

# Wikipedian posteditatzeko gidalerroak

—  
AMAIA SOLAUN MARTÍNEZ (UPV/EHU)

## Sarrera

Postedizioa (PE) gero eta ohikoagoa da itzulpen merkatuan, baina ez dugu ahaztu behar praktika hori ez dela eremu profesionalera mugatzen. Izan ere, itzultzaile profesionalak ez direnek ere erabiltzen dituzte sarean eskuragarri dauden doako itzultzaile automatikoak. Hori dela eta, erabaki genuen eraldaketa prozesu horren erakusgarri den eremu bat aztertzea: *crowdsourcing* plataformetako postedizioa. Azterketa horren emaitza UPV/EHUko Itzulpengintza eta Interpretazio ikasketetako Gradu Amaierako Lan (GRAL) gisa aurkeztu genuen, 2021ean, Elizabete Manterolaren tutoretzapean; eta, artikulu honetan, GRAL horren laburpena jaso dugu.

*Crowdsourcing* plataformetako edukia jende askoren elkarlanari esker sortzen da, eta gerta daiteke kolaboratzaile horiek adituak nahiz ez-adituak izatea. Horrek esan nahi du, itzulpenaren kasuan, itzultzaile profesional eta ez-profesionalek parte har dezaketela. Eredu horren adibideak dira, besteak beste, TED fundazioa eta Wikipedia.

Guk, zehazki, Euskal Wikipedia aztertzea erabaki genuen, neurri handi batean itzulpenaz elikatzen delako, eta erabiltzaileen artean PEa ohiko praktika delako. Ikerlana gidalerroen sorruntzan zentratu genuen, gaztelaniatik euskararako norabidean, bi arrazoiengatik: alde bate-tik, gidalerroak PEra ezinbesteko baliabide bat izanda ere, oraindik gidalerro publiko gutxi daudelako; eta, bestetik, Wikipediaren kasu konkretuan ez zegoelako gidalerrorik ez euskararako ez beste hizkuntza batzuetarako ere.

Gidalerroak sortzeko, PE proba bat diseinatu genuen: lehenik, gidalerro minimo batzuk sortu genituen; bigarrenik, gidalerro horiek, 7 wikilariren laguntzaz, PE proba batean frogatu genituen; eta, azkenik, probaren emaitzak aztertu, eta haien arabera gidalerro minimoei hobekuntzak egin genizkien. Bestalde, proban parte hartu zuten wikilariei galdetegi bat bidali zitzaien, pertsonen laginaren ezaugarriei behatzeko.

## Postedizio proba

### Oinarri teorikoa

PE probaren diseinurako, Linda Mitchellen doktorego tesian oinarritu ginen, gure lanarekin antzekotasun gehien zituen ikerlana zen heinean. Michellek, postedizio komunitaterako bere tesian, gidalerro batzuk sortu zituen lehenengo aldiz; zehazki, Norton enpresako erabiltzaile-foroko mezuak ingelesetik alemanera posteditatzeko.

Gidalerroak sortzeko, PE proba bat egin zuen, 18 partaideren laguntzarekin. Lehenik, parte hartzaileek galdetegi bat bete zuten, eta, galdetegi horretatik, bi informazio mota eskuratu zituen: informazio demografikoa, eta itzulpenaren eta postedizioaren inguruko iritziak. Horren ostean, partaide bakoitzak 12 testu posteditatu zituen, Michellek sortutako gidalerroen minimoei jarraituta. Ondoren, PEaren emaitzak aztertu ziren, hiru ebaluazio sistema erabilia: 1) akats anotazioa; 2) eremuko espezialistaren ebaluazioa; eta 3) komunitateko kideek egindako ebaluazioa.

Guk, PE probarako, berak jarraitutako eskemari eutsi genion, baina, GRAL baten dimentsioetara egokitzeko, testu eta pertsonen lagina murriztu, eta prozesua sinplifikatu genuen; guk, postedizioaren emaitzak ebaluatzeko hiru sistema erabili beharrean, bakarra erabili genuen: akats anotazioa. Halaber, azpimarratzekoa da, itzultzaile automatikoaren hautuari erreparatu, gure proba eguneratuagoa dagoela, itzultzaile estatistiko baten ordez itzultzaile neuronal bat erabili baikeuen.

### Itzultzaile automatikoaren aukeraketa

Lana hasi aurretik, Euskal Wikipediako Aiora Mujikarekin eta Galder Gonzalezekin, eta baita Elhuyarreko Itziar Cortésekin hitz egin genuen. Haiiek jakinarazi ziguten Elia Wikipediako testu editorean integratzeko asmoa zutela, eta dagoeneko proiektua martxan zegoela. Horregatik, probarako Elia itzultzailea erabiltzea erabaki genuen.

Gure aukeraketa prozesua, beraz, ezohikoa izan zen. Normalean, PE probak egiterakoan, hainbat itzultzaile automatikorekin egiten dira itzulpen probak, eta kalitate handieneko emaitzak sortzen dituen aukeratzeko da, PEaren eraginkortasuna lehen itzulpen gordin horren menpekoea baita. Gure kasuan, aldiz, ez genuen halako alderaketarik egin, eta zuzenean Elia erabili genuen, esan bezala itzultzaile hori izango baitzen Wikipedian integratuko zena.

Proba Wikipediako testu-editoretik kanpo egin genuen, proba egin genuenean Elia oraindik ez baitzegoen Wikipedian integratuta. Printzipioz, askoz egokiagoa da proba erabiltzaileen ohiko lan-tresnan egitea, egoera errealagoa baita, eta sor daitezkeen erabilera arazoak hauteman baitaitezke, baina, gure kasuan, ez zen posible izan. Hala ere, aipatu beharra dago dagoeneko Elia Wikipedian integratuta dagoela, eta, etorkizunera begira proba errepikatuz gero, testuinguru egokienean probatzeko aukera egongo litzatekeela.

## Testuen loginak

Testuen laginetarako, Wikipediako bi artikuluren azpiatal bana erabiltzea erabaki genuen: bata OME erakundearen ingurukoa eta bestea feminismoaren eta generoaren ingurukoa.

Biak gaztelaniazko Wikipediatik atera genituen, testuek Wikipediako artikuluen erakusgarri izan behar baitzuten. Hortaz, aukerarik onena testu errealak aukeratzea zen. Egoera are erreala goa izateko, euskaratu gabeko azpiatalak aukeratu genituen, eta aintzat hartu genuen jadanik euskaraz zegoen artikuluen kalitatea. Azken horretarako, ORES kalitate neurgailuari erreparatu genion, eta puntuazio baxuena zuten artikulua hobetsi genituen, horiek baitira euskarazko edukia falta duten artikulua.

Testuak aukeratzekoan, saiatu ginen orokortasunaren eta espezializazioaren artean oreka bat bilatzen. Argi geneukan testuak ezin zirela oso espezializatuak izan, bi arrazoiengatik: lehenik, parte hartzaileak jakintza alor batekoak baino gehiagotakoak izan zitezkeelako, eta, bigarrenik, gaiaren inguruan zuten jakintza maila desberdinak PEaren kalitatea baldintzatu zezakeelako, eta kasu horretan emaitzak ez zirelako guretzat erabilgarriak izango. Aldi berean, testuak ezin ziren orokorregiak izan, Wikipediako artikuluetan terminologia espezifikoa ager baitaiteke. Azkenean, hori guztia kontuan izanik, erabaki genuen egunerokotasun handiko gai sozialak jorratzen zituzten testuak aukeratzea, nolabait eskuragarriak eta interesgarriak izan daitezkeelako edonorentzat, baina aldi berean izen berezien eta terminologiaren agerpena ere bermatzen dutelako.

Proba gehiegi ez luzatzearen, artikuluen azpiatalekin lantzea erabaki genuen; izan ere, laburrak izanda, ideia edo azpiadar bat jorratzen dute, eta, neurri batean, testu independente gisa funtziona dezakete. Azpiatalen luzera 200-300 hitzera mugatu genuen, bi laginek antzeko luzera izan zezaten.

## Gidalerro minimoen sorkuntza

Kontuan hartuta itzultzaileek ez zutela zertan profesionalak izan, erabaki genuen gidalerroen helburua PE arina izan behar zela.

PE arinak amaierako testu ulergarri eta erabilgarri bat sortzea du helburu, nahiz eta hizkuntza eta estilo aldetik perfektua ez izan. Normalean, irakurleak atzeman dezake itzulpen automatiko (IA) posteditatu bat dela (DePalma 2013, Hu eta Cadwell 2016, 347. or.-an aipatua).

Probarako gidalerroak sortzeko, beste bi gidalerro multzotan oinarritu ginen: alde batetik, TAUSen (2016) PE arinerako gidalerroak, eta, bestetik, Mitchell (2015) bere tesian proposatutakoak.

TAUS (2016)	MITCHELL (2015)
1 Se intenta obtener una traducción correcta desde el punto de vista semántico.	1 Edit the text to make it more fluent and clearer based on your interpretation.
2 Se asegura que no se haya añadido ni omitido ninguna información accidentalmente.	2 Try to correct phrasal ordering and spelling, for example, if they make the text hard or impossible to understand.
3 Se modifica todo contenido ofensivo, inadecuado o no aceptable culturalmente.	3 Use words, phrases or punctuation as they are, if they are acceptable.
4 Se utiliza tanto como sea posible del resultado “en bruto” de la traducción automática.	4 If you are working with reference to the original text, make sure that no information has been added or deleted.
5 Se aplican las reglas básicas con respecto a la ortografía.	
6 No hay que implementar correcciones de naturaleza estilística.	
7 No hay que reestructurar oraciones solo para mejorar el flujo natural del texto.	

1. taula. TAUSen eta Mitchellen gidalerroak.

Gure gidalerroetan, beste gomendio bat ere barneratzea erabaki genuen (7. irizpidea): itzultzaile automatikoaren proposamena baztergarria baldin bada, hobe da errotik itzultzea. Pentsatu genuen PEaren inguruan ezagutza maila desberdinak egon zitezkeela eta horrenbestez komenigarria zela oinarrizko printzipio hori ezartzea. Itzultzaileei aurkeztutako gidalerroak hauek izan ziren:

1. Bermatu ez dela informaziorik 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.
4. Zuzendu ortografia akatsak.
5. Ahal den neurrian, mantendu itzultzaile automatikoak proposatutako hitzak, esaldiaren ordena eta puntuazioa.
6. Itzulpen automatikoak sortutako esaldia ez bada batere erabilgarria, itzuli errotik.

### Ebaluazio sistema

Itzultzaile automatiko baten kalitatea neurtzeko, batez ere hiru metodo erabiltzen dira: ebaluazio automatikoa, giza ebaluazioa eta erabilgarritasun ebaluazioa (Mitchell, 2015: 42). Proba honetarako, giza ebaluazioa erabili genuen, eta, zehazki, akats anotazioa egin genuen.

Akats anotazioa testu batean dauden akatsak hautematean eta sailkatzean datza. Itzultzaile automatiko batek sortutako emaitzaren kalitatea neurtzea ahalbidetzen du, eta baita emaitzak hobetzeko estrategiak diseinatzea ere (Comparin eta Mendes, 2017: 1).

Akats anotazioak, gainera, IA eta PEaren emaitza alderatzeko aukera ematen du (Comparin eta Mendes, 2017; Mitchell, 2015), baldin eta bi testuak anotatzen badira. Metodo horrekin, jakin daiteke zer akats zuzendu diren eta zein ez (Comparin eta Mendes, 2017: 2). Hortaz, oso informazio egokia da gidalero erabilgarri batzuk sortzeko orduan.

Akatsak sailkatzeko, ezinbestean errubrika bat osatu behar da. Gomendatzen da egindako aldaketa indibidual bakoitzari kategoria bat esleitzea. Kategoria eta azpikategoria kopurua aldakorra izan daiteke, gomendatzen baita errubrikaren kasuan kasuko beharretara egokitzea (Solokiev, 2017).

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

**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

1. irudia: Postedizio aldaketen sailkapena (De Almeida, 2013: 95).

Gure errubrika sortzeko, Giselle de Almeidak proposatutako errubrika hartu genuen oinarri (2013: 95), baina euskararako egokitzapen gramatikalak eginez. De Almeidaren errubrikari jarraituz, akats gramatikalak zehaztasun handiz sailkatu zitezkeen. Hori oso erabilgarria zitzaien; izan ere, euskarak eta gaztelaniak arau gramatikal desberdinak dituztenez, ziurrenik gramatikalak izango ziren akats gehienak. Gainera, errubrika egokia da erabiltzaileek sortutako edukiaren PEa ebaluatzeko orduan (Mitchell, 2015: 94).

Akatsen larritasun maila sailkatzeko, berriz, gure eskema propioa sortu genuen, TAUSen MQM eta Comparin eta Mendesen proposamenetik abiatuta (2017: 3). Gure proposamenean, «kudos» kategoria itzulpen egokia bonifikatzeko erabili beharrean, jatorrizkoaren edukia argitzen edo hobetzen laguntzen duten aldatetarako erabiltzen da.

Proposamen hori ez litzateke zentzuzkoa izango postedizio profesionaletarako, baina zentzuduna da Wikipedia bezalako kolaborazio-plataforma baterako. Wikipediako edukia era askotako erabiltzaileen artean sortzen da, eta, hortaz, ez da harrigarria akatsak eta kalitate maila eskaseko segmentuak aurkitzea jatorrizko testuetan. Wikilariek, ziurrenik, horrelako segmentu baten aurrean, edukiaren kalitatea eta ulergarritasuna bermatzearen, PEtik haratago doazen aldatetak barneratuko lituzkete, eta edukia bera ere moldatuko lukete. Dena den, guk egindako PE proban, jatorrizko testuaren eduki-akatsak ez zuzentzeak ez zuen penalizaziorik izan, uler-tzen baita aldaketa horiek PE jardunetik haratago doazela, batik bat akatsa jatorrizko testuan dagoelako eta ez hainbeste IAn.

AKATS MOTA	DEFINIZIOA	PENALIZAZIOA/ BONIFIKAZIOA
Arina	Ez dute informazioa aldatzen edo kaltetzen; era honetako akatsekin, testuak ulergarria izaten jarraitzen du. Hala ere, testuaren jariatortasunari kalte egiten diote.	-0,5
Larria	Xede testua nahasgarria edo anbigua bihurtzen dute, nahiz eta testuak jatorrizkoaren esanahia mantentzen duen.	-1
Kritikoa	Testua ulergaitza bihurtzen dute, edota jatorrizkoaren esanahia aldatzen dute. Jatorrizko testuarekin alderatuta, 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

2. taula: akatsen sailkapena larritasunaren arabera.

## Pertsonen lagina

Hasiera batean, proba 8 pertsonarekin egitea aurreikusi genuen: lau pertsonak testu bat posteditatuko zuten eta beste lauk beste bat. Tamaina horretako lagina egokia zen gure lanaren dimentsioetarako: gidalerroak doitzeko adina informazio eskuratu genezakeen, baina gehiegi

luzatu gabe. Dena dela, azkenean, 7rekin egin genuen proba, deialdian 8 parte hartzaile lortu bagenituen ere, batek azkenean huts egin baitzigun.

Deialdia Wikipediako Txokoa ataletik egin genuen. Mezu bat utzi, eta, gero, Galder Gonzalezek, Euskal Wikilarien Kultur Elkarteko sortzaileetako batek, mezua zabaltzen lagundu zigun.

Proba egiterakoan, partaide bakoitzari galdetegia eta beste hiru dokumentu bidali zitzaizkion, postaz: gidalerroak, jatorrizko bertsioa eta itzulpen automatikoaren (IA) emaitza. Horren ostean, bueltan galdetegia eta posteditatutako dokumentua bidaltzeko eskatu zitzaien.

Parte hartzaileei bidalitako galdetegiari esker, parte hartzaileen profila eratu genuen, ikusi ahal izateko ea euren hizkuntza mailak eta esperientziak euren postedizio emaitzetan eragiten zuen ala ez. Profil hori eratzeke, batez ere lau alderdiri erreparatu genion: informazio demografikoa, euskararen ezagutza eta erabilera maila, Wikipediako esperientzia, eta itzulpen eta postedizio esperientzia.

Parte hartzaile guztiak gizonak izan ziren, 35-62 urte artekoak, eta guztiek, batek izan ezik, denbora luzea zeramaten Wikipedian kolaboratzen (8-15 urte). Denek euskararen ezagutza maila altua zuten (C1 edo C2 ziurtagiria), eta, gainera, maiz erabiltzen zuten.

Esperientziari begira, guztiek zuten era bateko edo besteko esperientzia itzulpenean nahiz PEan, baina ez maila berean: gehienek kontsumo propiorako edo lanerako itzultzen zuten, eta soilik bik jarduten zuten itzulpen profesionalean. Dena dela, denek zuten esperientzia Wikipediako itzulpenean eta postedizioan. Aipatzekoa da, bestalde, guztiek zutela ezagutza ona eskuragarri dauden hizkuntza baliabideen inguruan.

PERako erabilitako tresnak aukeratzekoan, gehienek Euskal Herrian garatutako itzultzaile neuronaletara jotzen zuten (Elia, Batua eta Itzuli), eta bereziki Itzuli erabiltzen zuten. Orokorrean, tresna horiekin izandako esperientzia positiboa zela adierazi zuten (batez besteko nota: 8,14/10).

### Postedizio probaren emaitzak

Akats anotazioa: emaitza kuantitatiboak

Testuen ebaluaziorako, IAK sortutako bertsioa eta posteditatutako bertsioak akatsak anotatu eta alderatu genituen, lehen aurkeztutako bi ebaluazio errubrikeri jarraituta. Bestalde, kalitatearen ehunekoak (IQ) ateratzeko, TAUSek (2017) proposatutako formulak erabili genituen.

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 (%) = guztizko penalizazioa/jatorrizkoaren hitz kopurua
4	Itzulpen kalitatea (IQ) = % 100 – Penalizazioak hitzeko (%)

3.taula: TAUS (2017), itzulpen kalitatearen kalkulua.

	Zehaztasuna		Hizkuntza		Formatua		Gutzizkoa		
	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	98,80
A4	2	-1,5	4	-3	0	0	6	-4,5	97,83
<hr/>									
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	98,30
B3	1	-0,5	2	-1,5	1	-0,5	4	-2,5	98,94

4. taula: akats anotazioaren emaitzak.

4. taulan, IAK eta posteditore bakoitzak egindako akats kopurua (mantendutakoak nahiz posteditoreak gehiturakoak) eta kategoria bakoitzeko penalizazioa jasotzen dira. Azken zutabearen datuen guztizkoak biltzen dira: akats kopuruen eta penalizazioen guztizkoak, eta itzulpen kalitatea ehunekotan emanda.

Lehenik eta behin, aipatu behar dugu IAren bertsioek bi kasuetan kalitate puntuazio altua lortu zutela (IA IQ), eta akats kopuru aldetik ere parekoak zirela. Dena den, ikus daitekeen moduan, bietan akats mota desberdinak egon ziren: A testuko (OMEren ingurukoa) akats gehienak zehaztasun akatsak ziren; B testukoak, aldiz, hizkuntza motakoak. Horrez gain, A testuan ez genuen formatu akatsik (kakotxak, etzanak, etab.) aurkitu, B testuan ez bezala.

PE emaitza onak lortu ziren, kasurik txarrean ere jatorrizkoa 3 puntu hobetzea lortu baitzen. Hala ere, gogoan izan behar da testuak PE arineko ikuspuntutik zuzendu zirela eta horrenbestez estilo kontuak alde batera utzi zirela. Horiek kontuan hartuz gero, puntuazioa baxuagoa litzateke, bai IArena, bai posteditoreena.

Akats kopuruen guztizkoa ere kasu guztietan gutxitu zen, eta, gainera, IAaren kategoria penalizatuenera, emaitza hobetu zen. Horrek adierazten du itzultzaileak gai izan zirela akats larrienak hautemateko eta ondo ebazteko. Kategoria guztietan hobetu ziren emaitzak, edo, bestela, IAren emaitza berdindu zen, salbuespen bat kenduta: A4 posteditorearena. Taulan ikus daitekeenez, A4 posteditoreak IAK baino penalizazio altuagoa du hizkuntza kategorian, eta baita akats gehiago ere. Kasu konkretu horretan, ikusi genuen identifikatu zituela IAren akats larrienak, baina ez zituela beti behar bezala ebatzi; ondorioz, kasu batzuetan, zehaztasun akatsen ordez, hizkuntza akatsak gehitu zituen.



Orokorrean ez genuen ikusi posteditorearen profilak posteditatutako testuaren kalitatea baldintzatu zuenik, ez behintzat PE arinaren ikuspuntutik. Dena den, behatu genuen hizkuntza maila altuagoa zutenek eta itzulpen profesionalean zihardutenek desberdin jokatzen zutela: aldaketa gehiago egin zituzten, eta maizago proposatu zituzten estilo aldaketak, eta halakoak lotuago daude PE osoaren helburuekin. PE osoan, PE arinean ez bezala, giza itzulpenaren pareko kalitatea duen testu bat sortzea bilatzen da, hala hizkuntza aldetik nola estilo aldetik zuzena dena (DePalma, 2013 in Hu eta Cadwell, 2016: 347).

Anotazioa: azterketa kualitatiboa

Ikuspuntu kualitatibo batetik, bi motatako akatsak aztertu genituen: alde batetik, mantendutako akatsak, eta, bestetik, gehitutakoak. Horrez gain, itzultzaile automatikoaren ahulezia bat ere atzeman genuen: izen berezien itzulpena. Aurkitutako akats motak era argi batean azaltzeko, testuen lagin bateko baino gehiagotako adibideak jaso ditugu tauletan.

Mantendutako akatsetik hasita, behatu genuen batzuk oso arinak zirela, edukiari eta ulermenari kalterik egiten ez ziotenak. Esate baterako, artikulua mugagabearen ordeztatu artikulua mugatua erabiltzea (1. adibidea). Horrelako akatsak ez genituen zenbatu IAren eta PEaren emaitzetan, ez baitzioten kalterik egiten testuaren ulermenari.

Jatorrizkoa	Postedizioa
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>
Azalpena: artikulua mugagabea euskaraz artikulua mugatu gisa itzuli da.	

5. taula: 1. adibidea.

Beste batzuetan, ortotipografia arazoak sortu ziren; esaterako, hainbat kasutan, kakotxak berrezarri gabe utzi ziren (2. adibidea). Ikusi genuen IAk kakotx latindarrak ezabatzen zituela eta ingelesak aldiz ez. Elia itzultzailearekin hainbat saiakera egin ondoren, ikusi genuen akatsa beti egiten zuela. Kakotxak akats arin gisa sailkatu genituen, behatutako kasuetan ez baitzioten edukiari kalte handirik eragiten. Hala ere, beste kasu batzuetarako, garrantzitsuagoak izan daitezke; esaterako, aipuen kasuan.

Jatorrizkoa	Postedizioa
El uso del término « <i>género</i> » <i>también se generalizó una década más tarde en el mundo feminista español</i>	Genero terminoaren erabilera <i>ere</i> hamarkada bat geroago orokortu zen Espainiako mundu feministan, bereziki zentzu sozial eta kulturalarekin.
Azalpena: kakotx latindarrak ezabatu dira, eta <i>ere</i> partikula gehitu da okerreko kokalekuan. Izan ere, jatorrizko testuak adierazi nahi duena hauxe da: hamarkada bat geroago Espainiako mundu feministan <i>ere</i> zabaldu zela genero terminoaren erabilera.	

6. taula: 2. adibidea.

Beste akats batzuk mantentzeko arrazoa euskararen normalizazio eta ezagutza falta izan daitezke. Kasu konkretu batean ikusi genuen batek baino gehiagok arazoak zituela *ere* partikula posizio okerrean zegoela hautemateko eta zegokion lekuan jartzeko (2. adibidea). Gauza jakina da *ere* partikularen posizioak arazoak ematen dizkiola euskaldun askori, baita euskara maila aurreratua dutenei ere; beraz, ez da harrizkoa akats hori mantendu izana. Beste batzuek ere zailtasunak izan zituzten, besteak beste, atzerakarga arazoak (3. adibidea) eta ergatibo akatsak hautemateko (4. adibidea).

Jatorrizkoa	Postedizioa
[...] considera que los factores culturales inciden en <i>una socialización universal</i> que moldea los comportamientos de los hombres y las mujeres.	[...] uste du kultura-faktoreek eragina dutela gizonen eta emakumeen portaerak moldatzen dituen <i>sozializazio unibertsalean</i> .
Azalpena: erlatibozkoa oso atzean gelditu da, eta atzerakarga arazo bat dago.	

7. taula: 3. adibidea.

Jatorrizkoa	Postedizioa
Algunas entidades pueden también tener estatuto de observador, <i>como lo es la Santa Sede</i> .	Erakunde batzuek behatzaile-estatutua ere izan dezakete, <i>Egoitza Santua</i> , esaterako.
Azalpena: ergatiboaren marka falta du.	

8. taula: 4. adibidea.

Gehitutako beste akats batzuek arreta galtzearen ondorioa dirudite. Batzuetan, badirudi arreta akats handiago bat zuzentzean jarri dela eta bien bitartean akats txikiak gehitu direla, batez ere gramatika mailan (5. adibidea). Beste batzuetan, omisio txikiak identifikatzeko arazoak izan dituzte, ziurrenik xehetasun horietan ez delako hainbesteko arreta jarri (6. adibidea).

Jatorrizkoa	Las funciones principales del Consejo <i>son las de hacer efectivas las decisiones y las políticas de la Asamblea, aconsejarla y facilitar su trabajo</i> .
Itzulpen automatikoa	Kontseiluaren eginkizun nagusiak <i>Batzarraren erabakiak eta politikak gauzatzea, aholkatzea eta lana erraztea dira</i> .
Postedizioa	Kontseiluaren eginkizun nagusiak <i>Batzarraren erabakiak eta politikak gauzatzea, Batzarrari aholkuak ematea eta haren lana erraztea da</i> .
Azalpena: « Batzarrari » gehitu da subjektu aldaketa zuzentzeko, baina aditz komuntzadura hautsi da.	

9. taula: 5. adibidea.

Jatorrizkoa	Postedizioa
Los Estados miembros de la Organización Mundial de la Salud designan sus delegaciones a la Asamblea Mundial de la Salud, la cual se reúne <i>generalmente</i> en mayo de cada año, [...].	Osasunaren Mundu Erakundeko kide diren estatuek Osasunaren Mundu Batzarrera izendatzen dituzte beren ordekaritzak. Batzar hori urtero egiten da, maiatzean, [...].
Azalpena: <i>generalmente</i> hitza ezabatu da IAn, eta posteditorea ez da ohartu eta aldatu gabe utzi du.	

10. taula: 6. adibidea.

Orokorrean, tankera horretako akatsak arinak dira, eta ia ez diote eragiten edukiari eta ulermenari. Gehitutako beste akats batzuek, berriz, agerian uzten dute posteditoreak ez duela ondo jakin nola ebatzi IAn akatsa. Batzuetan, saiatu dira mantentzen IAn proposaturiko hitzak, baina, zuzendu nahian, beste akats mota batzuk gehitu dituzte, batik bat gramatikalak (7. eta 8. adibideak). Kasu horietan, lortu da jatorrizkoa zertxobait hobetzea eta argitzea, baina gelditu diren akats batzuk larriak dira oraindik ere.

Jatorrizkoa	Taiwán <i>se propone</i> como miembro observador
Itzulpen automatikoa	Kide behatzaile izateko <i>proposatu du</i> Taiwanek, [...].
Postedizioa	Kide behatzaile izateko <i>proposatu da</i> Taiwan [...]
Azalpena: jatorrizkoaren bihurkaritzaren zentzua galdu da IAn («Taiwán se propone». PEan, berriz, gaztelaniazko kalkoa dugu.	

11. taula: 7. adibidea.

Jatorrizkoa	Es considerada como una de las autoras que ha definido de la forma más clara y precisa el término género
Itzulpen automatikoa	Genero terminoa modu argiengan eta zehatzengan definitu duen <i>egiletzat jotzen da</i> [...].
Postedizioa	Genero terminoa modu argi eta zehatzengan definitu duen <i>egiletako bat jotzen da</i> [...].
Azalpena: jatorrizkoak dio «como una de las autoras». Xehetasun hori IAn galdu da. Posteditorea saiatu da informazio hori berreskuratzen, baina komunztadura hautsi du.	

12. taula: 8. adibidea.

Halaber, konprobatu dugu izen bereziak gatazkatsuak direla itzultzaile automatikoarentzat. Kasu batzuetan, ondo identifikatu eta itzuli ditu, baina beste batzuetan ez. Esaldi luzeetan, bereziki galdu da, eta oso esaldi aldrebestua sortu du (9. adibidea). B testuan, arazoak izan ditu obra baten izenburuarekin ere (10. adibidea).

Jatorrizkoa	Itzulpen automatikoa
[...] actualmente, si son aprobados por mayoría de <i>la asamblea Puerto Rico y Tokelau se convertirán en miembros asociados.</i>	[...] <i>gaur egun, kide horiek Elau kide bihurtuko dira, gehiengoarekin.</i>

13. taula: 9. adibidea.

Jatorrizkoa	Itzulpen automatikoa
en su libro <i>Política sexual</i> [...]	[...] bere <i>Política sexual</i> liburuan [...]

14. taula: 10. adibidea.

IAn, jatorrizkoan zetorren bezala mantendu da izena, hau da, gaztelaniaz. Posteditore gehienek ere hala jokatu dute, eta bakar batek euskaratu du. Orduan, posteditoreei sortu zaien galdera bera sortu zaigu: zelan jokatu halako kasuetan?

Printzipioz, guk proposatutako PE arinaren helburuari begira, bi konponbideak dira baliozkoak. Kontua da komeniko litzatekeela euskarazko Wikipedian nolabaiteko barne koherentzia bat izatea, 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.

## Amaierako gidalerroak

Pertsonen eta testuen laginak txikiak zirenez, gidalerroei egindako hobekuntzak minimoak dira. Gidalerro oso eta sakon batzuk sortzeko, komeniko litzateke lagin handiagoa aztertzea. Dena den, hemen egindako proposamenak oinarri bat izan daitezke etorkizunean gidalerroak hobetzeko eta osatzeko. Hona hemen amaierako gidalerroak:

1. Bermatu ez dela informaziorik 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 baitu dagokion tokian jartzen. Gogoratu beti gehitu nahi dugun elementu horren atzean doala.
  - Batzuetan, itzultzaileak baditu atzerakarga arazoak: esaldi luzeetan, aditza oso atzean uzten du. Horrelakoetan, posible bada, aurreratu aditza, esaldiaren esanahia argitzeko.
4. Zuzendu ortografia eta ortotipografia akatsak.
  - IAko tresnak kakotx ingelesak mantentzen ditu, baina latindarrak ez; beraz, kontuz.

5. Egiaztatu izen propioak ondo itzuli dituela.
  - Batzuetan, izen bereziak itzultzerakoan, arazoak izan ditzake. Gerta daiteke izena ezabatzea 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, itzuli errotik.

## **Ondorioak**

Wikipedian hainbat urtetatik hona IA erabiltzen bada ere, oraindik ez zegoen PE gidalerrik eskuragarri. Hori dela eta, gaztelaniatik euskarara posteditatzeko gidalerro batzuk proposatu genituen, posteditzio proba batean oinarrituta, zeinetan parte hartzaileak erabiltzaile errealak izan baitziren. Alde horretatik, gure lana berritzailea da.

Azpimarratzekoak dira proban lortutako emaitzak. Ikusi genuen itzultzaile ez-profesionalak ere gai direla IAren testuak orrazteko eta kalitate onargarri bateko emaitzak sortzeko, betiere PE arinaren ikuspuntutik, zeinetan onartzen baita akats guztiak zuzendurik ez egotea. Kasu guztietan, IAren kalitate emaitza hobetu zen, eta, kasurik txarrean, 3 puntuko hobekuntza lortu zen.

Horrez gain, alde txikiak ikusi genituen itzulpen esperientzia handiena eta txikiena zuten artean. Esperientzia handienekoen orokorrean aldaketa gehiago egiteko joera agertu zuten, eta estilo aldaketak maizago proposatu zituzten. Bestalde, parte hartzaileen euskara maila nahiko antzekoa zen, eta, beraz, ez genuen alderdi horretan desberdintasun handirik ikusi. Dena den, lanaren tamaina txikiegia da benetan hori guztia egiaztatzeko: ondorio argiako ateratzeko, beharrezkoa litzateke proba handiago bat egitea.

Etorkizunera begira, interesgarria litzateke horrelako posteditzio proba bat berriz egitea. Guk testu editoretik kanpo lan egin genuen, Elia ez baitzegoen oraindik Wikipedian erabiltzeko prest, eta ezin izan genituen gidalerroak dagozkien testuinguruan probatu. Hala ere, lehen aipatu dugun bezala, Elia dagoeneko Wikipedian integratuta dago, eta, beraz, egun posible litzateke PE proba Wikipediako testu editorean egitea.

## BIBLIOGRAFIA

- COMPARIN, Lucia eta MENDES, Sara. (2017). Using error annotation to evaluate machine translation and human postediting in a business environment. *EAMT 2017 konferentzia*. Praga. [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)
- DE ALMEIDA, Giselle. (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. [http://doras.dcu.ie/17732/1/THESIS\\_G\\_de\\_Almeida.pdf](http://doras.dcu.ie/17732/1/THESIS_G_de_Almeida.pdf)
- HU, Ke. eta CADWELL, Patrick. (2016). A Comparative Study of Post-editing Guidelines. *EAMT 2016 konferentzia*. Riga. 347. or. DOI: 10.13140/RG.2.1.2253.1446
- MITCHELL, Linda. (2015). *Community Post-Editing of Machine-Translated User-Generated Content* [Doktoretza Tesia]. Dublin City University. [https://doras.dcu.ie/20463/1/PhDThesis\\_LindaMitchell\\_Final.pdf](https://doras.dcu.ie/20463/1/PhDThesis_LindaMitchell_Final.pdf)
- SOLOKIEV, Kirill. (2017). *Measuring Content Quality with Error Typology: Step by Step Guide*. Blog.taus.net. <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 <https://www.taus.net/qt21-project#harmonized-error-typology> (Kontsulta: 2021-04-12)

### **Resumen**

El objetivo de este artículo ha sido crear unas pautas de posesición para poseditar en la Wikipedia del castellano al euskera. Para ello, hemos creado unas pautas mínimas y las hemos puesto a prueba en una prueba de posesición en la que han participado 7 colaboradores de la Wikipedia. Después se han analizado los resultados de la prueba y se han propuesto mejoras a las pautas mínimas iniciales.

### **Résumé**

L'objectif de cet article est de poser des lignes directrices pour la post-édition de contenus de l'espagnol au basque dans Wikipédia. Pour ce faire, nous avons créé quelques lignes directrices et les avons passées à l'épreuve d'un test de post-édition auquel 7 contributeurs de Wikipédia ont participé. Les résultats du test ont ensuite été analysés et des améliorations apportées aux lignes directrices initiales.

### **Abstract**

The aim of this article was to create guidelines for Spanish to Basque post-editing on Wikipedia. To do this, we created minimum guidelines and subjected them to a post-editing test in which 7 Wikipedia contributors participated. The test results were then analyzed and the initial minimum guidelines have been improved.