William Morales

Linguistics & NLP

wsmbuis5224@gmail.com | wmoral2@lsu.edu | github.com/WilliamMorales1 | williammorales1.github.io

---

Variationist sociolinguistics and quantitative corpus methods, with a focus on Spanish dialects and morphosyntactic change. Phonetics and speech perception, particularly dialectal effects on listener judgments. Language documentation and revitalization for endangered languages, especially Louisiana Creole (Kouri-Vini). Computational approaches to linguistic analysis, including what speech models encode about phonology and NLP tooling for under-resourced languages.

EDUCATION

Louisiana State University — B.Sc. Computer Science, Concentration in Data Science Fall 2023 – present

Minors: Linguistics; Spanish · College Honors (Ogden Honors College) · GPA: 4.10/4.0 · Baton Rouge, LA

- **Notable coursework:** CSC 4700 Special Topics in AI; CSC 4402 Database Systems; CSC 4330 Software Systems Development; LING 3150 Phonetics; LING 4750 Independent Research in Speech Science; SPAN 4005 Structure of the Spanish Language; SPAN 4917 Independent Research in Spanish Linguistics

RESEARCH EXPERIENCE

A Variationist Study of Spanish Intensifiers Fall 2023 – present

Mentor: Dr. Rafael Orozco · Louisiana State University

- Variationist analysis of 7,835 tokens across Medellín, Tunja (Colombia), and Louisiana Puerto Rican communities; examined periphrastic vs.
  morphological intensification conditioned by gender, age, education, and adjective type
- Developed a custom Python algorithm (TagAnt + manual verification) to extract and classify intensifier tokens from dialect transcripts
- Published in *Proceedings of the Linguistic Society of America* 11(1). doi:10.3765/plsa.v11i1.6091

Dialectal Variation and Perception of English Interdental Fricatives and Lateral Liquids Spring 2024 – present

Mentors: Dr. Hyunju Chung; Dr. Irina Shport · Louisiana State University

- Investigated whether sociophonetic stereotypes influence listener identification of ethnolect (AAE vs. SWE) and perception of /θ, ð/ via forced-choice and Visual Analog Scale tasks with 108 listeners
- Categorized stimuli (stopping and labialization variants), built Qualtrics survey, and contributed to Praat acoustic analysis
- Manuscript submitted to *Journal of Speech, Language, and Hearing Research*; currently under review

Honors Thesis: Phonological Feature Retention in Distilled Speech Models Fall 2026 – present

Mentor: Dr. Keith Mills · Louisiana State University

- Tests whether the linear phonological feature structure found in full-size self-supervised speech models (HuBERT, wav2vec2, WavLM) survives distillation, and which features degrade first

- Built a configuration-driven probing pipeline: per-layer hidden-state extraction, phone-level pooling over forced-aligned boundaries, and per-contrast scoring; PanPhon and Phoible feature backends are interchangeable without code changes
- Corrected an evaluation flaw in the standard approach, where cross-validation folds that are not phone-disjoint let a probe score by recognizing the phone rather than the feature

CONFERENCE PRESENTATIONS

2026

- **LSA Annual Meeting** — Tracking a Linguistic Innovation: A Sociolinguistic Investigation of Intensifiers in Three Spanish-speaking Communities. New Orleans, LA. January 2026.
- **Southeastern Conference on Linguistics (SECOL)** — A Variationist Study of Spanish Intensifiers. Atlanta, GA. March 2026.
- **13th Annual LSU Discover Day** — Tracking a Linguistic Innovation: A Sociolinguistic Investigation of Intensifiers in Three Spanish-speaking Communities. Baton Rouge, LA. April 2026.
- **XXI Congreso de la Asociación de Lingüística y Filología de América Latina (ALFAL)** — Una investigación variacionista de la intensificación en el español colombiano. Lima, Peru. August 2026.
- **2026 American Speech-Language-Hearing Association (ASHA) Convention** — Effect of Dialectal Variation on Listeners' Perception of English Interdental Fricatives and Liquids. Virtual. November 2026.

2025

- **188th Meeting of the Acoustical Society of America (ASA) / 25th International Congress on Acoustics** — Effect of Dialectal Variation on Listeners' Perception of English Interdental Fricatives and Lateral Liquid Sounds. New Orleans, LA. May 2025.
- **12th Annual LSU Discover Day** — A Sociolinguistic Study of Intensification in the Spanish of Louisianian Puerto Ricans. Baton Rouge, LA. April 2025.
- **Fifth Annual Emory Undergraduate Linguistics Conference (EULC5)** — A Variationist Analysis of Intensification in Spanish. Virtual. April 2025.
- **Seventh Annual Tulane Conference on Linguistics (T-CoL)** — A Sociolinguistic Study of Spanish Intensifiers across Two Dialects. New Orleans, LA. March 2025.
- **Department of Communication Sciences and Disorders Roundtable, LSU** — Effect of Dialectal Variation on Listeners' Perception of English Interdental Fricatives and Lateral Liquid Sounds. Baton Rouge, LA. 2025.

2024

- **Summer Undergraduate Research Forum, LSU** — A Sociolinguistic Study of Intensification in the Spanish of Louisianian Puerto Ricans. Baton Rouge, LA. August 2024.

SELECTED PROJECTS

Linguistics & NLP

Variationist Intensifier Corpus Pipeline github.com/WilliamMorales1/IntensifierResearch

Python · pandas · R

- Processes POS-tagged sociolinguistic interview transcripts across 6 Spanish dialect corpora (Cali, Cartagena, Medellín, Tunja, Louisiana Puerto Rican, and historical) to extract intensifier tokens with adjective/adverb hosts

- Classifies predication type (predicative vs. attributive), resolves apocope, handles -ísimo superlatives and archi- prefixation, detects double-intensifier constructions, and identifies proximate verb
- Merges token-level output with speaker metadata (age, sex, dialect) for quantitative analysis; R and Python analysis layer produces rate-by-age plots, Zipf distributions, log-adjusted frequency graphs, and Bayesian models
- Used in multiple conference presentations across national and regional venues; published in *Proceedings of the Linguistic Society of America* 11(1). doi:10.3765/plsa.v11i1.6091

Louisiana Creole Online Dictionary creole-dictionary.fly.dev · github.com/WilliamMorales1/louisiana-creole-online-dictionary

Python · Django · SQLite · Fly.io

- Searchable web edition of the Valdman dictionary of Louisiana Creole (Kouri-Vini): 5,139 entries, 6,519 senses, and 5,055 spelling variants, with accent-insensitive and whole-word search; built with the LSU Creole Club

Praat Data Analysis Scripts github.com/WilliamMorales1/PraatDataAnalysis

Praat scripting language

- Batch WAV processing and automated acoustic measurements for a study on dialectal variation and perception of English sounds

Software & App Development

Multilingual Wordle multilingual-wordle.fly.dev · github.com/WilliamMorales1/Multilingual-Wordle

Go · TypeScript

- Wiktionary-based Wordle supporting 70+ languages, including scripts with grapheme clusters, abugidas, and RTL text; word lists cached locally after first download

Vivan — Language Learning Platform github.com/claytakiler/Heritage-Project

Django · React · Tailwind · ShadCN

- Language-learning platform for predefined and user-created courses; originally built for Louisiana Creole (Kouri-Vini); supports course creation, collaboration, and 2FA; built as a CSC 4330 team project

Natural Syntax LSP github.com/WilliamMorales1/Natural-Syntax-LSP

Go · ONNX Runtime · Python · TypeScript (VS Code Extension)

- Parts-of-speech, semantic-embedding, or dependency-parse highlighting for VS Code via a local Go LSP server running ONNX inference (BERT/MobileBERT, MPNet/MiniLM, or a biaffine UD parser); hover any word for its tag, head/dependents, and a Wiktionary definition

GuideOverride github.com/CSC4700-BRCC/Version-2

Django · SQLite · React · OpenAI · Tailwind

- Collaboration with Baton Rouge Community College to streamline student course-override requests using an LLM recommendation workflow
- Featured: Business Report — LSU AI Showcase

LEADERSHIP AND ORGANIZATIONS

President — Student Linguistics Association (SLA), LSU 2025 – present

Vice-President — Creole Club (Organizasyon Langaj é Lakilchi Kréyol Lalwizyàn), LSU 2025 – present

Featured in LSU Reveille

Vice-President — Student Linguistics Association (SLA), LSU 2024 – 2025

Historian — French Club (Le Cercle Français), LSU 2024 – 2025

Event Coordinator — Arabic Language Club, LSU 2024 – 2025

OTHER ACTIVITIES

Hackathon (WICS & SASE) Fall 2025

github.com/g34rmeister/2r-in-strawberry

Hackathon (SASE) Spring 2024

Jill Brody's Linguistic Meetup Day — Student Linguistics Association, LSU Spring 2023

- Student-organized event to introduce undergraduates to linguistics research and encourage participation in the field

AWARDS AND FUNDING

Hispanic Scholarship Fund Fall 2026 - Spring 2027

Baker Hughes Endowed Scholarship Fall 2026

Adrian Virginia Lazarus Memorial Scholarship Spring 2026

Gulf Scholars Program, LSU Summer 2024

President's Honor Roll Spring 2024, Fall 2024, Fall 2025, Spring 2026

Honors TOPS Scholarship Fall 2023 – present

President's Student Aid Fall 2023 – present

SKILLS AND LANGUAGES

Programming Languages: Python; SQL; Go; Java; R; TypeScript/JavaScript
Libraries & Frameworks: pandas; numpy; PyTorch; Hugging Face Transformers; Django; React; SpaCy; SQLite; PostgreSQL; Ollama
Domain Tools: Corpus linguistics methods; Praat scripting; acoustic analysis; IPA
Natural Languages: English (Native); Spanish (Proficient); French (Intermediate); Japanese (Intermediate); Chinese (Intermediate)