Damian Yukio | Romero Diaz
Tucson, Arizona, USA, 85719 | damianiji@email.arizona.edu
https://linguistics.arizona.edu/user/damian-yukio-romero-diaz
https://spanish.arizona.edu/people/damianiji

Ph.D. candidate at the University of Arizona experienced in language analysis with a focus on corpus linguistics transitioning into computational linguistics and natural language processing. Interested in improving communication outcomes among humans and with computers.

Experience in group and individual projects in English, Spanish, and Portuguese both for academia and industry. Currently working on the automated classification of documents using machine learning. Keen, hardworking, and eager to learn. Passionate about computational linguistics and big data. Able to work under high-demand situations in several projects at one time. Excellent communication skills.

SKILLS

● **Collection and annotation of language data:** Creation of corpora and dictionaries | XML annotation | Use of regular expressions, FSAs and other tools for semi-automation of language annotation | Creation and implementation of workflows and annotation conventions for language corpus analysis
 annotation tools such as Salt and Prodigy, and corpus tools such as LancsBox and AntConc

● **Linguistic data analysis:** Use of APA and IPA conventions for spoken data transcription and analysis | Experience in morpho-syntactic, phonological and pragmatic analyses of Spanish and English | Familiarity with WordNet and semantic categories | Basic use of R software for statistical tests

● **Scripting languages & version control:** Perl & Python (mainly) | Bash | MATLAB | IBM's Annotation Query Language | SQL (basic prior knowledge) | Git & GitHub

● **Natural language processing and machine learning:** Python libraries for ML and NLP**:** spaCy, Natural Language Toolkit (NLTK), scikit-learn, PyTorch | Kaldi ASR (basic knowledge)

● **Natural language proficiency:** Spanish (Native and Language Expert for Professional Purposes) | English (C2/High proficiency in technical, academic, and everyday language) | Portuguese (B1/Intermediate conversational and written proficiency) | French (A1/Basic, studying at the moment).

● **Pedagogical expertise:** Collaborative learning | Online teaching and learning | Classroom diversity and inclusion | Community-based learning | Task-based learning (learn by doing) | Student-oriented facilitation

WORK & RESEARCH EXPERIENCE

Dissertation work, The University of Arizona, U.S.A.
Dr. Carmen King de Ramirez, *Dissertation committee chair*,   **07/2019 - today**
- Currently working on a theoretical and computational study of people's (instructors' and crowdsourced) perceptions of student emails.
- Creation of corpora, datasets, workflows, annotation conventions, experiments and evaluation (precision, recall, F1) for the supervised multilabel classification of documents using Natural Language Processing, as well as conventional linguistic analysis methods.

Department of Spanish and Portuguese, The University of Arizona, U.S.A.
Research assistant for Dr. Julieta Fernandez   **08/2017 - 12/2017**
- Designed workflows and conventions of two oral corpora of learners of Spanish.
- Transcribed and annotated oral language interviews using some of these conventions.

IBM Deutschland - Research and Development GmbH, Böblingen, Germany
Intern, Watson Analytics for Social Media   **05/2017 - 08/2017**

- Creation of vocabularies, grammars and other linguistic resources for Spanish and Portuguese.
- Computational analysis of Brazilian Portuguese data using rule-based approaches.

Management Information Systems Department, The University of Arizona, U.S.A.
Research assistant for Dr. David Kauchak, Pomona College   **08/2015 - 07/2016**
- Worked with a team of researchers from the University of Arizona and Pomona College.
- Performed language analysis and processing through rule-based approaches and database (MySQL) queries for text simplification on a project funded by the National Institute of Health (U.S.A.)

General Hospital Dr. Manuel Gea González, Mexico City, Mexico
Research assistant for Dr. Alejandra Auza Benavides   **08/2014 - 05/2015**
- Worked with a team of researchers in the creation of a screening test for developmental language disorder (also known as specific language impairment) for Mexican children.
- Conducted language interviews and performed standardized language and IQ tests on several schools in Mexico City and Querétaro.
- Transcribed speech samples using the Salt software and managed the project database (MS Excel).

EDUCATION

Ph.D. in Hispanic Linguistics
The University of Arizona, Tucson, Arizona, U.S.A.
Expected graduation: Spring, 2022, GPA: 3.922/4.0
Dissertation: Instructor Perceptions of College Student Emails: Success Factors in Spanish and English

M.S. in Human Language Technology
The University of Arizona, Tucson, Arizona, U.S.A.
Graduation date: Spring, 2020, GPA: 3.922/4.0

M.A. in Hispanic Linguistics
The University of Arizona, Tucson, Arizona, U.S.A.
Graduation date: Spring, 2014, GPA: 3.903/4.0
Outstanding Graduate Teaching Assistant Award, College of Humanities
Outstanding Graduate Student Award, Hispanic Alumni Association

B.A. in Linguistics and Spanish Teaching as a Second Language
Autonomous University of Querétaro (UAQ), Querétaro, México
Graduation date: Summer, 2012, Cumulative grade: 9.359/10
Sponsored study abroad program at Concordia University (Montreal, Canada) in Fall 2010 & Spring 2011

FEATURED PUBLIC PROJECTS ON GITHUB

Python exercises: https://github.com/damian-romero/python_exercises
I continuously develop Python exercises in Jupyter Notebooks for myself and other graduate students.

covid19_mx: https://github.com/damian-romero/covid19_mx
(Spanish) A simple Jupyter Notebook tutorial to explore some of the data for COVID-19 from the ECDC. An interactive version can be launched through the following link and clicking on `covid19mx_notebook.ipynb`:
https://mybinder.org/v2/gh/damian-romero/covid19_mx/dev_pandas

Cuarentena 2020: https://github.com/damian-romero/cuarentena_2020
(Spanish) Code and live coding video tutorial for writing a video game with Pygame.