

Roger Yu-Hsiang Lo

LinkedIn | GitHub | Google Scholar
+1-778-223-5208 | roger.lo@ubc.ca | b13472003@gmail.com

EDUCATION

U OF BRITISH COLUMBIA

PHD IN LINGUISTICS

Jun 2022 | Vancouver, BC

LEIDEN UNIVERSITY

MA IN LINGUISTICS

Aug 2016 | Leiden, The Netherlands
Cum Laude

NATIONAL TAIWAN U

BSc IN CHEMISTRY

MINOR IN LINGUISTICS

Jun 2013 | Taipei, Taiwan
GPA: 3.79 / 4.0

COURSEWORK

GRADUATE

Computational Linguistics
Computational Morphology
Syntactic Theory and Analysis
Semantic Theory and Analysis
Phonological Theory and Analysis
Phonetic Theory and Analysis
Advanced Machine Learning
Exploratory Data Analysis
Bayesian Analysis and Decision Theory
Information Visualization

UNDERGRADUATE

Machine Learning and Data Mining
Methods for Statistical Learning
Software Construction
Basic Algorithm and Data Structures
Intermediate Algorithm and Data Structures
Numerical Computation for Algebraic Problems
Introduction to Relational Databases

SKILLS

PROGRAMMING

Python (PyTorch, scikit-learn, NumPy, SciPy) • R • Javascript • Julia • Praat • Java • Stan • C++ • SQL • Linux • Git • \LaTeX

LANGUAGES

Mandarin: native
English: fluent (TOEFL: 113 / 120)
Dutch: intermediate
Japanese: intermediate

SUMMARY

A linguist that uses machine learning, Bayesian statistical modeling, NLP, and experimental/corpus methods to study and analyze natural language. 6-year programming experience, familiar with Red Hat Linux Server and Python/R data science stack.

EXPERIENCE

AMAZON | APPLIED SCIENTIST

Sep 2022 – Present | Vancouver, BC
• Alexa Daily Essentials

UNIVERSITY OF BRITISH COLUMBIA | RESEARCH ASSISTANT

Sep 2016 – Aug 2022 | Vancouver, BC

- Supported various research projects from Prof. Molly Babel, Prof. Kathleen Currie Hall, Prof. Márton Sósokuthy, and Prof. Anne-Michelle Tessier.

UNIVERSITY OF BRITISH COLUMBIA | TEACHING ASSISTANT

Sep 2019 – May 2022 | Vancouver, BC

- Supported teaching for the *Master of Data Science in Computational Linguistics* program: *Computational Morphology, Computational Semantics, Corpus Linguistics, NLP for Low-Resource Languages, Sentiment Analysis, Syntactic Parsing*.
- Guided 30+ students in weekly programming lab assignments.

DATA SCIENCE FOR SOCIAL GOOD | FELLOW

May 2020 – Aug 2020 | Vancouver, BC

- Processed traffic and demographic data from censuses and local governments with R and Python.
- Ran various statistical models, such as Gaussian Markov random fields and generalized additive models.
- Developed a **web-based traffic visualization tool** for the City of Surrey with R Shiny.

PROJECTS

XLING CORPUS Sep 2020 – Present

Worked with **Prof. Márton Sósokuthy** to create a **Laravel-based website**, a web interface for collecting speech from speakers of various languages. Annotated, forced-aligned, and analyzed speech data with Praat, Python, and R.

PHONOLOGICAL CORPUS TOOLS Sep 2016 – Present

Worked with **Stanley Nam** and **Prof. Kathleen Currie Hall** to improve functionality of **Phonological CorpusTools**, a software tool written in Python for various phonological analyses on transcribed corpora.

GRAPHEME-TO-PHONEME CONVERSION April 2021 – May 2021

Partnered with **Garrett Nicolai** to develop an LSTM-based sequence-to-sequence model that won the first place in the SIGMORPHON 2021 Grapheme-to-Phoneme Shared Task in the low-resource setting.

SELECTED PUBLICATIONS

- [1] R. Y.-H. Lo and G. Nicolai. Linguistic knowledge in multilingual grapheme-to-phoneme conversion. In *Proceedings of the Seventeenth SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, pages 131–140, 2021.

