

# Stella Li

Baltimore, MD 21218 | +1(650) 660-3990 | [sli136@jhu.edu](mailto:sli136@jhu.edu) | [LinkedIn](#)

## RESEARCH INTERESTS

Natural Language Processing, Computational Linguistics, Code-Mixing, Low-Resource, Multilingual NLP, Explainable AI

## EDUCATION

Aug. 2019 – Current	<b>Johns Hopkins University</b>   Baltimore, MD BS in Computer Science Thesis: Learning from Gibberish: Code-Mixing Data Augmentation for Sentiment Analysis Other Majors: Cognitive Science (linguistics focus), Applied Mathematics (statistics focus) Minor: Mathematics Cumulative GPA: 3.99/4.0, Major GPA: 4.0/4.0
Relevant Coursework	Natural Language Processing, Machine Learning: Deep Learning, Machine Translation, Human Language Technology, Data Science, Human-Computer Interaction, Math Models of Language

## PUBLICATIONS

2022	[1] <b>Shuyue Stella Li</b> , Hannah Peeler, Andrew N. Sloss, Kenneth N. Reid, and Wolfgang Banzhaf. 2022. Genetic improvement in the shackleton framework for optimizing LLVM pass sequences. In Proc. of the Genetic and Evolutionary Computation Conference Companion (GECCO'22).
2022	[2] Hannah Peeler, <b>Shuyue Stella Li</b> , Andrew N. Sloss, Kenneth N. Reid, Yuan Yuan, and Wolfgang Banzhaf. 2022. Optimizing LLVM pass sequences with shackleton: a linear genetic programming framework. In Proc. of the Genetic and Evolutionary Computation Conference Companion (GECCO'22).
2022	[3] <b>Shuyue Stella Li</b> and Kenton Murray. 2022. Learning from Gibberish – Synthetic Language Agnostic Data Augmentation for Code-Mixing Sentiment Analysis. Under review at EACL'23.
2022	[4] <b>Shuyue Stella Li*</b> , Xiangyu Zhang*, Shu Zhou, Hongchao Shu, Ruixing Liang, Hexin Liu, Leibny Paola Garcia. 2022. PQLM - Multilingual Decentralized Portable Quantum Language Model for Privacy Protection. Under review at ICASSP'23.
2022	[5] Xiangyu Zhang, <b>Shuyue Stella Li</b> , Zhanhong He, Roberto Togneri, and Leibny Paola Garcia. 2022. End-to-End Lyric Recognition with Self-Supervised Learning Models. Under review at ICASSP'23.
2022	[6] Yu Xuan, Xiangyu Zhang, <b>Shuyue Stella Li</b> , Zihan Shen, Leibny Paola Garcia, and Roberto Togneri. 2022. A new approach to extract fetal electrocardiogram using convex combination of adaptive filters. Under review at ICASSP'23.

## RESEARCH EXPERIENCE

### Language & Speech Processing

May 22 – Current	<b>Multilingual Generalizability of Self-Supervised Models</b>   Johns Hopkins CLSP This study uses large self-supervised models trained on English as a feature extractor for ASR tasks in languages from diverse language families to investigate multilingual generalizability of these models. We use deepCCA to examine whether semantics or phonology is learned by SSL.
May 22 – Oct. 22	<b>Language Agnostic Code-Mixing Data Augmentation</b>   Johns Hopkins CLSP This work analyzes the underlying linguistic phenomenon of Code-Mixing. I invented a language agnostic data augmentation algorithm for Code-Mixing languages by leveraging cross-lingual information. Results are under review at EACL'23.

Mar. 22 – Oct. 22	<b>Quantum Language Model</b>   Johns Hopkins CLSP This project addresses the concern of data privacy in large pre-trained language models. We train language embeddings on quantum machines to preserve privacy without sacrificing performance. We also propose a portable method to transfer quantum information to classical machines and achieve competitive performance on downstream tasks. Results are under review at ICASSP'23.
Feb. 22 – Oct. 22	<b>Self-Supervised Learning for Lyric Recognition</b>   Johns Hopkins CLSP This work establishes an end-to-end lyrics recognition baseline and uses self-supervised learning models to outperform state-of-the-art by 2.4%. Moreover, we study the generalization ability of the SSL features using attention weights. Results are under review at ICASSP'23.
<b>General Machine Learning</b>	
Sep. 20 – Current	<b>Delineo Covid-19 Modeling</b>   Johns Hopkins & University of Tsukuba & Microsoft Research Designed expansion-based clustering algorithm for community detection using covid-19 mobility data. Built agent-based stochastic model for disease spread simulation and intervention modeling.
Jul. 22 – Current	<b>Cryptographic Code Optimization</b>   Michigan State University & University of Adelaide Optimizing cryptographic code generation with genetic improvement. This project aims to automatically reduce the computation time for encryption algorithms.
Mar. 22 – Oct. 22	<b>Fetal Electrocardiogram Extraction</b>   Johns Hopkins School of Medicine This project uses a convex combination of adaptive filters to detect the fetal heart rate from the ECG signals from the mother's abdominal and thorax. Our algorithm out-performs both single adaptive filters and convex combination of RLS filters. Results are under review at ICASSP'23.
May 21 – Jul. 22	<b>Genetic Programming – LLVM Optimization</b>   Michigan State University Designed and implemented novel GP algorithm for LLVM compiler flag optimization (20%) Published work at GECCO; second author of <a href="#">GP paper</a> ; first author of <a href="#">GI paper</a>
Jan. 20 – Jan. 22	<b>Language and Cognition Lab</b>   Johns Hopkins Language and Cognition Lab Investigated developmental spatial cognition using Lego Block building Created ML model for movement prediction and stability analysis using motion sensor data

## TEACHING EXPERIENCE

Sep. 2022 – Current	<b>Human-Computer Interaction (EN.601.490)</b>   Johns Hopkins University   Baltimore, MD <b>Course Assistant</b> Supervised design projects, held weekly office hours, graded homework & exams
May. 2022 – Aug. 2022	<b>Computer Ethics (EN.601.104)</b>   Johns Hopkins University   Baltimore, MD <b>Head Course Assistant</b> Finding supplemental materials, organizes in-class presentations, graded homework & exams
Sep. 2021 – May 2022	<b>Introduction to Statistics (EN.553.430)</b>   Johns Hopkins University   Baltimore, MD <b>Teaching Assistant</b> Taught recitation sections, held weekly office hours, graded homework & exams
Sep. 2021 – May 2022	<b>Intermediate Programming (EN.601.220)</b>   Johns Hopkins University   Baltimore, MD <b>Course Assistant</b> Oversaw in class exercises, held weekly office hours, graded homework & exams

## INDUSTRY EXPERIENCE

May 22 – Aug. 22	<b>Yext</b>   Arlington, VA <b>Software Engineering Intern</b> Integrated client data to Yext platform for real-time site information updates using Go Created a Figma Style Picker to improve developer workflow and scalability using ReactJS
------------------	--

May 20 – Aug. 20	<b>Bytedance (TikTok) AI Lab</b>   Beijing, China <b>Algorithms Engineering Intern</b> Trained neural networks for text normalization in TTS tasks Implemented algorithms for theme clustering and complexity ranking for TikTok videos
May 19 – Aug. 19	<b>IBM AI Doctor</b>   Beijing, China <b>Data Science Intern</b> Created ML models to predict diseases from symptoms using EHR records Improved classification accuracy from 74% to 99% with a hybrid algorithm of GA with SVM

## HONORS

2022	Upsilon Pi Epsilon International Computer Science Honor Society
2021 & 2022	Omicron Delta Kappa National Leadership Honors Society
2022	Omega Psi National Cognitive Science Honors Society

## AWARDS

Jul. 2022	GECCO-GI Best Presentation Award
Jul. 2022	ACM Student Travel Grant
May 2022	PAJH Greek Scholars Award
Sep. 2021	Grace Hopper Scholarship Award
Every Semester	Dean's List (GPA > 3.50)

## LEADERSHIP

Oct. 19 – Current	<a href="#"><u>HopHacks</u></a> at Johns Hopkins, Organizer & Head of Sponsorships
Oct. 19 – Current	<b>Omega Psi Cognitive Science Honor Society (JHU Chapter)</b> , President (22)
Oct. 19 – Current	<b>Alpha Phi Omega National Service Fraternity</b> , Executive Board Member (20)
Feb. 20 – Current	<b>Phi Mu Fraternity</b> , Academics Excellence Chairwoman (20-21), Honors Member (22)
Apr. 20 – Current	<b>Hopkins Undergraduate Society of Applied Mathematics</b> , President (22)
Oct. 20 – Current	<b>Women in Computer Science (WiCS)</b> , Executive Board Member (21)
Nov. 21 – Current	<b>Theta Tau Engineering Fraternity</b> , Social Chair (22)

## SKILLS

Programming Languages	Python; C/C++; Java, R, MATLAB, HTML/CSS, JavaScript, ReactJS, Go
Languages	English (native); Mandarin Chinese (fluent); Spanish (intermediate)