GIANG LE

gianghl2@illinois.edu

EDUCATION

Ph.D. Candidate, Linguistics *expected 2023*
University of Illinois at Urbana Champaign

M.S. Statistics *2023*
University of Illinois at Urbana Champaign

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

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

RESEARCH INTERESTS

Experimental linguistics, computational linguistics, psychoacoustics, machine translation, natural language processing, phonology, phonetics

RESEARCH PROJECTS

Fairness in Artificial Intelligence: A New Paradigm for the Evaluation and Training of Inclusive Automatic Speech Recognition

- PI: Prof. Mark Hasegawa-Johnson, Prof. Zsuzsanna Fagyal

- Abstract

- Mine large, publicly available text and speech corpora (e.g., a 100,000-podcast corpus) in collaboration with the research team working on inclusive speech recognition

Tone Perception in Noise

- Designing perception experiments for tone perception in noise following our production results of Lombard speech

PUBLICATIONS

Le, Giang; Tang, Yan. 2023. The Lombard Effect on the Vowel Space of Northern Vietnamese In *Proceedings of the 20th International Congress of Phonetic Sciences, Prague 2023.*

Le, Giang; Shih, Chilin; Tang, Yan. 2022. A Laryngographic Study on the Voice Quality of Northern Vietnamese Tones under the Lombard Effect In *INTERSPEECH 2022.*

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.*

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.*

INVITED TALKS

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

CONFERENCE PRESENTATIONS

The Lombard Effect on the Vowel Space of Northern Vietnamese - The 20th International Congress of Phonetic Sciences, Prague 2023

A Laryngographic Study on the Voice Quality of Northern Vietnamese Tones under the Lombard Effect - INTERSPEECH, Sep 2022

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

Department of Linguistics Conference Presentation Awards, University of Illinois, Fall 2021
Teacher Ranked Excellent by Students, University of Illinois, Fall 2021
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

ACADEMIC APPOINTMENTS

Research Assistant *UIUC*
Fairness in Artificial Intelligence: A New Paradigm for the Evaluation and Training of Inclusive ASR

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

UNIVERSITY & PROFESSIONAL SERVICE

Reviewer
2023 Eighth Conference on Machine Translation (WMT23)   Fall 2023

Reviewer
2022 Seventh Conference on Machine Translation (WMT22)   Fall 2022

Organizer   *UIUC*
Computational Linguistics Reading Group   Fall 2021 - Spring 2022

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

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

Treasurer   *UIUC*
Linguistics Student Organization   2020-2021

INDUSTRY EXPERIENCE

Dictionary Database Manager   *The Language Conservancy*
July 2022 - present

Analytical Linguistic Project Manager   *Google via Adecco, Tokyo Japan*
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.

Speech Data Operations Project Manager   *Google via Beyondsoft, Taipei Taiwan*
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.

PROFESSIONAL MEMBERSHIP

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

COMPUTATIONAL SKILLS

Programming Languages	Python, shell scripting, R, Wolfram Language, Praat scripting yodalog, gqui, SQL, dremel Java, Lua (basic)
Software & Tools	LaTeX, RMarkdown, vim, Git, Mathematica, SPSS, SAS, Praat, TLex
Libraries	pandas, numpy, scikit-learn

NATURAL LANGUAGES

Vietnamese	native
English	full professional proficiency
Japanese	JLPT N2 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
STAT 448 Advanced Data Analysis
STAT 446 Data Science Programming Methods
STAT 510 Mathematical Statistics
STAT 527 Advanced Regression Design

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
LING 446 Fundamentals for Speech Signal Processing and Analysis

(* indicates a grade of A+ earned)

REFERENCES

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