Anuk Centellas

anukzc@gmail.com | anukzc.github.io | linkedin.com/in/anuk-centellas

Computational linguist building LLM and NLP tools from evaluation pipelines to user-facing apps

Education

University of Washington Sept 2025 – June 2028 (expected)
MS in Computational Linguistics *4.0 GPA*

Western Washington University Sept 2021 – June 2025
BS in Mathematics/Computer Science, BA in Linguistics *4.0 GPA*
Honors Interdisciplinary Studies Minor; Outstanding Graduate in Linguistics

Coursework: Machine Learning, Deep Learning for NLP, Advanced Statistical Methods for NLP, Algorithms and Data Structures

Experience

LLM Distillation Research Intern Aug 2026 – Present
Applied Research Laboratory for Intelligence and Security (ARLIS) *University of Maryland*
- Built reusable evaluation tooling for distillation research that quickly evaluates current and future models
- Developed a configurable MATH-500 evaluation using Inspect AI; users pick difficulty levels, subjects, and number of problems
- Wrote a script compiling results from teacher, base, and student models into tables that compare them from multiple angles
- Shipped the tooling in days rather than weeks by directing Claude Code and Codex as coding agents

Research for Intelligence & Security Intern June 2026 – Aug 2026
LLM-Powered Content Collection Tool – National Foreign Language Center (NFLC) *University of Maryland*
- Built an LLM pipeline from scratch to collect lesson content in 7 languages, reducing manual collection time from hours to minutes
- Refactored codebase for optimal modularity and organization, ensuring future work can be integrated smoothly
- Presented updates to other interns and leaders throughout the internship and addressed their questions and feedback
- Implemented third-party APIs for LLM calling, web searching, text scraping, and audio transcription
- Applied in-context learning and prompt engineering to boost LLM performance on classification and query generation
- Incorporated output saving for each LLM call to provide a descriptive, human-readable record for accountability
- Created a Gradio UI with instructions for entering desired content specifications to maximize user-friendliness

Language Revitalization Tooling Research Feb 2026 – July 2026
Bridging Digital Tools for Linguistic Documentation and Revitalization *University of Washington*
- Collaborated with student researchers to create langlit, a language revitalization app that addresses previous versatility gaps
- Performed user testing and documentation to improve organization, built-in instruction, and overall user-friendliness of langlit
- Created comprehensive tutorials for langlit users of different backgrounds, including documentary linguists, teachers, and learners
- Co-authored a paper titled "Bridging Digital Tools for Linguistic Documentation and Revitalization" published in **AmericasNLP**

Language Revitalization Lab Feb 2025 – June 2026
Documentation, Analysis, and Revitalization in Linguistics (DARLing) – Dr. Maura O'Leary *Western Washington University*
- Analyzed and structured under-documented linguistic data for integration into a morphological transducer system
- Improved code consistency and correctness by identifying and correcting errors affecting downstream translation accuracy
- Collaborated with the team to determine representations for ambiguous verb morphology that prioritize the speaker community

Alpha Game Tester Dec 2025 – May 2026
Atlas Arena *Nexus Games*
- Collaborated with unfamiliar players and adjusted my role within the team during testing sessions to optimize team dynamic
- Adapted to a frequently changing schedule, communicated regularly to change session times and reorganized when needed
- Provided game play and interface feedback to game developers to maximize the product's impact on fostering human intelligence

Machine Learning Research – Team Lead April 2024 – June 2025
Hutch Research – Dr. Brian Hutchinson *Western Washington University*
- Collaborated with the Speech Pathology Department on an applied ML research project to transcribe audio of toddler speech
- Built an end-to-end data preprocessing pipeline in Python to transform noisy transcripts into structured, model-ready datasets
- Listened to toddler speech recordings to verify transcript and timestamp accuracy against the audio
- Independently researched and tested existing forced alignment and diarization models to identify components for our system
- Communicated weekly with non-technical stakeholders to ensure our project was maximally useful for speech pathologists
- Coordinated task assignments, communication, and collaborative research sessions to ensure project milestones were met

Projects

Vetology Data Science Training Sept 2026 – Present
Veterinary Radiology Report Analysis *Python, Excel, LangChain, Claude Code*

- Manually scored radiology reports by condition and reconciled labels with peers to build a shared gold-standard dataset
- Engineered LLM prompts to classify conditions in radiology reports, then compared model output against the gold standard
- Found LLMs over-labeled bronchitis by inferring it from lower airway disease; adding explicit-mention rules improved performance
- Built a Python app that exports a confusion matrix including metrics like sensitivity and specificity per condition to Excel
- Rebuilt the app with Claude Code, analyzing differences in approach, structure, and code quality

Data Formatting and Plotting App – Honors Senior Capstone Aug 2024 – June 2025
Language Learning Research Group (LLRG) *Python, Streamlit, pandas*

- Built and deployed a Streamlit web application enabling users to upload raw Excel data and generate structured outputs
- Collaborated with a non-technical team and users to design the app, and independently learned Streamlit to implement it
- Designed and implemented a backend data normalization pipeline using pandas to standardize inconsistent formatting
- Generated exportable visualizations and reformatted data used by the LLRG to improve the Spanish program
- Presented the app and design process to a diverse, non-technical audience as my Honors Senior Capstone Project

Phoneme Classifier Nov 2024
Deep/Machine Learning Project *Python, scikit-learn, WandB, GitHub*

- Trained and evaluated an ensemble model of twelve neural network classifiers on 1.7 million training instances
- Improved ensemble model performance by giving the top four performing NNs more votes
- Conducted extensive hyperparameter tuning and compared performance across model architectures to find the optimal model
- Achieved a validation accuracy of 0.456 and a test accuracy z-score of 1.340, the top performance for this task

Arbitrarily Deep Neural Network Oct 2024
Deep/Machine Learning Project *Python, NumPy, argparse, GitHub*

- Built a fully customizable neural network framework from scratch through pair programming
- Implemented support for both classification and regression tasks, minibatching, and an optional verbose mode
- Designed the neural network to accept data files and hyperparameters as command-line arguments for flexible experimentation
- Incorporated periodic evaluation on the development set with real-time performance reporting

Skills

Programming: Python, Java, SQL; NumPy, pandas, scikit-learn, Matplotlib, argparse; Streamlit, Gradio, Tkinter

AI & LLMs: LLM APIs, LangChain, Inspect AI; prompt engineering, in-context learning, LLM evaluation; Claude Code, Codex

Machine Learning: neural networks, ensemble models, hyperparameter tuning, experiment tracking (Weights & Biases), model evaluation (confusion matrices), data preprocessing

Speech & Linguistics: phonetics, phonology, morphology, syntax, semantics, pragmatics, sociolinguistics; IPA transcription, Praat; forced alignment, speaker diarization, morphological transducers

Annotation & Evaluation: annotation guidelines, gold-standard labeling, label reconciliation, rubric-based evaluation of AI outputs, written rating justifications

Tools: Git, GitHub, Linux/Unix, VS Code, Vim, LaTeX, Excel

Languages: English, Spanish, Catalan

Collaboration: user testing, requirement gathering, stakeholder communication, technical documentation and tutorials, presenting to technical and non-technical audiences