

# TYLER W. CARSON

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

## WORK EXPERIENCE

---

### 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. Created 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 iterated on a web tool for audio data collection written in Javascript & HTML.
- Designed, set up, and maintained a MySQL database for cataloging text and audio data used in creating text to speech models.

#### English Data Intern

Aug 2019 - Mar 2020

- Organized and lead scripting lessons and office hours for other interns and team members.
- Developed a collection of 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

- Conducted research on the relationship between surprisal and entropy reduction as measures of complexity for natural language.
- Developed software in OCaml for calculating the complexity of grammar in natural language sentences using concepts based in information-theory.
- Contributed code for calculating the entropy of a probabilistic formal grammar.

## 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 backness harmony.

### Web Scraper for Corpus Work

Python, Scrapy, Tkinter

- Built 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 was done using the open-source maximum entropy algorithm MegaM.

### Text Based / Rogue-Like RPG

C++, NCurses

- Designed and implemented a game-engine for a terminal-based RPG using C++ and nCurses library.

## SKILLS

---

Tools / Environments:	Git, UNIX, Jira, Jenkins, Docker, Kubernetes, Audacity, Emacs, Excel, $\text{\LaTeX}$
Programming Languages:	Python, Bash, C, C++
Natural Languages:	English (fluent, native) German (working proficiency)

## EDUCATION

---

### University of California, Los Angeles

Los Angeles, CA

M.A. Linguistics GPA: 4.00

Sep 2018 - Jun 2019

B.A. Linguistics and Computer Science GPA: 3.69

Sep 2015 - Jun 2019

### Crafton Hills College

Yucaipa, CA

A.A. Japanese Language and Culture GPA: 3.91

Sep 2013 - Jun 2015

## COURSEWORK

---

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