatea, politika eta generoa aztertzen ditu ikuspuntu biologiko, soziologiko, antropologiko, ekonomiko, mitologiko eta literariotik, besteak beste. Milleten aburuz, sexua biologikoa da eta generoa konstruktua soziala. Giza emea "femeninoa" izan dadin hezten da, eta gizakia "maskulinoa" izan dadin. Genero terminoa modu argiengan eta zehatzengan definitu duen egiletzat jotzen da "sexu jakin bati buruz 'egiten' den gizarte-konstruktua, nahitaez sexu horri lotuta dauden ezaugarri eta alderdiak naturalizatuz". Milleten aburuz, generoa kultura-eraikuntza baten emaitza da, eta uste du kultura-faktoreek eragina dutela gizonen eta emakumeen portaerak moldatzen dituen sozializazio unibertsalean.