Borui (Bri) Zhang

zhan3829@umn.edu | 320-761-4876 | https://boruizhang.github.io/

Education

University of Minnesota Minneapolis, MN
Ph.D. in Linguistics (minor in Computer Science) May 2021 (expected)
Award: Graduate Fellowship
Relevant Coursework: Computational Linguistics, Syntactic Theory, Semantics, Phonetic and Phonology
Artificial Intelligence, Data Structure and Algorithms

Tianjin Foreign Studies University Tianjin, China
B.S. in Educational Technology July 2011

Research Experience

United Language Group | *Computational Linguist Research Intern* Minneapolis, MN

- Built parallel corpora for machine translations in Chinese, Japanese, and Korean Jan - Aug 2017
- Segmented the data and evaluated the models (rule-and-NN based), using Moses, NLTK, NLPIR

Fallon Worldwide | ***Language consultant*** Minneapolis, MN

- Provided semantic knowledge to the design team, and created a semantic entertainment program 2017
- Used Twitter IPA and Python and implemented a program that captures keywords in customers' tweets

Artificial intelligence course project

- Examined whether linguistic functors have effects on the vector space modeling (TensorFlow-word2vec)
- The results show functors have negative effects on vector space modeling due to their syntactic variations

Computational linguistics course project

- Compared the efficiency of regex taggers and bigram taggers on LCMC corpus, using HMM method
- Regex approach performed better on time-saving, achieved the same accuracy: 85% on unseen data

MA thesis project

- Examined Entropy Reduction predictions on Mandarin relative clauses, used CTB7 corpus and Tregex tool
- Extracted data with Tregex coding; found and normalized the sets of context free grammar rules that generate these data using CCPC; calculated the probabilities of the target sentence structures
- The resulting predictions of ER model reflect the results from human language processing experiments

Leadership Experience

Institute of Linguistics at U of M | *Research Fellow/Teaching Assistant* 2015 - present

- Working with community members of low-resource languages, eliciting accurate data in scientific and ethnical ways, store and analyze the data for linguistic research
- Leading review sessions for two 50-minute classes of 15~30 junior college students; prepare lectures and small class activities; grade assignments and exams, and provide feedbacks

College of Science and Engineering at U of M | *MnDRIVE Scholars* May - Aug 2019

- Worked in a passionate team of engineering students; designed and taught basic robotic courses to five groups of 30~40 students from Grade 7 to 9 in a one-week long summer camp during the summer months
- Taught the courses in programming Ozobots, Inkscape graphic editor, and intro to virtual reality

Skills

Programming: Python, R, Matlab, Haskell, Bash, JavaScript, HTML, CSS, RegEx, Tregex | Packages and deep learning tools: TensorFlow, PyTorch, NLTK, Moses, WordNet

Languages: Mandarin, Cantonese, Korean | Fieldwork in: Nepal Bhasa, Somali, Sino-Tibetan language family

Publications

3 single-author publications: CLS 54th (2018) | LSA 92nd (2018) | BEAL (2016)

2 first-author publications: Glow-in-Asia XII (2019) | FASAL 7th (2017)

Volunteers

Organizer of UMN Syntax-Semantic Group | Officer of International Latin Dance Group | Former student organizer of UMN Linguistics Colloquium