

# Ikerkuntza

Tesiak

Argitalpenak

## Tesiak

1. *Oesophageal Speech: Enrichment and Evaluations*

**Sneha Raman**

Zuzendariak: Inma Hernaez Rioja and Eva Navas Cordon

2. *Txosten klinikoak euskararen eta gazteleraren artean itzultzen laguntzeko corpusaren bilketa eta itzultzaile automatikoaren garapena [Corpus compilation and development of a machine translation system for translating clinical reports between Basque and Spanish]*

**Xabier Soto Garcia**

Zuzendariak: Gorka Labaka Intxauspe and Maite Oronoz Anchordoqui

## Argitalpenak

### ALDIZKARIAK

1. Alberto Blanco, Alicia Pérez, Arantza Casillas

*Exploiting ICD Hierarchy for Classification of EHRs in Spanish Through Multi-Task Transformers*

(IEEE Journal of Biomedical and Health Informatics, v. 26, n. 3, pp. 1374-1383.)

center

Language

2. Xabier Arregi

*Hizkuntza-teknologiak euskaldunon artean*

Hermes aldizkaria, n. 69, pp. 78-82.

3. Beatriz Pereda-Goikoetxea, María Isabel Elorza-Puyadena Mikel Lersundi-Ayestaran Joseba Xabier Huitzi-Egilegor María José Uranga-Iturrioz Blanca Marín-Fernández

*Emakumeen emozio-zurrunbiloa erditzean*

Ekaia, n. 41, pp. 31-48.

4. Igone Zabala, María Jesús Aranzabe, Izaskun Aldezabal

*Retos actuales del desarrollo y aprendizaje de los registros académicos orales y escritos del euskera*

Círculo de Lingüística Aplicada a la Comunicación, v. 88, pp. 31-50.

5. Ainara Estarrona, Izaskun Etxeberria, Ricardo Etxepare, Manuel Padilla-Moyano, Ander Soraluze

*The First Annotated Corpus of Historical Basque*

Digital Scholarship in the Humanities, Advance Article, pp. 1-14.

6. Anastasia Klimovich-Gray, Ander Barrena, Eneko Agirre, Nicola Molinaro

*One Way or Another: Cortical Language Areas Flexibly Adapt Processing Strategies to Perceptual And Contextual Properties of Speech*

Cerebral Cortex, v. 31, n. 9, pp. 4092-4103.

7. Irune Ibarra, Unai Atutxa, Mikel Iruskieta

*Idazteko zailtasunak detektatzeko eta esku hartzeko proposamena Lehen Hezkuntzan*

Ikastorratza, n. 27, pp. 1-29.

8. Unai Bermejo, Aitor Almeida, Aritz Bilbao-Jayo, Gorka Azkune

*Embedding-based real-time change point detection with application to activity segmentation in smart home time series data*

Expert Systems with Applications, v. 185, December 15, pp. 1-12.

9. Rodrigo Agerri, Roberto Centeno, María Espinosa, Joseba Fernández de Landa, Álvaro Rodrigo

*VaxxStance@IberLEF 2021: Overview of the Task on Going Beyond Text in Cross-Lingual Stance Detection*

Procesamiento del Lenguaje Natural, n. 67, September, pp. 173-181.

10. Joseba Fernandez de Landa, Rodrigo Agerri  
*Social analysis of young Basque-speaking communities in twitter*  
Journal of Multilingual and Multicultural Development, Advance Article, pp. 1-15.
11. Pablo Gamallo, Gorka Labaka  
*Using Dependency-Based Contextualization for transferring Passive Constructions from English to Spanish*  
Procesamiento del Lenguaje Natural, n. 66, March, pp. 53-64.
12. Xavier Gómez Guinovart, Itziar Gonzalez-Dios, Antoni Oliver, German Rigau  
*Multilingual Central Repository: a Cross-lingual Framework for Developing Wordnets*  
arXiv preprint arXiv:2107.00333, pp. 1-11.
13. Unai Atutxa, Alejandro Molina-Villegas, Mikel Iruskieta  
*Generación automática de meta-resúmenes para la evaluación del manejo de estructuras discursivas y coherencia en el alumnado*  
Procesamiento del Lenguaje Natural, n. 66, March, pp. 165-175.
14. Iakes Goenaga, Xabier Lahuerta, Aitziber Atutxa, Koldo Gojenola  
*A Section Identification Tool: towards HL7 CDA/CCR Standardization in Spanish Discharge Summaries*  
Journal of Biomedical Informatics, v. 121, September, pp. 1-11.
15. Jan Deriu, Alvaro Rodrigo, Arantxa Otegi, Guillermo Echegoyen, Sophie Rosset, Eneko Agirre, Mark Cieliebak  
*Survey on Evaluation Methods for Dialogue Systems*  
Artificial Intelligence Review, v. 54, June 25, pp. 755-810.
16. Ander Barrena, Aitor Soroa, Eneko Agirre  
*Towards Zero-Shot Cross-Lingual Named Entity Disambiguation*  
Expert Systems With Applications, v. 184, December 1, pp. 1-9.
17. Jon Alkorta  
*Hacia el análisis de sentimientos en euskera*  
Procesamiento del Lenguaje Natural, n. 66, March, pp. 201-204.

18. Joseba Fernandez de Landa, Rodrigo Agerri  
*Euskarazko on-line artikuluetan aipatutako izendun entitate nabarmenen identifikazioa denbora errealean*  
Ekaia, n. 40, pp. 315-328.
19. Elena Zotova, Rodrigo Agerri, German Rigau  
*Semi-automatic generation of multilingual datasets for stance detection in Twitter*  
Expert Systems with Applications, v. 170, May 15, pp. 1-13.
20. Cristina Aceta, Johan Kildal, Izaskun Fernández, Aitor Soroa  
*Towards an optimal design of natural human interaction mechanisms for a service robot with ancillary way-finding capabilities in industrial environments*  
Production & Manufacturing Research, v. 9, n. 1, pp. 1-32.
21. Ainhoa Serna, Aitor Soroa, Rodrigo Agerri  
*Applying Deep Learning Techniques for Sentiment Analysis to Assess Sustainable Transport*  
Sustainability, v. 13, n. 4, pp. 1-19.
22. Aitzol Elu, Gorka Azkune, Oier Lopez de Lacalle, Ignacio Arganda-Carreras, Aitor Soroa, Eneko Agirre  
*Inferring spatial relations from textual descriptions of images*  
Pattern Recognition, v. 113, May, pp. 1-10. [Q1]
23. Sara Santiso, Alicia Pérez, Arantza Casillas  
*Adverse Drug Reaction extraction: Tolerance to entity recognition errors and sub-domain variants*  
Computer Methods and Programs in Biomedicine, v. 199, February, pp. 1-9.
24. Sneha Raman, Xabier Sarasola, Eva Navas, Inma Hernaez  
*Enrichment of Oesophageal Speech: Voice Conversion with Duration-Matched Synthetic Speech as Target*  
Applied Sciences, v. 11, n. 13, pp. 1-14.
25. Luis Serrano, Sneha Raman, Inma Hernández, Eva Navas, Jon Sanchez, Ibon Saratxaga  
*A Spanish multispeaker database of esophageal speech*  
Computer Speech and Language, v. 66, March, pp. 1-12.

## AKTAK

1. Bernardo Magnini, Begoña Altuna, Alberto Lavelli, Manuela Speranza, Roberto Zanolì

### *The E3C Project: European Clinical Case Corpus*

Proceedings of the Annual Conference of the Spanish Association for Natural Language Processing: Projects and Demonstrations (SEPLN-PD), pp. 17-20.

2. Horacio Saggion, Dimitar Shterionov, Gorka Labaka, Tim Van de Cruys, Vincent Vandeghinste, Josep Blat

### *SignON: Bridging the gap between and Sign and Oral Languages*

Proceedings of the Annual Conference of the Spanish Association for Natural Language Processing: Projects and Demonstrations (SEPLN-PD), pp. 21-24.

3. Itziar Gonzalez-Dios, Kepa Bengoetxea

### *MultiAzterTest@VaxxStance-IberLEF 2021: Identifying Stances with Language Models and Linguistic Features*

Proceedings of the Iberian Languages Evaluation Forum (IberLEF), co-located with the Conference of the Spanish Society for Natural Language Processing (SEPLN), pp. 192-201.

4. Oscar Sainz, Oier Lopez de Lacalle, Gorka Labaka, Ander Barrena, Eneko Agirre

### *Label Verbalization and Entailment for Effective Zero- and Few-Shot Relation Extraction*

Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1199-1212.

5. Yi-Ling Chung, Marco Guerini, Rodrigo Agerri

### *Multilingual Counter Narrative Type Classification*

Proceedings of The 8th Workshop on Argument Mining, pp. 125-132.

6. Sara Santiso

### *Adverse Drug Reaction extraction on Electronic Health Records written in Spanish: A PhD Thesis overview*

IberSPEECH 2020, pp. 155-159.

7. Kepa Sarasola, Iñaki Alegria, Olatz Perez de Viñaspre

### *Language Technology for Language Communities: an overview based on our experience (2020)*

In: Language and Technology in Wales, v. 1. Delyth Prys (ed.). University of Bangor, Bangor, pp. 10-20.

8. Sergio Santana, Alicia Pérez, Arantza Casillas, Maite Oronoz

*Erlazio-erazketa testu klinikoetan hizkuntzaren prozesamenduaren bidez*

In: IV. Ikergazte Nazioarteko Ikerketa Euskaraz. Olatz Arbelaitz, Ainhoa Latatu, Miren Josu Omaetxebarria, Blanca Urgell (eds). UEU, Bilbo, pp. 59-66.

9. Amir Zeldes, Yang Janet Liu, Mikel Iruskieta, Philippe Muller, Chloé Braud, Sonia Badene

*Proceedings of the 2nd Shared Task on Discourse Relation Parsing and Treebanking (DISRPT 2021)*

Association for Computational Linguistics.

10. Iker García-Ferrero, Rodrigo Aggerri, German Rigau

*Benchmarking Meta-embeddings: What Works and What Does Not*

Proceedings of the Findings of the Association for Computational Linguistics (EMNLP), pp. 3957-3972.

11. Mikel Iruskieta, Maria del Mar Boillos

*Los modelos discursivos de los trabajos finales de grado del área de educación en euskera*

Congreso Universitario Internacional sobre Contenidos, Investigación, Innovación y Docencia, p. 250.

12. Irune Ibarra, Mikel Iruskieta

*Smartpen and language corpora: handwriting interventions in children with learning difficulties*

e-LADDA CON 2021: Understanding Language and Literacy Development in the Digital Age

13. Cristina Aceta, Izaskun Fernández, Aitor Soroa

*Ontology Population Reusing Resources for Dialogue Intent Detection: Generic and Multilingual Approach*

Proceedings of the International Conference on Recent Advances in Natural Language Processing, pp. 10-18.

14. Cristina Aceta, Izaskun Fernandez, Aitor Soroa

*TODO: A Core Ontology for Task-Oriented Dialogue Systems in Industry 4.0*

In: Further with Knowledge Graphs (Proceedings of the 17th International Conference on Semantic Systems). Mehwish Alam, et. al. (eds), Studies on the Semantic Web (v. 53), IOS Press Ebooks, pp. 1-15.

15. Alberto Blanco, Sonja Remmer, Alicia Pérez, Hercules Dalianis, Arantza Casillas  
*On the Contribution of Per-ICD Attention Mechanisms to Classify Health Records in Languages With Fewer Resources than English*

Proceedings of the International Conference on Recent Advances in Natural Language Processing, pp. 165-172.

16. Ander Cejudo, Owen Trigueros, Alicia Pérez, Arantza Casillas, Daniel Cobos  
*Verbal Autopsy: First Steps Towards Questionnaire Reduction*

In: Text, Speech, and Dialogue (24th International Conference TSD). Kamil Ekštejn, František Pártl, Miloslav Konopík (eds), Lecture Notes in Artificial Intelligence, v. 12848. Springer, Cham, pp. 111-123.

17. Oscar M. Cumbicus-Pineda, Itziar Gonzalez-Dios, Aitor Soroa

*Linguistic Capabilities for a Checklist-based evaluation in Automatic Text Simplification*

In: Proceedings of the First Workshop on Current Trends in Text Simplification (CTTS), co-located with the 37th Conference of the Spanish Society for Natural Language Processing (SEPLN). Horacio Saggion, Sanja Štajner, Daniel Ferrés, Kim Cheng Sheang (eds), Online, pp. 70-83.

18. Kepa Bengoetxea and Itziar Gonzalez-Dios

*MultiAzterTest@Exist-IberLEF 2021: Linguistically Motivated Sexism Identification*

Proceedings of the Iberian Languages Evaluation Forum (IberLEF), co-located with the Conference of the Spanish Society for Natural Language Processing (SEPLN), pp. 449-457.

19. Kepa Bengoetxea, Itziar Gonzalez-Dios

*MultiAzterTest: a Multilingual Analyzer on Multiple Levels of Language for Readability Assessment*

Proceedings of the Iberian Languages Evaluation Forum (IberLEF), co-located with the Conference of the Spanish Society for Natural Language Processing (SEPLN), pp. 1-33.

20. Oscar Cumbicus, Itziar Gonzalez-Dios, Aitor Soroa

*A Syntax-Aware Edit-based System for Text Simplification*

Proceedings of Recent Advances in Natural Language Processing, pp. 324-334.

21. Jon Alkorta, Koldo Gojenola, Mikel Iruskietea

*Ezeztapena identifikatzeko Murriztapen Gramatikako erregelaksentimenduen analisiaren testuinguruan*

In: IV. Ikergazte Nazioarteko Ikerketa Euskaraz. Olatz Arbelaitz, Ainhoa Latatu, Miren Josu Omaetxebarria, Blanca Urgell (eds). UEU, Bilbo, pp. 169-176.

22. Aitor Ormazabal, Mikel Artetxe, Aitor Soroa, Gorka Labaka, Eneko Agirre

*Beyond Offline Mapping: Learning Cross Lingual Word Embeddings through Context Anchoring*

Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 6479–6489.

23. Kepa Sarasola, Nora Aranberri

*Language-centered AI will allow additional EU official languages for 2025*

Proceedings of the Conference on Linguistic Rights and Language Varieties in Europe in the Age of AI. Artificial Intelligence for European Integration (AI4EI), pp. 1–11.

24. Itziar Gonzalez-Dios, Uxoa Iñurrieta, Igone Zabala

*General and Specialised Corpora to Raise Linguistic Awareness in a Language Undergoing the Normalisation Process: Academic Writing in Basque*

In: Proceedings of AELFE-TAPP (19th AELFE Conference, 2nd TAPP Conference). Elisabet Arnó, Marta Aguilar, Judith Borràs, Guzmán Mancho, Balbina Moncada, Dietmar Tatzl (eds), Universitat Politècnica de Catalunya, Vilanova i la Geltrú, pp. 1–7.

25. Oscar Sainz, German Rigau

*Ask2Transformers: Zero-Shot Domain labelling with Pretrained Language Models*

Proceedings of the 11th Global WordNet Conference, pp. 44–52.

26. Igone Zabala, Izaskun Aldezabal, Maria Jesus Aranzabe

*Academic Research Works and Domain Dynamics: Resources and Tools for Basque Academic Writing*

18th International Conference on Minority Languages.

27. Ander Salaberria, Jon Ander Campos, Iker García, Joseba Fernandez de Landa

*Itzulpen Automatikoko Sistemen Analisia: Genero Alborapenaren Kasua*

In: IV. Ikergazte Nazioarteko Ikerketa Euskaraz. Olatz Arbelaitz, Ainhoa Latatu, Miren Josu Omaetxebarria, Blanca Urgell (eds). UEU, Bilbo, pp. 153–160.

28. Joseba Fernandez de Landa, Iker García, Ander Salaberria, Jon Ander Campos

*Twitterreko Euskal Komunitatearen Eduki Azterketa Pandemia Garaian*

In: IV. Ikergazte Nazioarteko Ikerketa Euskaraz. Olatz Arbelaitz, Ainhoa Latatu, Miren Josu Omaetxebarria, Blanca Urgell (eds). UEU, Bilbo, pp. 137–144.



29. Eneko Agirre, Marianna Apidianaki, Ivan Vulić (Editors)  
*Proceedings of Deep Learning Inside Out (DeeLIO): The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures*  
Association for Computational Linguistics, Online edition.
30. Cecilia Domingo, Tatiana Gonzalez-Ferrero, Itziar Gonzalez-Dios  
*What is on Social Media that is not in WordNet? A Preliminary Analysis on the TwitterAAE Corpus*  
Proceedings of the 11th Global Wordnet Conference, pp. 234-242.
31. Ionut-Teodor Sorodoc, Madhumita Sushil, Ece Takmaz, Eneko Agirre (Editors)  
*Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop*  
Association for Computational Linguistics, Online edition.
32. Lana Yeganova, Dina Wiemann, Mariana Neves, Federica Vezzani, Amy Siu, Inigo Jauregi Unanue, Maite Oronoz, Nancy Mah, Aurélie Névéol, David Martinez, Rachel Bawden, Giorgio Maria Di Nunzio, Roland Roller, Philippe Thomas, Cristian Grozea, Olatz Perez de Viñaspre, Maika Vicente Navarro, Antonio Jimeno Yepes  
*Findings of the WMT 2021 Biomedical Translation Shared Task: Summaries of Animal Experiments as New Test Set*  
Proceedings of the Sixth Conference on Machine Translation, pages 664–683.
33. Víctor García, Inma Hernández, Eva Navas  
*Implementation of neural network based synthesizers for Spanish and Basque*  
IberSPEECH 2021, pp. 225-229.
34. Agustín Alonso, Víctor García, Inma Hernández, Eva Navas, Jon Sánchez  
*Automatic Speaker Adaptation Assessment Based on Objective Measures for Voice Banking Donors*  
IberSPEECH 2021, pp. 210-214.
35. Inma Hernández, José Andrés González-López, Eva Navas, José Luis Pérez Córdoba, Ibon Saratxaga, Gonzalo Olivares, Jon Sánchez de la Fuente, Alberto Galdón, Víctor García Romillo, Miriam González-Atienza, Tanja Schultz, Phil D. Green, Michael Wand, Ricard Marxer and Lorenz Diener  
*Voice Restoration with Silent Speech Interfaces (ReSSInt)*  
IberSPEECH 2021, pp. 130-134.

36. Víctor García Romillo, Inma Hernaéz, Eva Navas

*The AHOLAB Text-to-Speech system for Blizzard Challenge 2021*

The Blizzard Challenge

## LIBURU KAPITULUAK

1. Igone Zabala

*Euskaltzaindiaren Hiztegiaren ekarpena lexiko espezializatuaren eta ez-espezializatuaren harmonizazioan*

In: Arantzazutik mundu zabalera. Euskararen normatibizazioa: 1968-2018. Andres Urrutia (ed.). Euskaltzaindia-Iberoamericana Vervuert, Bilbao, pp. 285-299.

2. Adrián Núñez-Marcos, Gorka Azkune, Ignacio Arganda-Carreras

*Exploiting Egocentric Cues for Action Recognition for Ambient Assisted Living Applications*

In: Emerging Technologies in Biomedical Engineering and Sustainable TeleMedicine. Jihad Alja'am, Somaya Al-Maadeed, Osama Halabi (eds). Springer, Cham, pp. 131-158.

3. Cristina Cumbreño, Nora Aranberri

*What Do You Say? Comparison of Metrics for Post-editing Effort*

In: Explorations in Empirical Translation Process Research. Michael Carl (eds), Machine Translation: Technologies and Applications, v. 3. Springer, Cham, pp. 57-79.

## HITZALDIAK

1. Mikel Iruskieta

*Helduen hizkuntza i(ra)kaskuntzan hizkuntza-teknologiak: zenbait arrakasta eta euskararen etorkizuneko erronkak*

Euskarabidea 2020-2021 ikasturtearen irekitze-ekitaldira.

2. Mikel Iruskieta

*Herramientas Digitales para las Humanidades Digitales en la e-infraestructura CLARIN*

Curso de Verano: Creación de un Proyecto en Humanidades Digitales basado en el análisis de textos: Modelado y Procesamiento