GIANG LE

(+1) 217-250-1700 ⋄ gianghl2@illinois.edu

EMPLOYMENT EXPERIENCE

Analytical Linguistic Project Manager *Google*

June 2016 - July 2019

As Analytical Linguistic Project Manager at Google, I led Vietnamese syntax annotation, segmentation annotation, message development, and linguistic evaluation projects. Since Q2 2017, I worked on cross-linguistic projects and became co-owner of a major grammar domain for APAC languages at the end of Q4 2017. I created tools and workflows for data extraction and manipulation to speed up the process of grammar induction. I also worked on two lexicon building efforts, specifically I was the main contributor to pronunciation data analysis, function words/mwe extraction, mapper for data conversion of Japanese Unidic. I also worked on data collection for i18n lexical resources to improve grammar quality.

Own and lead CFG rules development for a major domain for APAC region
Lead a working group to develop guidelines for message generation
Lead Vietnamese syntax annotation project
Lead Vietnamese message generation project
Design the workflow for grammar data processing
Develop tools for data processing using Python
Manage Vietnamese Treebanks, creating annotation guidelines for dependency relations, semantic, syntactic, and morphological projects
Lead various Vietnamese NLU, NLG projects for the Assistant
Train, manage, and oversee the work of a team of junior linguists
Evaluate and analyze data quality, consult with researchers and engineers on the development of linguistic databases
Train and/or evaluate NLP or dialogue systems, identify errors/regressions and propose solutions
Monitor data acquisition and curation; design quality control methodology, experiments, implementation and improvements

Speech Data Operations Project Manager *Google*

October 2015 - June 2016

As manager of the Vietnamese speech data operations team, I managed a team of evaluators to work on the launch of the first Vietnamese TTS voice. I led the creation of the phonemic guidelines, quality control, development of the speech database, development of the pronunciation dictionary, and evaluation of the final product. Due to limited resources, we optimized our performance with creative solutions. Our first Google voice was well-received by the media in Vietnam.

Leading Vietnamese speech data operations team
Lead Text to Speech projects including lexical transcription, studio recording, text normalization
Develop studio guidelines for recording quality
Create guidelines for phonemic transcriptions
Study loanwords adaptation and develop phonemic transcription guidelines for foreign words
Manage the speech database and oversee quality control efforts
Designing Vietnamese phonemic inventory for Google's TTS voice
Write ASR grammars for verbalizing numeric strings

COMPUTATIONAL SKILLS

Programming Languages	Python, shell scripting, C++, Java, R,
	Wolfram Language, Praat scripting, MATLAB
	yodalog, gqui, SQL, dremel
Software & Tools	LaTeX, vim, Git, Mathematica, SPSS, Praat
Libraries	pandas, numpy, scikit-learn

NATURAL LANGUAGES

Vietnamese	native
English	full professional proficiency
Japanese	JLPT 2 certified
Mandarin Chinese	conversational
German	knowledge of grammar

SELECTED COURSEWORK

* STAT 400 (MATH 463) Statistics and Probability I
STAT 410 (MATH 464) Statistics and Probability II
STAT 425 Applied Regression and Design
STAT 432 Basics of Statistical Learning
STAT 431 Applied Bayesian Analysis
LING 402 Tools and Techniques of Speech and Language Processing
* LING 514 Design and Methods in Linguistic Research
LING 506 Introductory Machine Learning
* LING 401 Introduction to General Phonetics
* LING 520 Acoustic Phonetics
* LING 506 Neural Machine Translation

(* indicates a grade of A+ earned)

RESEARCH INTERESTS

Computational linguistics, experimental linguistics, psychoacoustics, machine translation, natural language processing, phonology, phonetics

RESEARCH PROJECTS

Distortion in Tone Production due to the Lombard Effect

- Designing production experiments to investigate the effect of noise levels on tones articulation in Northern Vietnamese

- Conducting data analysis and modeling from acoustic and laryngograph waveform signals

- Studying the implications for speech perception and designing follow-up perceptual studies

- Abstract

Multilingual NMT Evaluation (in collaboration with Brennan Dell)

- Developing evaluation methods to examine NMT systems' performance on specific linguistic phenomena

PUBLICATIONS

Le, Giang; Mori, Shinka; Schwartz, Lane. 2021. Illinois Japanese ↔ English News Translation for WMT 2021. In *Proceedings of the Sixth Conference on Machine Translation (WMT)*. pp 149-158. November 10-11, 2021. Association for Computational Linguistics.

Le, Giang. 2021. L1 Vietnamese L2 English Speakers' Cues to the Perception of Stress, *Studies in the Linguistic Sciences: Illinois Working Papers 2021*

in preparation Le, Giang; Shih, Chilin; Tang, Yan. 2021. Distortion in Tone Production due to the Lombard Effect. In *Proceedings of the First International Conference on Tone and Intonation.*

INVITED TALKS

Text-to-Speech and the Linguist's Role in Industry - League of Linguists, Fall 2020

CONFERENCE PRESENTATIONS

Distortion in Tone Production due to the Lombard Effect - 1st International Conference on Tone and Intonation, Dec 2021

Illinois Japanese ↔ English News Translation for WMT 2021 - EMNLP 2021 Sixth Conference on Machine Translation (WMT 2021), Nov 2021

Do we change the tone quality spontaneously when speaking in non-ideal conditions? - Mid-continental Phonetics and Phonology Conference 26, University of Chicago, Sep 2021

L1 Vietnamese L2 English Speakers' Cues to the Perception of Stress - Illinois Language and Linguistics Society 13 and Asian Conference on Languages, Spring 2021

ACADEMIC HONORS AND AWARDS

Graduate College Conference Presentation Awards, University of Illinois, Spring 2021
LSO Service Recognition Award, University of Illinois, Spring 2021
Human Subjects Award, University of Illinois, Spring 2020
Departmental Fellowship, University of Illinois, Fall 2019
Provost's Fellowship, University of Southern California, 2009
Travel Award, University of Southern California, 2012
Phi Beta Kappa, Ohio Wesleyan University, 2009
Graduation with University Honors, Magna Cum Laude, Ohio Wesleyan University, 2009
Omicron Delta Epsilon (International Honor Society in Economics), 2009
Outstanding Major in Economics, Ohio Wesleyan University, 2009
Florence Leas Freshman Mathematics Prize, Department of Mathematics and Computer Science, Ohio Wesleyan University, 2007
Phi Eta Sigma (National Freshman Honor Society), 2006
Presidential Scholar, Ohio Wesleyan University, 2006

TEACHING APPOINTMENTS

Instructor   *UIUC*
LING 100 - Introduction to Linguistic Science   Spring 2022

Head Teaching Assistant   *UIUC*
LING 100 - Introduction to Linguistic Science   Fall 2021
Instructor: Professor Ryan Shosted

Teaching Assistant *UIUC*
LING 302 - Elements of Phonology Spring 2021
Instructor: Professor Ryan Shosted

Teaching Assistant *UIUC*
LING 100 - Introduction to Linguistic Science Fall 2020
Instructor: Professor Ryan Shosted

Teaching Assistant *UIUC*
LING 490 - Computational Morphology Spring 2020
Instructor: Professor Lane Schwartz

PROFESSIONAL MEMBERSHIP

Association for Laboratory Phonology
Association for Computational Linguistics
Linguistic Society of America

UNIVERSITY & PROFESSIONAL SERVICE

Organizer *UIUC*
Computational Linguistics Reading Group Fall 2021 -

Abstract Reviewer *UIUC*
Slavic Linguistics Society 16 Spring 2021
Illinois Language and Linguistics Society 14 Spring 2022
Illinois Language and Linguistics Society 13 Spring 2021
Illinois Language and Linguistics Society 12 Spring 2020

Organizing Committee Member *UIUC*
Illinois Language and Linguistics Society 13 2020-2021
Illinois Language and Linguistics Society 12 2019-2020

Treasurer *UIUC*
Linguistics Student Organization 2020-2021

EDUCATION

University of Illinois at Urbana Champaign *current, expected 2023*
Ph.D. Candidate in Linguistics
Current GPA: 4.00/4.00

University of Southern California
M.A. in East Asian Languages and Cultures *2014*

Ohio Wesleyan University
B.A. in Economics, East Asian Studies *2009*

REFERENCES

Professor Chilin Shih cls@illinois.edu
Professor Lane Schwartz lanes@illinois.edu
Professor Yan Tang yty@illinois.edu
Professor Ryan Shosted rshosted@illinois.edu