ALEJANDRO CIUBA

<u>Address</u>: 8 Riverside Drive, Hawley, PA 18428

<u>Phone</u>: (412) 874-1838 ・ <u>Email</u>: alejandrociuba@gmail.com

<u>LinkedIn</u>: [alejandro-ciuba](#) ・ <u>Github</u>: [AlejandroCiuba](#)

Professional Profile

I am a 4th year undergraduate student pursuing a major in computer science with minors in Spanish and linguistics. I possess a wide range of computer science skills ranging from experience with various popular languages like Java, Python and C, to familiarity with useful tools/applications such as the Selenium IDE and the Git/Github ecosystem. I am also comfortable in many OS environments, having years of experience using Windows, Ubuntu, and MacOS. In regards to linguistics, I have taken a wide range of courses which has given me a breadth of knowledge in the field. My main focuses include: natural-language processing, corpus linguistics and internet sociolinguistics. I have also done four research projects: *[Pragmatics in Video Game Dialogue](#)*, *[JSALT 2022 Summer Workshop](#)*, organizing the *[PELIC Speech Dataset](#)*, and transcribing Pittsburgh speech for *APLS*.

Education

University of Pittsburgh, Pittsburgh, PA　　　　　　　　　　　Anticipated Graduation Month: April 2023
Bachelor of Science in Computer Science　　　　　　　　　　　　　　　　　　　　　GPA 3.97
Minors: Spanish, Linguistics

Relevant Coursework for Computer Science: Data Structures and Algorithms I & II, Systems Software, Computer Organization & Assembly Language, Introduction to Computer Graphics, Software Quality Assurance, Introduction to Operating Systems

Relevant Coursework for Mathematics: Calculus I & II, Introduction to Statistics, Discrete Structures, Computational Methods

Relevant Coursework for Linguistics: Introduction to Linguistics, Introduction to Sociolinguistics, Introduction to Computational Linguistics, Introduction to Translation Theory[*], Data Science for Linguists, Overview of Spanish Linguistics[*]

[*]Course taught in Spanish

Experience and Personal Projects

Linguistics Data Science Research – Audio Transcriber, Pittsburgh, PA　　　　August 2022 – Present
Research Assistant for Dr. Dan Villareal
- Transcribes audio recordings between field researchers and interviewees using ELAN
- Follows transcription guidelines to clean-up previous transcripts
- Experiments with using CLOx speech-to-text to speed up transcription process for multi-speaker audio

Teaching Assistant – Computational Linguistics, Pittsburgh, PA　　　　August 2022 – Present
Teaching Assistant to Professors Na-Rae Han at the University of Pittsburgh
- Assists Dr. Han in managing in-class exercises
- Hosts office hours for students to ask questions about concepts, homeworks and exercises
- Helps give constructive feedback during the grading process by analyzing student code submissions
- Gave a guest lecture on machine learning and translation with Jupyter notebook demo

[Frederick Jelinek Memorial Summer Workshop](#), Baltimore, MD　　　　June 2022 – August 2022
Undergraduate Student and Researcher – Machine Translation for Low Resource Languages
- Undergraduate researcher under Dr. Anthony Larcher of Le Mans Université
- Participated in the Summer school portion of the workshop; introduced to a variety of NLP technologies and concepts: LabSE, Fairseq, xVector, SpeechBrain, etc.
- Created a trilingual corpus featuring a low resource language (English-French-Tamasheq/Tamahaq) used in training and testing
- Ran models (such as T5-Small and mBART) on the Johns Hopkins University Compute Cluster
- Read numerous academic papers and publishings about the latest NLP technology

Linguistics Data Science Research – Pragmatics, Pittsburgh, PA   February 2022 – April 2022

Research guided by Dr. Na-Rae Han for Spring 2022 "Data Science for Linguists" Class
- Gathered data from previously establish Github repository and reformatted for research use
- Extracted raw data from games using custom-parsing scripts
- Applied pragmatics knowledge onto video-game dialogue: formalized speech act phrases into regular expressions and Matcher patterns, named-entity and hapax examination, and pronoun frequency analysis

Linguistics Data Science Research – PELIC, Pittsburgh, PA   January 2022 – April 2022

Research Assistant for Dr. Alan Juffs and Dr. Na-Rae Han
- Aided in the clean-up of the unpublished *PELIC Speech Dataset*
- Organized the data through searching the organizations archives and connected data points together
- Used Python to create and run scripts/programs to better manage the dataset

Teaching Assistant – Algorithms & Data Structures I, Pittsburgh, PA   August 2021 – April 2022

Teaching Assistant to Professors John Ramirez at the University of Pittsburgh
- Instructed recitations with upwards of 20 students on algorithm and data structure concepts
- Managed exercises given by Dr. John Ramirez and walked through solutions
- Assisted students one-on-one during office hours

Provost Academy Leader, Pittsburgh, PA   May 2020 – Present

Councilor in The Provost Academy Program
- Helps plan and monitor activities virtually and in-person for incoming first-year students
- Advises and helps first-year college students in scheduling and resident-life
- Maintains connections with first-year students throughout the year via wellness checks

Flatland: The Game, Pittsburgh, PA   September 2020 – November 2020

CSC Hacks 2020 Winner + Best Beginner Project
- Created a 2D arcade-style shooter with retro effects for CSC Hacks 2020
- Used the Unity Game Engine and C# to program the game in 6 weeks
- Made all assets in Aseprite and used LMMS to make the game's music

Carnegie Library ESL Volunteer Instructor, Pittsburgh, PA   August 2019 – March 2020

Volunteer Teaching Assistant for Carnegie Public Library ESL Class
- Assisted main instructor with lesson plans
- Guided adult learners through vocabulary and grammar exercises
- Worked one-on-one with students to ensure they were able to practice their English in a safe environment

Skills

Spanish
- Conversational in Spanish
- Has translated technical math documents into the language
- Tutored math and computer science in the language

Programming
- Languages: Python, C# (Unity Game Engine), Java, and C with Makefiles and Bash Scripting
- Python Experience: PyTorch, SKLearn, spaCy, Hugging Face API, NLTK, Pandas, Numpy, Matplotlib, etc. through scripts, shell and Jupyter Notebook environments
- C Experience: OpenGL class projects, low-level operating system programming class projects
- Java Experience: JUnit, Gherkin + Cucumber, through both the Eclipse IDE and terminal

Mathematics
- Knowledge of Statistics, both descriptive and inferential
- Knowledge of Calculus: derivatives, integrals, differential equations, multivariable problems, etc.
- Knowledge of Linear Algebra through computer graphics: matrices/vectors and their operations
- Knowledge of Formal Language Theory: Automatons & Grammars, Chomsky Hierarchy, Complexity Classes, etc.

Other Skills

- Experience with Git and Github in individual, team and organization settings
- Experience with cluster computing: Sun Grid Engine and Slurm
- Experience with VMWare and OpenVM
- Experience using ELAN for audio recording transcription
- Currently uses a Linux environments (Ubuntu) as daily operating system
- Knowledgeable with LaTeX using Texmaker

Hobbies and Interests

University of Pittsburgh Spanish Club, *President, Vice President & Mesa Coordinator*	August 2019 – Present
University of Pittsburgh Computer Science Club, *Member*	October 2020 – Present
PyLing – Python Linguistics Club, *Member*	August 2021 – Present
Interest in Natural-Language Processing	August 2021 – Present
Interest in Computer Graphics and Game-Design	August 2021 – Present