Tyler W. Carson

San Jose, CA · tyler.w.carson@gmail.com · (909) 756-5942
linkedin.com/in/twcarson · github.com/twcarson

Work Experience

Sensory, Inc. — Portland, OR (remote)
Product Linguist — May 2023 - May 2025
- Created, documented, and maintained internal software tools written in Python and Bash.
- Trained and tested artificial neural networks built for wake word and command recognition for major automotive and consumer electronics companies.
- Classified and categorized large language datasets for use in model training.
- Used research methods to find optimal model parameters and best practices for data curation.
- Balanced the work of several simultaneous projects with tight deadlines.

SoundHound, Inc. — Santa Clara, CA
Linguistic Engineer — Aug 2021 - Jan 2023
- Created experimental and production acoustic models for several languages, often delivering models on tight time-frames. Produced scripts to automate and streamline multi-step training processes.
- Lead meetings and coordinated presentations for a team of engineers and linguists.
- Worked in a small team to port a legacy repository from C to Python.

Language Data Specialist — Mar 2020 - Aug 2021
- Trained and debugged acoustic models and language models for automatic speech recognition.
- Developed and maintained internal company web tool for audio data collection (Javascript, HTML) and audio corpus database for speech models (MySQL, Python).

English Data Intern — Aug 2019 - Mar 2020
- Curated language datasets, annotated speech data in IPA and SAMPA, lead scripting lessons and office hours, and developed tools in Python for scraping and conditioning text data for use in a linguistic corpus.

UCLA Linguistics — Los Angeles, CA
Undergraduate Researcher — Jul 2017 - Oct 2018
- Developed software in OCaml for calculating the complexity of grammar in natural language sentences using concepts based in information-theory.

Projects

Analysis of Uyghur Vowel Harmony *PRAAT, Python, Excel*
- Measured phonetic qualities of consultant-provided audio data using PRAAT.
- Performed corpus study of written Uyghur analyzing properties of backenss harmony.

Web Scraper for Corpus Work *Python, Scrapy, Tkinter*
- Constructed a web scraper using Python's Scrapy module to generate a corpus of written Uyghur.
- Implemented corpus data processing tools including regex search tool and Arabic/Latin transliterator.

Part of Speech Tagger *Python, NLTK*
- Developed a part of speech tagger for untagged corpora using Python Natural Language Tool-Kit.
- Model training used the open-source maximum entropy algorithm MegaM.

Skills

Tools / Environments:	Git, UNIX, Jira, Docker, Google Workspace, Audacity, Excel, LaTeX
Programming Languages:	Python, Bash, SQL, C, C++
Natural Languages:	English (fluent, native)
	German, Spanish (basic proficiency)

Education

University of California, Los Angeles — Los Angeles, CA
M.A. Linguistics — Sep 2018 - Jun 2019
B.A. Linguistics and Computer Science — Sep 2015 - Jun 2019

Coursework

Phonetics · Phonology · Syntax · Semantics · Computational Linguistics · Programming Languages · Formal Languages & Automata Theory · Algorithms · Mathematical Structures in Language