

# Goi-mailako testu akademikoak lantzeko baliabideak eta tresnak

María Jesús Aranzabe, Igone Zabala, Izaskun Aldezabal

Euskal Hizkuntza eta Komunikazioa Saila; HITZ zentroa, Ixa Taldea  
Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)

{maxux.aranzabe,igone.zabala,izaskun.aldezabal}@ehu.eus



## Testuingurua eta motibazioa

- Nahitaezkoa da hizkuntza bat erabiltzea eremu akademikoetan, eremu horretako adierazpide bereizgarriak garatzeko.
- Hizkuntza-komunitate batek adierazpide akademikoak garatzeari eta erabiltzeari uzten dionean, erabilera-eremu espezializatuak galtzeko arriskua du (Lauren et al., 2002).
- Eremu akademikoaren konkista erabakigarria izan da euskara biziberritzeko (Zabala, 2019).
- Baina euskararen erregistro akademikoak oraindik garapen- eta egonkortze-prozesuan daude (Zabala et al., 2011; Zabala et al., 2021).

## Helburua

- Euskararen erregistro akademikoaren garapena ikertzeko eta garapenean laguntzeko sortu ditugun tresnak eta baliabideak aberasten jarraitzea.
- Alor akademikoan erreferentziako lan-ingurunea izatea (Aranzabe et al., 2022a) eta lankidetzak bilatzea, sare dinamiko nahiz kolaboratiboak sortzeko.



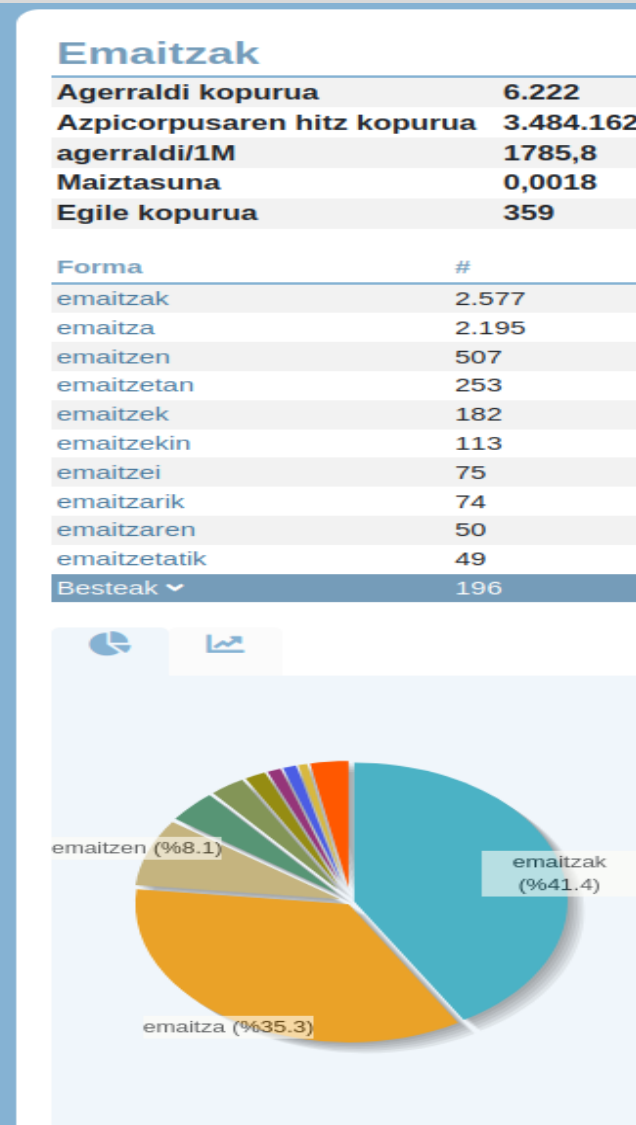
# LAGUNTZA-TRESNAK ETA KONTSULTA-BALIABIDEAK

## Tresnak eta baliabideak garatzeko ABIAPUNTUA

Unibertsitateko komunikazio-egoera espezializatueta erabiltzen diren testuak, terminologia eta fraseologia.

**CORPUSAK: Garaterm (Zabala et al., 2013) eta HARTAEus (Aranzabe et al., 2022b)**  
<http://garaterm-corpora.ixaxa.eu/>

### Adibidea: "emaitza" hitzaren bilaketa arrunta corpusean



Komplementatzele-Sintagma nahasiak euskara-gaztelania kode-alfabetizazioan (Master amaierako lana [MAL] [Arañak eta Humanitateak] [Zientzia Teknologikoak] [2])  
Lau **EMAITZA** hitzak dira zuzenean 1) kontzeptu posible guztien artean, ...  
...aztertua bada osatze adieraz, baina ez bada **EMAITZA** berdinak aurkitu egingo diren uztarretan.  
BRUNO ZABALTEZ PROZESUA ETA ARKITEKTURA ARTETERIAK HISTORIKO GRADUA (Gradu amaierako lana [GRAL] [Arañak eta Humanitateak] [Arteak eta Letren Zientziak] [2])  
...diseinatu garatzen den garapenaren emaitza izan da **EMAITZA** hitzak eta hurrengo harrizketan jada agertzen eta uztarretan ...  
...kutsuak izan da eremu espezializatuaren emaitza izan da **EMAITZA** hitzak eta hurrengo harrizketan jada agertzen eta uztarretan ...  
Euskarazko perpaus erlatiboaren joserka, jostailu zaharrentzat gutxiak erabiltzen (Master amaierako lana [MAL] [Arañak eta Humanitateak] [Zientzia Teknologikoak] [3])  
...diseinatu garatzen den garapenaren emaitza izan da **EMAITZA** hitzak eta hurrengo harrizketan jada agertzen eta uztarretan ...  
...kutsuak izan da eremu espezializatuaren emaitza izan da **EMAITZA** hitzak eta hurrengo harrizketan jada agertzen eta uztarretan ...  
Performanceakurte bildumetan (Gradu amaierako lana [GRAL] [Arañak eta Humanitateak] [Arteak eta Letren Zientziak] [4])  
...kutsuak izan da eremu espezializatuaren emaitza izan da **EMAITZA** hitzak eta hurrengo harrizketan jada agertzen eta uztarretan ...  
...kutsuak izan da eremu espezializatuaren emaitza izan da **EMAITZA** hitzak eta hurrengo harrizketan jada agertzen eta uztarretan ...  
Elikanoko Jostailu Lizaragarren lexikokoaren azterketa (Master amaierako lana [MAL] [Arañak eta Humanitateak] [Zientzia Teknologikoak] [2])  
...kutsuak izan da eremu espezializatuaren emaitza izan da **EMAITZA** hitzak eta hurrengo harrizketan jada agertzen eta uztarretan ...  
...kutsuak izan da eremu espezializatuaren emaitza izan da **EMAITZA** hitzak eta hurrengo harrizketan jada agertzen eta uztarretan ...

- Garaterm corpora:**
  - Euskarazko corpus akademiko bakarra.
  - Irakaskuntza-materialez eta ikerketa-lan akademikoak osatutakoa.
  - Corpus monitoria: azterketa diakronikoak → Etengabeko aberastea.
  - Testu espontaneoak eta kontrolatutakoak: egokiak idatzizko eta ahozko testuak aztertzeko eta irakaskuntzan erabiltzeko.

- HARTAEus corpora:**
  - 3.285.098 hitz  
398 egile
  - Ikasleen lan akademikoak (GrAL eta MAL) osatutako corpora.
  - Gaztelaniako HARTA-noveles del español (Villayandre, 2018) corpusarekin konparagarria.

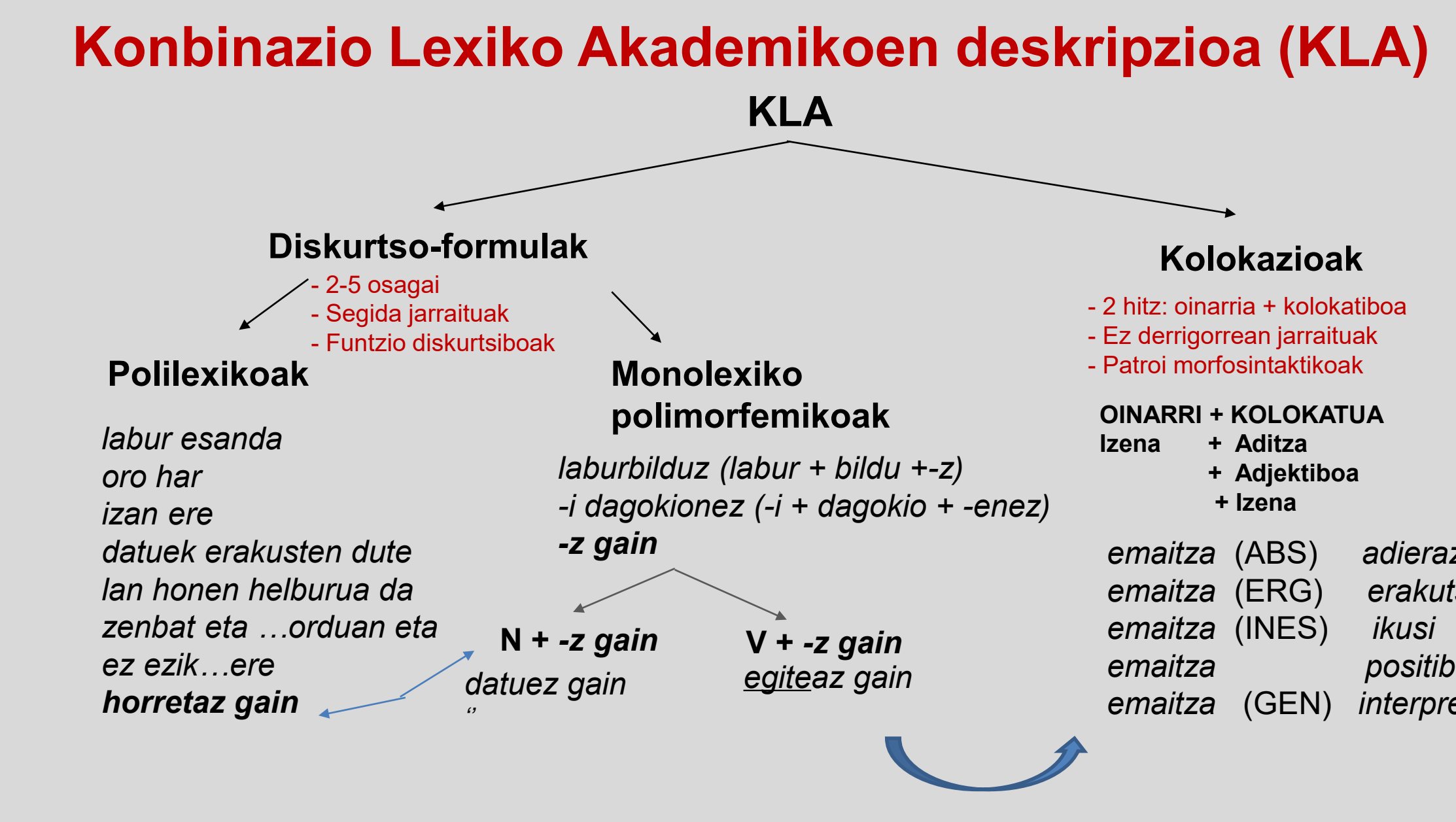
## TZOS terminologia-datubasea (Aldezabal et al., 2022; Arregi et al., 2013)

<http://tzos.ehu.eus/>

- Komunikazio akademikoan erabiltako terminologiaren deskripzioa (deskripzio aktiboa).
- Corpusak gidatutako terminologia-lana.
- Garaterm corpusarekin konektatuta: terminoak beren testuinguruan kontsultatzeko aukera (*in vivo*).

## HARTA/TAILA: Testu Akademikoaren Idazketarako Laguntza tresna elebiduna

<http://harta.ixaxa.eu/>



### Erreferentzia bibliografikoak

- Aldezabal, I., Arriola, J.M., Otegi, A. (2022). TZOS: an Online Terminology Database Aimed at Working on Basque Academic Terminology Collaboratively. En N. Calzolari (Conference chair), F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, J. Ojidi, S. Piperidis (Eds.), *Proceedings of the 13th Language Resources and Evaluation Conference* (pp. 1353-1359). European Language Resources Association (ELRA).
- Aranzabe, M.J., Aldezabal, I. y Zabala, I. (2022a). Recursos y Herramientas de Lingüística de Corpus y PLN para la Monitorización e Investigación de los Usos Académicos del Euskera. III. workshop de INTELE (Infraestructura de Tecnologías del Lenguaje).
- Aranzabe, M.J., Gurrutxaga, A. y Zabala, I. (2022b). Compilación del corpus académico de noveles en euskera HARTAEus y su explotación para el estudio de la fraseología académica. *Procesamiento del Lenguaje Natural*, 69, 95-103. <https://doi.org/10.28342/2022-69-8>
- Arregi, X., Arruarte, A., Artola, X., Zabala, I. y Lersundi, M. (2013). TZOS: An On-Line System for Terminology Service. En *Actualizaciones en Comunicación Social. Actas XIII Simposio Int. de Comunicación Social* (400-404). Centro de Lingüística Aplicada.
- Laurén, Ch., Myking, J. y Picht, H. (2002). Language and domains: a proposal for a domain domains taxonomy. LSP and Professional Communication, 2(2), 23-30.
- Zabala, I., San Martín, I., Lersundi, M. y Elordui, A. (2011). Graduate teaching of specialized registers in a language in the normalization process: Towards a comprehensive and interdisciplinary treatment of academic Basque. En S. Maruenda-Bataler y B. Clavel-Arriola (Eds.), *Multiple voices in academic and professional discourse* (pp. 208-218). Cambridge Scholars.
- Zabala, I., Lersundi, M., Leturia, I., Manterola, I. y Santander, G. (2013). GARATERM: euskararen erregistro akademikoaren garapenean ikerketarako lan-ingurunea. En X. Alberdi y P. Salaburu (Eds.), *Ugarteburu terminologia jardunaldiak (V). Terminologia naturala eta terminologia planifikatuak euskararen normalizazioari begira* (98-114). Servicio Editorial de la UPV/EHU.
- Zabala, I. (2019). The elaboration of Basque in Academic and Professional Domains. En L. Grenoble, P. Lane y U. Rögnvald (Eds.), I. Igarua y L. Oñederra (Basque Eds.), *Linguistic Minorities in Europe Online*, De Gruyter Mouton. <https://doi.org/10.1515/ime.9612443>.
- Zabala, I., Aranzabe, M.J. y Aldezabal, I. (2021). Retos actuales del desarrollo y aprendizaje de los registros académicos orales y escritos del euskera. *Círculo de Lingüística Aplicada a la Comunicación*, 88, 31-50.

### Finantziak:

- EUSKO JAURLARITZA GOBIERNO VASCO
- IXA TALDEA
- HARTAEus proiektua A motako PID2019-109683GB-C22), ikertalde finkatua Ministerioko finantziatuta (IT-1570-22).
- CLARIAH-EUS II. workshop. Donostia, 2023-11-23