

Euskal itzulpen ikasketak eta humanitate digitalak

TRALIMA-ITZULIK ikerketa-taldearen ekarpenak eta beharrak

Itzulpena, literatura, ikus-entzunezkoak
Euskal Herriko Unibertsitateko ikerketa-taldea (GIU21/060)

Sarrera

TRALIMA/ITZULIK ikerketa-taldearen egitekoa kulturen arteko itzulpena eta hartu-emanak ikertzea da, nola literatura testuetan hala ikus-entzunezko produktuetan, ikuspegi historikoa eta deskribatzailea erabilia, eta teknologia berrietako tresnez baliatuz (Humanitate Digitalak). Hala, orain arte landu ditugun ikerketa-lerroen ardatza izan da idatziz nahiz ikus-entzunezkoetako itzulpenek agertzen dituzten mekanismo linguistikoak, kulturalak eta sozialak aztertzea. Ikerketa horiek ikuspegi historikoa erabilia egin ditugu, metodologia deskribatzaile komun batez baliatuta, eta teknologia berriak garatu eta aplikatu ditugu. Hain zuzen ere, lan egiteko modu hori bihurtu da taldearen ezaugarri nagusi. Horiek horrela, ikerketa-taldearen ikerlerro nagusiak honako hauek dira: corpusetan oinarritutako itzulpen ikasketak, euskaratik eta euskararako itzulpena, literatura itzulpena, ikus-entzunezko itzulpena eta irisgarritasuna, eta itzulpengintzaren didaktika. Ekarpenean honetan, ikerlerro horien testuinguruan ikerketa-taldea burutu dituen ekimenak eta abian dituenak aurkeztuko dira, euskarazko datu digitalekin lan egiten dugun talde gisa ditugun beharrak edo gabeziak azpimarratuz.

Ikerketa-lerro nagusiak

- ⇒ Corpusean oinarritutako itzulpen ikasketak deskribatzaileak
- ⇒ Euskaratik eta euskararako itzulpena
- ⇒ Literatura itzulpena
- ⇒ Itzulpenaren didaktika
- ⇒ Ikus-entzunezko itzulpena

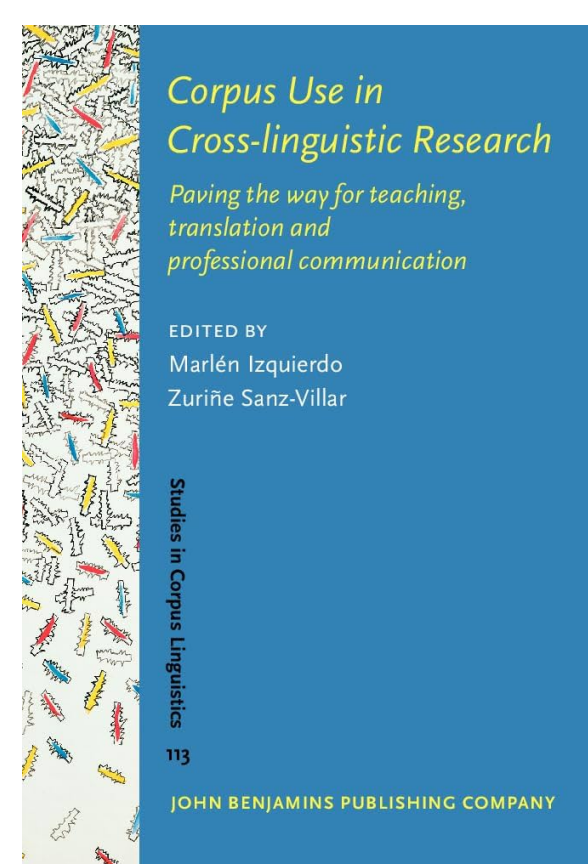
Kongresua eta argitalpena

2021eko ekainaren 23tik 25era bitartean, corpus paraleloei buruzko PaCor 2021 nazioarteko kongresua antolatu zuten TRALIMA-ITZULIK ikerketa-taldeko kideek.



<https://www.ehu.eus/en/web/pacor2020/>

Kongresu horretatik sortutako argitalpena 2023an eman da argitara:



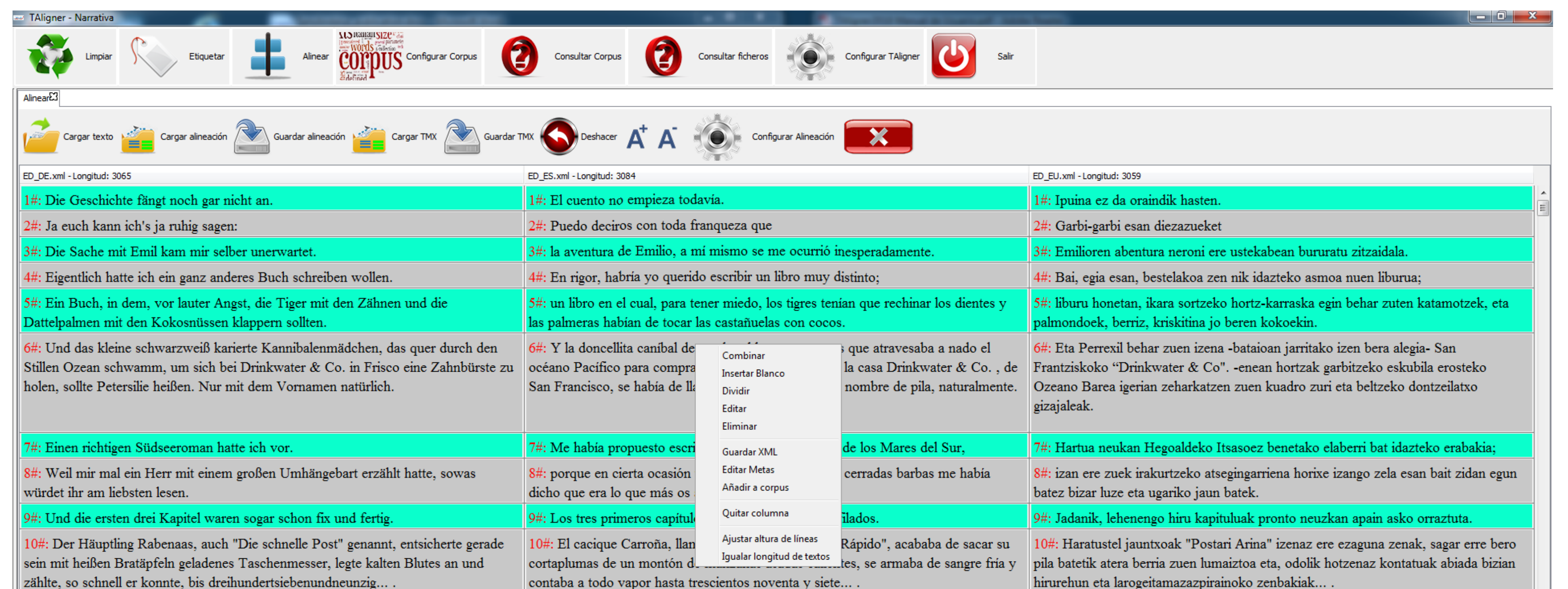
Etorkizunean

- Ikus-entzunezko itzulpenaren automatizazioa, euskaraz ardatz hartuta
- Itzulpengintza/interpretazioaren didaktika eta teknologia berriak
- Itzulpen tresna automatikoetako genero alborapena

TAligner 3.0 (eskuragarri: ADDI, CorpusNet)

Informatika eta itzulpengintza alorretako ikertzaileen **elkarlana** funtsezkoa izan da TAligner sortzeko. **Corpus eleaniztun eta paraleloak sortzeko eta ustiatzeko tresna** da. Honakoak ditu **ezaugarri nagusiak**:

- Erraz abia daiteke, eta erabiltzaileek ez dute laguntza teknikorik beharko. Hainbat motatako testuak biltzen dituzten corpusak sor daitezke.
- TAlignerren bitartez hainbat testu lerroka daitezke aldi berean, ikertzaileak bere lanerako behar dituen adina.
- Sortuko den corpusa tresnan bertan ustia daiteke.



Hala, **hainbat motatako corpusak sortu** dituzte dagoeneko taldekideek. Corpus horiek ezinbestekoak izan dira itzulitako testuen deskribapena egin eta horiek aztertzeko. Jarraian aipatuko ditugu corpus horietako batzuk:

- Euskaratik beste hizkuntza batzuetara egindako itzulpenen corpusa (Manterola 2012)
- Alemanetik euskararako literatura testuak, helduei zein haurrei zuzenduak, biltzen dituen corpusa (Sanz-Villar 2015; Zubillaga 2014)
- Euskaratik gaztelaniara edo frantsesera itzuliriko literatur testuen corpusa (Arrula 2019)
- Ingelesetik gaztelaniarako antzezlanen corpusa (Andaluz-Pinedo 2022)
- Leonard Cohenen abestien gaztelaniarako itzulpenen corpusa (Ros Abaurrea 2023)

Bi dira etorkizunera begira aurreikusten ditugun **erronkak**:

- Tresna garatzen jarraitzea. Horretarako, elkarlana ezinbestekoa izango da.
- Sortuta dauden corpusak sarreratzea.

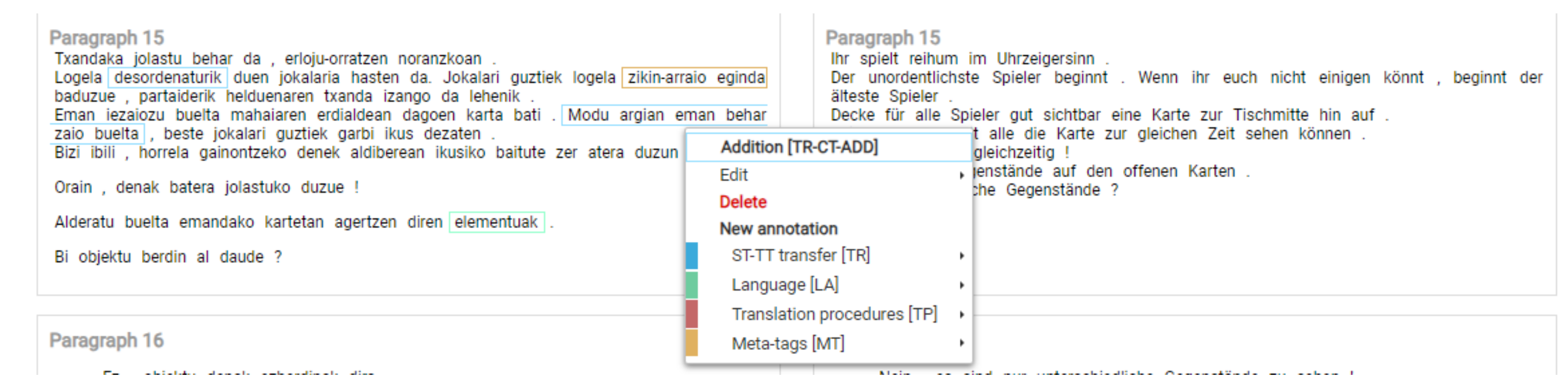
MUST

MUST (Multilingual Student Translation) nazioarteko proiektuko kide gara TRALIMA-ITZULIK ikerketa-taldeko hainbat kide. Ikasleen itzulpenez osaturiko corpus eleaniztun bat sortu eta anotatzea dugu helburu nagusi eta corpus horrek aplikazio praktikoa izatea itzulpen-eskoletan.

Orain arte hainbat corpus sortu ditugu, eta corpus horien azterketen emaitzak argitaratu (Izquierdo eta Zubillaga, bidalia; Sanz-Villar, onartua):

Pair	Total texts	NEW texts	PROGRESS texts	FINISHED texts	ST tokens	TT tokens	Active students	Annotations
en-es	65	48	0	17	22691	26307	0	399
eu-es	53	0	0	53	20275	29875	0	531
de-es	0	0	0	0	0	0	0	0
en-eu	122	38	0	84	49813	40942	16	1534
de-eu	104	36	0	68	38240	33692	9	1525

Corpusak sortzeaz gain ikasleen itzulpenak anotatzeko aukera ere badugu:



Mugak:

- Metadatuak ezin ditugu gure beharretara egokitu.
- Ezin ditugu euskarazko testuak POS mailan etiketatu.
- Corpusa esportatu eta Sketch Engine-ra inportatzeko zailtasunak ditugu.

