

Elena Álvarez Mellado

INFORMATION SCIENCES INSTITUTE
UNIVERSITY OF SOUTHERN CALIFORNIA

✉ ealvarezmellado@gmail.com
💻 lirondos.github.io
🐦 [@lirondos](https://twitter.com/lirondos)
🔗 github.com/lirondos
in linkedin.com/in/elenaalvarezmellado/

Research interests

Computational Linguistics • Natural Language Processing • Corpus Linguistics

Education

- 2018-2020 *Computational Linguistics MS*, Brandeis University (GPA: 3.95/4.00)
MS Thesis: *Lázaro: An Extractor of Anglicisms in Spanish Newswire*. Advisor: Constantine Lignos.
Coursework: *Fundamentals of Computational Linguistics; Statistical Neural Approaches to NLP; NLP Systems: Named-Entity Recognition; NLP in Python; Information Extraction; Linguistic Annotation for Machine Learning; Data Structures Algorithms; Advanced Programming in Java.*
- 2008-2010 *Linguistics BA*, Universidad Complutense de Madrid (Grade: 7.59/10)
Erasmus scholarship at Université Paris Diderot (Paris 7) in 2009.
Final project: *AZRAEL. A-Z Reconocedor Automático de Español*. Advisor: Eduardo Basterrechea.

Research & Professional Experience

- 06/2020 RESEARCH PROGRAMMER
Present *Information Sciences Institute, University of Southern California (USC ISI)* – MA, USA
Supervised by Liz Boschee
- Programmer at USC ISI Center for Vision, Image, Speech and Text Analytics (VISTA).
- 06/2019 RESEARCH DATA ANALYST
12/2019 *McLean Hospital, Harvard Medical School* – MA, USA
Supervised by Dr. Mei-Hua Hall, advised by prof. James Pustejovsky
- Applied Machine Learning and NLP techniques to health records to predict readmission risk.
- 06/2017 JUNIOR RESEARCHER
06/2018 *Digital Humanities Lab, Universidad Nacional de Educación a Distancia (UNED)* – Madrid, Spain
Supervised by prof. Elena González-Blanco and prof. Salvador Ros
- Researcher in POSTDATA, ERC project on Poetry Standardization & Linked Open Data
 - Developed annotation guidelines for Named Entity Recognition on medieval texts.
- 12/2014 LINGUISTIC DATA ANALYST
01/2016 *Fundación del Español Urgente (Fundéu)* – Madrid, Spain
- Conducted ARACNE, a Corpus Linguistics study on the evolution of the Spanish language on the media during the 20th century.
- 10/2010 ANALYTICAL LINGUIST
1/2016 *Molino de Ideas* – Madrid, Spain
- Developed, annotated and evaluated linguistic resources for Spanish language.

01/2010	INTERNSHIP STUDENT
06/2010	<i>Lingway-Eptica</i> – Paris, France <ul style="list-style-type: none"> Annotated keywords for automatic CV matching.
06/2009	INTERNSHIP STUDENT
09/2009	<i>Molino de Ideas</i> – Madrid, Spain <ul style="list-style-type: none"> Provided morphological annotation of a database of Spanish conjugated verbs.
10/2008	RESEARCH ASSISTANT
03/2009	<i>Department of Linguistics, Universidad Complutense de Madrid</i> – Madrid, Spain

Skills

Coding: Python (spaCy, sklearn, PyTorch), Java. Also familiar with: Prolog, Haskell, MySQL.

Languages: Spanish (native), English (fluent; CEFR: C2; TOEFL: 112; CPE: A), French (basic).

Publications

CONFERENCE PAPERS

2021	Elena Álvarez Mellado. Extracting English lexical borrowings from Spanish newswire. <i>Society for Computation in Linguistics (SCiL 2021)</i> (Forthcoming).
2020	Elena Álvarez Mellado. Two models for Named Entity Recognition in Spanish. <i>Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2020)</i> , colocated with SEPLN 2020. [pdf]
2020	Elena Álvarez Mellado. A corpus of Spanish political speeches from 1937 to 2019. <i>Proceedings of The 12th Language Resources and Evaluation Conference (LREC)</i> . Marseille, France. ISBN: 979-10-95546-34-4. [pdf]
2020	Elena Álvarez Mellado. An Annotated Corpus of Emerging Anglicisms in Spanish Newspaper Headlines. <i>Proceedings of the 4th Workshop on Computational Approaches to Code Switching</i> , colocated with LREC 2020. Marseille, France. ISBN: 979-10-95546-66-5. [pdf]
2019	Elena Álvarez Mellado, Eben Holderness, Nicholas Miller, Fyonn Dhang, Philip Cawkwell, Kirsten Bolton, James Pustejovsky and Mei-Hua Hall. Assessing the Efficacy of Clinical Sentiment Analysis and Topic Extraction in Psychiatric Readmission Risk Prediction. <i>Proceedings of the 10th International Workshop on Health Text Mining and Information Analysis</i> , colocated with EMNLP 2019. Hong Kong, China. ISBN 978-1-950737-77-2. [pdf]
2017	Elena Álvarez Mellado, Leticia Martín-Fuertes Moreno. Aracne: estudio de la variación lingüística en la prensa española entre 1914 y 2014. In <i>III Congreso de la Sociedad Internacional de Humanidades Digitales Hispánicas</i> . Málaga, Spain. ISBN: 978-84-697-5692-8. [pdf]

JOURNAL ARTICLES

IN PRESS	Elena Álvarez Mellado, M. Luisa Díez Platas, Pablo Ruiz Fabo, Helena Bermúdez, Salvador Ros Muñoz, Elena González-Blanco. TEI-friendly annotation scheme for medieval named entities. Currently under review for <i>Language Resources and Evaluation</i> . Springer. ISSN: 1574-0218. JCR Impact factor: 1.014
2020	M. Luisa Díez Platas, Salvador Ros Muñoz, Elena González-Blanco, Pablo Ruiz Fabo, Elena Álvarez Mellado. Medieval Spanish (12th–15th centuries) Named Entity Recognition and attribute annotation system based on contextual information. <i>Journal of the Association for Information Science and Technology</i> . ISSN: 2330-1643. JCR Impact Factor: 2.41 [pdf]

- 2018 **Elena Álvarez Mellado** and Leticia Martín-Fuertes Moreno. *Aracne: estudio de la variación lingüística en la prensa española entre 1914 y 2014* (*Aracne: a study of the linguistic variation in the Spanish press between 1914 and 2014*). *Artnodes*, (22), Fundació per la Universitat Oberta de Catalunya. ISSN: 16955951. SJR: 0.125 [pdf]

BOOKS

- 2016 **Elena Álvarez Mellado**, *Anatomía de la Lengua*, (*The Anatomy of Language*), Larousse-Vox, ISBN 978-84-9974-216-8

Scholarships & Awards

- 2020 [Outstanding Corpus Thesis Award](#) (Institute for Corpus Research, Incheon National University).
 2020 Karen Spärck Jones Award for Outstanding Achievement in NLP (Brandeis University).
 2017 [Premio Nacional de Periodismo Miguel Delibes](#) (Asociación de Prensa de Valladolid).
 2017 LaCaixa Foundation scholarship to pursue studies in the US (Fundación LaCaixa).
 2009 First award at the *Arquímedes Contest for Young Researchers* (Ministry of Science, Spain).
 2005 Excellence scholarship (Madrid Regional Government).

Selected Talks & Seminars

- Nov 2020 *Lingüística computacional en español*, Universität Zürich.
 Jun 2018 *Nacimiento, vida y muerte de una metáfora*, Natural Language Processing & IR Group, UNED.
 Jun 2018 *Aplicaciones del PLN para en humanidades*, Digital Humanities Summer School, UNED.
 Mar 2018 *Radiografía lingüística de la prensa*, Jornadas de Estudios de Lingüística, Universidad de Alicante.
 Mar 2017 *La lematización de textos*, Seminario Filologías Digitales Hoy, Universidad de Extremadura.
 Nov 2016 *Lingu ¿qué? Compu ¿cómo?*, Universidad de Alicante.
 Jul 2015 *Curso de expresiones regulares para profesionales de humanidades*, coLenguando, Madrid.
 Mar 2015 *Vivir (de) la lengua*, Universitat Rovira i Virgili.
 Mar 2014 *Cacharreando con la lengua. Recursos informáticos para filología*, Universidad de Zaragoza.
 Jun 2014 *From 1856 to 2014: resurrecting the dictionary*, APIdays Mediterranea, Barcelona.
 Dec 2013 *Todo lo que un filólogo no sabe que sabe hacer*, grupo Psylex, Universidad de Zaragoza.
 Dec 2012 *Nuevas tecnologías aplicadas a la enseñanza de español*, Universidad de Alcalá de Henares.
 Nov 2011 *Lenguas y Tecnología*, XI Semana de la Ciencia de la Comunidad de Madrid.

Personal Projects

- 2020- [Observatorio Lázar](#)
 • An observatory of anglicism usage in the Spanish press.
 2018-2019 [Corpus of political speeches from 1937 to 2019](#)
 • Analysis, visualizations and Python interface of a corpus of Spanish political speeches.
 2009-2016 [AZRAEL: A-Z RECONOCEDOR AUTOMATICO DE ESPANOL](#)
 • An automatic language detector in Prolog based in the syllable structure.
 • Current supported languages: Spanish, French, Italian, Portuguese, Catalan, Latin and Basque.

Dissemination & Outreach

- 2018- Editorial board member and columnist at [Archiletras](#), magazine about language and Linguistics.
 2018- Scientific committee member for [MariCorners](#) conference on LGBT language and communication.
 2017- Columnist for Spanish national newspaper [elDiario.es](#) with a monthly column about Linguistics.
 2012-2015 Radio contributor at RNE (Spanish National Radio) on a weekly section about Linguistics.