Han ZHOU

Kerry001.hi@163.com | (+86) 15221567568 | Homepage: https://hanzhou925.github.io/

Education

MSc. University College London (Language Science in Neuroscience, Language & Communication) Sept.2022 **(**Prospective)
B.A. Shanghai Jiao Tong University (German) Sept.2018-Jun. 2022
Courses: Natural Language Processing (90); Probability and Statistics (88); Introduction to Calculus (97); Syntax (91); Introduction to German Linguistics (94); Frontiers in Linguistics (98); AI Programming Framework (audit); Introduction to AI.
Awards: SJTU Scholarships (Top 20%); Third Prize of National English Competition for College Students (Top 5%).

Work Experience

NLP Product Operation, AI-LAB, ByteDance Technology Co., LTD Feb. 2022-Jun. 2022

- Led the TikTok Cross-Language Assessment project, surveyed factors affecting users' satisfaction with the live subtitle function on TikTok, and analysed data from over 1000 captioned videos with problems based on different angles like positions, translation accuracy and stickers, thus determining the priorities to improve the live captioning function;
- Interviewed 5 groups of German monolinguals and researched their habits and opinions of using subtitles when watching videos in foreign languages on TikTok to locate ways to improve the TikTok subtitle function based on users' need;
- Assessed and analysed over 1500 corpus items corrected by Feishu's auto-correct function from 5 dimensions such as grammar, semantics and pragmatics, and compared the statistical results with Grammarly, thus producing structured reports on the existing problems with the auto-correct function and providing feasible suggestions to upgrade Feishu.

TTS Linguistic Tech Intern, Ximalaya Technology Co., LTD (top Internet company in AI voice with over 2000 employees) Jun. 2021-Sept. 2021

- Led the program of emotion annotation in the data team, annotated the emotions in dialogues of four novel types into 8 categories, and analysed the emotion polarity of narration, thus establishing the emotion dataset with over 24,000 data for the production of human emotion in Text-To-Speech (TTS);
- Proofread over 70,000 polyphone/homograph Mandarin data, and added over 3000 missing data accordingly through BCC Chinese Corpus after analysing the polyphone coverage with Python and Excel, thus expanding the coverage of the dataset by 2% and improving the accuracy of machine learning;
- Evaluated and fixed over 1000 bad cases of AI pronunciation in Mandarin from the angle of prosody, phonetics and emotions, regularised and cleaned the newly added Mandarin data with Python, enhancing the pronunciation of AI Voice.

Online Course Editor, Shanbay Internet Co., LTD (famous for English apps with over 20M. DAU) Jan. 2021-Feb. 2021

- Designed analysis of 5 latest test papers of the CET-4 and CET-6, and recorded online video courses reaching over 100,000 subscribers, helping students improve their grades by 15% and gained over 85% of satisfaction from students;
- Annotated and labelled 1050 sentences for the lexicon of Shan Bei App and improved the accuracy of online translation by 20% in cooperation with R&D department.

Research Experience

The Regulation of Relative Clause Types on Causality Expectation in Real-time Discourse Processing
Independent Research | Advisor: Professor Fuyun Wu May. 2021-Jul. 2022

- Compared and analysed over 10 most recent papers on discourse processing, thus finding out the research gap in English research of Hoek (2020&2021), i.e., the effect of concessive relative clause (RC) on expectation of coherence, and expanding the experiment to Chinese to explore the effect of three RC types on prediction of causality in discourse;
- Designed 30 sets of Chinese stimuli and 60 fillers based on English experiments by Hoek (2021), each including a lead-in sentence, a matrix sentence with a neutral/causal/concessive relative clause, a connective clause, and a wrap-up sentence, and programmed with Experiment Builder for the eye-tracking experiment.
- Gathered data of eye movements during online sentence reading after recruiting 50 participants, extracted the targeted indexes and cleaned data with R, controlled the skipping rate and then built the Mixed-linear Regression Model to calculate the impact of RC types on the sentence processing during online reading.

The Impact of Quantificational Cues on L2 Subject-verb Agreement Processing: Evidence from P600
National-level Undergraduate Innovation Program (Highest level of school funding) Sept. 2018-Sept. 2020

- Analysed over 50 related theses in the last 5 years, thus proposing the research topic based on the currently competing theories including the *Unified Competition Model* and *Shallow Structure Hypothesis*, and investigated the subject-verb agreement processing mode of L2 speakers from the perspective of quantifications, which is a novel angle in this topic;
- Designed 60 sets of test materials for the 3×2 experiment including quantifications (definite quantifier *three*, indefinite quantifier *many*, and article *the*) and grammaticality (grammatical/ungrammatical), and conducted the behavioural test and ERP experiment among Chinese-English speakers;
- Calculated the participants' accuracy of behavioural test and recorded P600 effect, and proposed two explanations based on the ANOVA analysis for the discussion part, thus accomplishing a 10-thousand-word dissertation which achieved an A.

Other Information

Leader, Public Relations Department, Student International Organization Association Sept. 2018-Sept. 2020

- Invited several IO officials, thus organizing the Training Camp for IO Talents with an attendance of over 50 people;
- Organised the visiting program to offices of IOs in Bangkok and Beijing and expanded the visit to Tokyo.

Computer Skills: Microsoft Office; Python; RStudio; Experiment Builder; Data Viewer; Gorilla; Azure.
Language Skills: English (IELTS: 8.0; Certificate of Advanced English Interpretation); German (Zertifikat C1, TestDaf:18)