

Damián Yukio Romero Díaz, M.A.

Department of Spanish and Portuguese
Department of Linguistics
University of Arizona
Tucson, AZ 85721

Pronouns: he/him/they
<https://d-romero.com>
damianiji@email.arizona.edu

EDUCATION

Ph.D., Hispanic Linguistics, University of Arizona, Expected: Spring, 2022

Dissertation:

Spanish and English Academic Emails:

Systematic Review, Corpus Collection, and Instructors' Perceptions

Committee:

Dr. Carmen King de Ramirez (Spanish & Portuguese, UArizona, USA)

Dr. Julieta Fernandez (Spanish & Portuguese, UArizona, USA)

Dr. Sandiway Fong (Linguistics, UArizona, USA)

Dr. Helena Montserrat Gomez Adorno (IIMAS, UNAM, Mexico)

Dissertation summary:

My dissertation is a theoretical and computational study of people's perceptions of student emails. A question I aim to answer is how perceptions of appropriateness, formality, imposition, and others are triggered by the linguistic, communicative, contextual, and interpersonal factors contained in those emails. I argue that the relationship between emails and perceptions provides a measure of pragmatic competence. A second question I aim to answer is what linguistic and extralinguistic factors help students achieve their intent as expressed in their emails?

M.S., Human Language Technology, University of Arizona, Expected: Fall, 2021

M.A., Hispanic Linguistics, University of Arizona, 2014

B.A., Linguistics and Spanish Teaching as a Second Language, Autonomous University of Querétaro, México, 2012

TEACHING INTERESTS

Corpus linguistics, natural language processing, pragmatics, computational methods for the humanities and social sciences, online learning and teaching, instructional technology, language pedagogy, general linguistics courses, Spanish as a second language (all levels and formats), Spanish for heritage learners (all levels and formats), Portuguese (beginner levels), languages for specific purposes (including professional and academic purposes).

TEACHING EXPERIENCE

Spanish teacher as a Graduate Teacher Assistant (main instructor position in all cases):
The University of Arizona, Spring 2014 – Summer, 2021

SPAN 205 and 251 levels at the basic language program

SPAN 102, 202 and 251 at the Spanish Online program

Design and implementation of SPAN 425 – Advanced Grammar and Composition
in the Spanish Online program.

Creation, design and implementation of the Spanish Club Online: The University of
Arizona, Fall 2018 –Today

Spanish tutor in the study abroad program *Arizona in Costa Rica, Session 1*. Summer
2018. My teaching duties included:

Language tutoring: basic, intermediate, and advanced Spanish.

A culture class on salsa dancing and the Spanish language.

Spanish teacher as a Graduate Teacher Assistant (main instructor position in all cases):
The University of Arizona, Fall 2012 – Spring, 2014

SPAN 102, 201, 202, 206, and 325 levels at the basic language program

SPAN 333 at the Heritage Language program

English teacher: Autonomous University of Queretaro, Qro., MX, Spring 2011, Fall 2014
& Spring 2015

Basic and intermediate levels for the M.A. in literature

Intermediate and advanced levels for B.A. students at the department of natural
sciences

Spanish teacher: Instituto Tecnológico De Monterrey, Qro., MX, Fall 2014

Basic 1 and intermediate 1 level, as well as Communication Skills # 3

I also created the Spanish Club of this institution.

Spanish Teacher: Autonomous University of Queretaro, Qro., MX, Fall 2011 - Spring 2012
Individual and group teaching of several levels for foreign students using either
the communicative or the TPRS approach.

PUBLICATIONS

Romero Díaz, D. Y., Wang, H., Xu, W., and Jia, H. (*forthcoming*). Design and collection
challenges of building an academic email corpus for linguistics and computational
research.

- Romero Díaz, D. Y., Xu, W., Viale, M. L., Wang, H., Su, P., and Jia, H. (*submitted*). The Early Stages of the Multilingual College Email Corpus (MCEC): A Resource for Researchers and Language Instructors. *Proceedings of the LAEL-PG 2021*. Lancaster University.
- Harley, Heidi B; Romero Diaz, Damian Yukio; Whitney, Rebecca M; Ni, Tianyi; Yu, Jianrong; Hall Hartley, Lucy Jane; et al. (2021): *39th West Coast Conference on Formal Linguistics*. University of Arizona Research Data Repository. Collection. <https://doi.org/10.25422/azu.data.c.5325401.v2>
- Nitschke, Remo, Damian Y. Romero Diaz, Gabriela De La Cruz Sánchez, John W. W. Powell, Kristina Mihajlović, Luis A. Irizarry-Figueroa, George-Michael Pescaru, Florian Hafner (2021). *Coyote Papers 23: Proceedings of the Arizona Linguistics Circle 14*. Arizona: University of Arizona Linguistics Circle. <http://hdl.handle.net/10150/659844>
- Nitschke, Remo, Damian Y. Romero Diaz, Gabriela De La Cruz Sánchez, John W. W. Powell (2020). *Coyote Papers 22: Proceedings of the Arizona Linguistics Circle 13*. Arizona: University of Arizona Linguistics Circle. <http://hdl.handle.net/10150/641480>
- Mukherjee, P; Gondy Leroy; David Kauchak; Srinidhi Rajnarayanan; Damian R. Diaz; Nicole P Yuan; Gail Pritchard; Sonia Colina (2017). NegAIT: A New Parser for Medical Text Simplification Using Morphological, Sentential and Double Negation. *Journal of Biomedical Informatics*.
- Mukherjee, P; Gondy Leroy; David Kauchak; B. A. Navarrete; Damian R. Diaz; S. Colina (2017). The Role of Surface, Semantic and Grammatical Features on Simplification of Spanish Medical Texts: A User Study. *AMIA Fall Symposium*, Washington DC.
- Romero Díaz, Damián Y. (2013) "Eufemismos y tabú en las expresiones de la cultura mexicana con respecto a la muerte." *Divergencias. Revista de estudios lingüísticos y literarios*. Vol. 11, 1.
- Romero Díaz, Damián Y. (2012) "Creacionismo y deshumanización: los sentimientos en el jardín de "Altazor". *Divergencias. Revista de estudios lingüísticos y literarios*. Vol. 10, 2.

CONFERENCE PRESENTATIONS & WORKSHOPS

- "The Early Stages of the Multilingual College Email Corpus (MCEC): A Resource for Researchers and Language Instructors" Damian Yukio Romero Diaz, Wei Xu, Maria Laura Viale, Hui Wang, Peiwen Su & Hanyu Jia. Presentation at the Lancaster Postgraduate in Linguistics and Language Teaching (LAEL PG), 2021. Lancaster University, July 23rd-24th, 2021. <https://drive.google.com/file/d/1xeTk8bLj0j6lWrAMl0zRT6xS60Ek-IaY/view>
- "Basics of text processing with spaCy" Damian Yukio Romero Diaz, Workshop presented at the Arizona Research Bazaar 2021. The University of Arizona, May 17th-21st, 2021. <https://researchbazaar.arizona.edu/resbaz/resbazTucson2021/>
- "Helping online students cope with confusion, uncertainty, and anxiety through institutional values" Damian Yukio Romero Diaz, Presentation at the L2 Digital

Literacies Symposium (L2DL): Critical Transnational Dialogue and Virtual Exchange. Panel 1: Online Education during Times of Crisis: Challenges of Teaching during COVID-19. The University of Arizona, October 19th-24th, 2020. https://whova.com/embedded/speaker/ldicc_202010/12692728

- “Multilingual College Email Corpus Project” Hanyu Jia, Mohammed K Alshakhori, & Damian Yukio Romero Díaz, Presentation at the 19th Annual Second Language Acquisition and Teaching Interdisciplinary Roundtable. The University of Arizona, February 7th-8th, 2020.
- “The importance of studying online instructor/student relationships: A linguist's perspective” Damian Yukio Romero Díaz, Online education workshop at the Department of Spanish and Portuguese. The University of Arizona, 2019
- “Impositives on Student Emails to an Instructor: An Exploratory Analysis” Damian Yukio Romero Díaz, Poster presentation at the 18th Annual Second Language Acquisition and Teaching Interdisciplinary Roundtable. The University of Arizona, 2019
- “Multidimensional Analysis for Social Media Analytics” Damian Yukio Romero Díaz, *CoLi group colloquium of Watson Social Media Analytics*, IBM- Deutschland, 2017
- “Use of English in Sociolinguistic Interviews Among Bilinguals and Accommodation to Interviewer's L1” Courtney Cannon and Damian Yukio Romero Díaz, *26th annual symposium on Hispanic and Luso-Brazilian Literature, Language and Culture*, The University of Arizona, 2016
- “The Need for Variation: Rhotics in Mexican Spanish Male Gay Speech” Damian Yukio Romero Díaz, *26th annual symposium on Hispanic and Luso-Brazilian Literature, Language and Culture*, The University of Arizona, 2016
- “Taboo and Euphemisms in the Expressions of Mexican Culture with Regard to the Death” Damian Yukio Romero Díaz, *23th annual symposium on Hispanic and Luso-Brazilian Literature, Language and Culture*, The University of Arizona, 2013
- “What Was the Flood for? The Complaint in the Meta-Referential Function in Alejo Carpentier's *Los Advertidos*” Damian Yukio Romero Diaz, *IX ENELL (National Conference of Students of Linguistics and Literature)*, Autonomous University of Queretaro, 2011
- Poetry Reading: “*Puerto de Entrada*”, Damian Yukio Romero Diaz, *IX ENELL (National Conference of Students of Linguistics and Literature)*, Universidad Autónoma de Querétaro, 2011
- “The Art of Living in Unamuno's San Maue! Bueno, Mártir”, Damian Yukio Romero Diaz, *VII ENELL (National Conference of Students of Linguistics and Literature)*, Autonomous University of Aguascalientes, 2009
- “Mexican Spanish From Mexico City in Cellphone Text Messages”, Damian Yukio Romero Diaz, *VI ENELL (National Conference of Students of Linguistics And Literature)*, Autonomous University of Mexico (UNAM), 2008

Poetry Reading: “Sulphur, Experimental Poetry Reading”, Damian Yukio Romero Diaz, *V ENELL (National Conference of Students of Linguistics And Literature*, Autonomous University of Yucatán, 2007

RESEARCH EXPERIENCE

Research assistant for Dr. Tracy Crane (College of Nursing) and Dr. Steven Bethard (School of Information), University of Arizona, Current

As a research assistant, I am tasked with using the data in the Lifestyle Intervention for Ovarian Cancer Enhanced Survival (LIVES) dataset to help improve outcomes of cancer prevention interventions for U.S. Latina ovarian cancer survivors. This research is a step forward in democratizing the tools created by the LIVES project by giving access to automated fidelity monitoring to Hispanics, the largest ethnic minority in the U.S. We expect to inspire further research in other minority languages and cultures.

NIH Grant Project #1R21CA256680-01: Using Natural Language Processing to Determine Predictors of Healthy Diet and Physical Activity Behavior Change in Ovarian Cancer Survivors

Principal Investigator, Multilingual Email Corpus Project, University of Arizona, Fall 2019 - Today

The Multilingual College Email Corpus Project (MCEC) is an ongoing collection of student and instructor emails across several departments from the University of Arizona. It is a first-of-its-kind project given its wide scope (several disciplines and academic levels) and the sociolinguistic information collected (L1, L2, linguistic experiences, heritage languages, and more). I initiated the MCEC as my Ph.D. dissertation project and, since then, it has grown in orders of magnitude. Along with my co-PI and co-director Hanyu Jia, I am currently leading a team of six graduate students from three different departments.

Research assistant for Dr. Benjamin R. Fraser (Department of Spanish and Portuguese), University of Arizona, Fall 2019

I served as a research assistant for the Journal of Urban Cultural Studies, where I worked with Dr. Fraser on finding new authors and content, as well as on several publications for the Urban Cultural Studies blog. I also contributed to the publication of two research articles and published a short-form article on urban linguistic landscapes.

Research assistant for Dr. Julieta Fernandez (Department of Spanish and Portuguese), University of Arizona, Fall 2017

I worked with Professor Julieta Fernandez in the creation of two oral corpora of learners of Spanish in study abroad contexts.

Language Technology Intern at IBM Deutschland - Research and Development GmbH, May - August 2017

I worked as an intern for the IBM division of Watson Analytics for Social Media in the automated analysis of social media content in Portuguese.

Research assistant for Dr. Gondy Leroy (Management Information Systems Department), University of Arizona, Fall 2015 - Summer 2016

I worked with Professor Gondy Leroy (University of Arizona), Professor David Kauchak (Pomona College) and a team of researchers on a project funded by the NIH (U.S.A.) for evidence-based readability and text simplification. Our objective was the improvement of health literacy in English and Spanish speakers within the U.S.A. My main duties are a combination of both Natural Language Processing (NLP) tasks and cognitive science-based research aimed to the discovery and quantification of linguistic features that help reduce text complexity and maximize text understanding.

NIH Grant Project #1R01LM011975-01A1: Evidence-based Strategy and Tool to Simplify Text for Patients and Consumers

Research assistant for Dr. Alejandra Auza Benavides (Hospital General Dr. Manuel Gea González), Mexico City, Fall 2014, Spring 2015, and Fall 2019

Dr. Alejandra Auza is one of the top psycholinguists in Mexico and an expert on developmental language disorder (DLD) and language and autism. As her assistant, I worked on two projects. The first was the creation of a DLD screening test¹ for Mexican children, where I conducted language interviews and performed standardized language and IQ tests on several schools in Mexico City and Querétaro. On this project I also transcribed the children's interviews using the SALT software. I also managed the project's database (excel only).

For the second project, I was in charge of the experimental design, data collection, and analysis of words and part-words in children with DLD and children with typical development. This project we aimed to test how derivational morphology could be an indicator of atypical linguistic development.

On Fall 2019, Dr. Auza and I started working on a project for the automatic detection of DLD from speech using Machine Learning. The objective is to automate the detection of DLD using smart-phone recordings of children's book narrations. The project has been stopped due to the pandemic of 2020 and we hope to continue our work once the funding for this type of projects restarts.

Student, Psycholinguistics and Laboratory Phonology lead by Dr. Miquel Simonet (Department of Spanish and Portuguese), University of Arizona

I have experience in experimental design and inferential statistics of speech data from my graduate courses with Dr. Miquel Simonet, a lead expert in the field of Spanish laboratory phonetics and language acquisition:

SPAN 583A Spanish Phonetics I, SPAN 583B Spanish Phonetics II & SPAN 583C Spanish Phonetics III

PROFESSIONAL DEVELOPMENT

¹ Auza, A., Murata, C. Márquez-Caraveo, M.E. & Morgan, G. (2018). Manual de la prueba de Tamiz para detectar Problemas del Lenguaje (TPL). México: Manual Moderno (pp. 70). ISBN: 978-607-448-6506

Copyright First Responders, instructed by Kyle K. Courtney, the copyright advisor at Harvard's Office for Scholarly Communication. Spring 2021

Webinar. Delivering Distance Learning: How to Implement Effective Technology Quickly (Virtual Machine deployment for education). Amazon Web Services. Spring 2020.

Workshop. Writing Effective Data Management Plans Workshop. The University of Arizona - Research Data Management Services Team. Spring 2020.

Workshop. Advanced PlayPosit. The University of Arizona - Instructional Technology Team, Digital Learning. Spring 2020.

Workshop. Best Practices for Research Software: US Research Software Sustainability Institute (USRSSI Winter School, December 17–19, 2019 eScience Institute at the University of Washington: <https://escience.washington.edu/>)

Workshop. Building Communities in Online Courses. University of Arizona - Department of Spanish and Portuguese, imparted by M.A. Heidi Coronado, Fall 2019

Seminar. Conversifi - <https://conversifi.com>, Fall 2019

Mini Course. Collaborative Learning. University of Arizona – Office of Instruction and Assessment (OIA), Summer 2019

Workshop. Graduate Writing Institute, University of Arizona – Writing Skills Improvement Program (College of Humanities), Summer 2019

Certificate. Leader in Classroom Diversity & Inclusion, University of Arizona - Office of Inclusion and Multicultural Engagement, Spring 2019

Boot Camp. GTA Online Course Design Boot Camp, University of Arizona - Office of Digital Learning, Spring 2019

MOOC. Corpus Linguistics: Method, Analysis, Interpretation, Lancaster University & FutureLearn, imparted by Tony McEnery & Vaclav Brezina, Fall 2019

Workshop. Zoom for Conference Training, University of Arizona - Department Linguistics, imparted by Gregory Anderson, Video Technology Support Analyst for University Information, Spring 2019

Webinar. Verbal immediacy in online courses. University of Arizona - Department of Spanish and Portuguese, imparted by Dr. King de Ramirez, Fall 2018

Workshop. Online Instructor Training. University of Arizona - Department of Spanish and Portuguese, imparted by Dr. King de Ramirez, Summer 2018

Certificate. Teaching Proficiency Through Storytelling and Reading, Blaine Ray workshops at Pittsburgh, P.A., 2011

ADMINISTRATIVE EXPERIENCE

Social Media Chair. Arizona Linguistics Circle 15 conference (ALC15), Remote Linguistics: Experimentation, Fieldwork and Collection Methods in an Uncertain World, University of Arizona, 2021

Organization Committee Lead Tech Coordinator and Research Data Repository Lead. 39th meeting of the West Coast Conference on Formal Linguistics (WCCFL), University of Arizona, 2021

Organization Committee Vice-chair, Technology Team Chair, Webmaster, Outreach & Social Media Chair, Presenter Communication Chair, and Abstract Review Co-Chair. Arizona Linguistics Circle 13 conference (ALC13): Research Across Linguistic Subfields, University of Arizona, 2019

Outreach coordinator for the student club Ritmos Latinos. University of Arizona, Fall, 2019.

Organization Committee Member. *Todos Somos Otros, Festival Internacional de la Diversidad Social*: International Short-Film Festival, Mexico City- University of Arizona, 2014-2015

Director. DSP Film Festival, Department of Spanish and Portuguese, University of Arizona, 2017

EDITORIAL EXPERIENCE

Associate Editor / Secretary / Digital Archivist / Sustainability Coordinator: Coyote Papers, University of Arizona, 2020 - 2021.

Associate Editor / Webmaster / Author Communications: Coyote Papers, University of Arizona, 2019 - 2020.

Reviewer: Arizona Linguistics Circle 9 (ALC9), University of Arizona, 2015.

Reviewer: Divergencias: Magazine of Literary and Linguistic Studies, University of Arizona, Fall, 2013.

Editorial Assistant: Arizona Journal of Hispanic Studies, University of Arizona, Fall, 2012 & Spring, 2013.

Reviewer: Winter Edition, Divergencias: Magazine of Literary and Linguistic Studies, University of Arizona, 2012.

OTHER WORKING & PROFESSIONAL EXPERIENCE

Natural Language Processing Intern at ExplosionAI, GmbH (Remote), May 2021 - Today

Language transcription, annotation, and analysis. Contractor. Lionbridge Technologies, Inc. Spring 2019 - Today.

Software development (backend), data curation, and campus outreach. University Libraries - Office of Digital Innovation and Stewardship, Fall 2020 & Spring 2021.

Professional Mentoring and Course Evaluation of Online Courses (Quality Matters), University of Arizona – SPAN 201, Department of Spanish and Portuguese, Spring 2020.

Volunteer Salsa Instructor, University of Arizona – Ritmos Latinos Student Club, Summer 2019, Fall 2020.

Assistant to the Director for the study abroad program "Arizona in Costa Rica" (Session 1). Summer, 2018.

Translation agency entrepreneur- Summer, 2010 – Today

A few years ago, I started my own business in Mexico City; a translation agency and I have worked as a freelance translator over the years. The most recent translation work I did was for a company named MotorMexico on 2018:
<https://www.motormexico.com/spanish/>

Meditation Teacher, Sivana Yoga, Qro., Mx and other Meditation Centers, 2009 – 2014

Customer Service and Special Sales: Librerías Gandhi, Qro., Mx., 2007-2010

SERVICE

Volunteer for the Arizona Linguistics Circle 14 (ALC14), Department of Linguistics, The University of Arizona. Fall 2020.

Technology & Live Streaming Chair for the Human Language Technology Homecoming, Department of Linguistics, The University of Arizona. Spring 2020.

Assistant to the site coordinator / proctor for the North American Computational Linguistics Open Competition 2020, Department of Linguistics, The University of Arizona. Spring 2020.

Technology & Live Streaming Assistant for the Second North American Conference in Iranian Linguistics (NACIL2), Department of Linguistics, The University of Arizona. Spring 2019.

Technology Assistant for the Arizona Linguistics Circle 12 (ALC12), Department of Linguistics, The University of Arizona. Fall 2018.

Recruiter: Study abroad programs, Department of Spanish and Portuguese, The University of Arizona. 2016, 2017, & 2018.

Volunteer: Hacking Marathon “Hackathon”, The University of Arizona. Spring 2016.

M.A. Student Representative: Department of Spanish And Portuguese, The University of Arizona. Fall 2012 & Spring 2013.

GRANTS, FELLOSHIPS & SCHOLARSHIPS

College of Humanities Graduate Research Grant, The University of Arizona, 2020

Department of Spanish and Portuguese Graduate Fellowship, The University of Arizona, 2020

Graduate and Professional Student Council - GPSC Travel Grants - February 2020 Round

Digital Humanities Summer Institute (DHSI) Scholarship, University of Victoria, Canada, 2020

College of Humanities Fellowship, The University of Arizona, 2016

College of Humanities Fellowship, The University of Arizona, 2012

FOPER Grant, Autonomous University of Queretaro, 2012

PRONABES Fellowship, Autonomous University of Queretaro, 2010

Study Abroad Grant, Autonomous University of Queretaro, 2010

PRONABES Fellowship, Autonomous University of Queretaro, 2008-2009

HONORS AND AWARDS

Outstanding Graduate Teaching Assistant, College of Humanities, University of Arizona, 2014

Outstanding Graduate Teaching Assistant, Department of Spanish and Portuguese, University of Arizona, 2014

Outstanding Graduate Student Award, Hispanic Alumni Association, University of Arizona, 2014

TECHNICAL EXPERTISE

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 LancesBox 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 statistics

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

LANGUAGE PROFICIENCY

Spanish, Native & Language Expert for Professional Purposes (e.g. Creation of language resources for automated voice assistants, lexicon databases, etc.)

English, C2

Portuguese, B1

PROFESSIONAL MEMBERSHIPS

Copyright First Responders Arizona

Learner Corpus Association

Association for Computers and the Humanities (ACH)

National Postdoctoral Association