

# Henry Li Xinyuan

✉ xinyuan.li1301@gmail.com • 🌐 hstehstehste.github.io

🔗 <https://scholar.google.com/citations?user=tyKMv8MAAAAJ&hl=fr>

## About

- PhD Student at Johns Hopkins University, Center for Speech and Language Processing.
- Advised by Sanjeev Khudanpur
- Focus areas : Voice Conversion, TTS, Adversarial Robustness, Speech Recognition, Speech Translation

## Education

<b>Johns Hopkins Whiting School of Engineering</b> <i>PhD in Electrical and Computer Engineering</i>	<b>Baltimore, Maryland</b> Sep. 2022–
<b>Johns Hopkins Whiting School of Engineering</b> <i>MSE in Computer Science</i>	<b>Baltimore, Maryland</b> Jan. 2021– May. 2022
<b>University of Oxford</b> <i>BA Mathematics and Computer Science</i>	<b>Oxford, United Kingdom</b> 2016–2019

## Publications

<b>2025</b>	
<b>Scalable Controllable Accented TTS</b> <i>Henry Li Xinyuan, Z. Cai, A. Garg, K. Duh, L.P. García-Perera, S. Khudanpur, N. Andrews, M. Wiesner</i>	<b>ASRU 2025</b> Aug. 2025
<b>HLTCOE Submission to the VoicePrivacy Attacker Challenge</b> <i>Henry Li Xinyuan, Z. Cai, A. Garg, K. Duh, L.P. García-Perera, S. Khudanpur, N. Andrews, M. Wiesner</i>	<b>ICASSP 2025</b> Apr. 2025
<b>GenVC: Self-Supervised Zero-Shot Voice Conversion</b> <i>Z. Cai, Henry Li Xinyuan, A. Garg, K. Duh, L.P. García-Perera, S. Khudanpur, N. Andrews, M. Wiesner</i>	<b>ASRU 2025</b> Feb. 2025
<b>ShiftySpeech: A Large-Scale Synthetic Speech Dataset with Distribution Shifts</b> <i>A. Garg, Z. Cai, Henry Li Xinyuan, K. Duh, L.P. García-Perera, S. Khudanpur, N. Andrews, M. Wiesner</i>	<b>ASRU 2025</b> Feb. 2025
<b>2024</b>	
<b>HLTCOE JHU Submission to the Voice Privacy Challenge 2024</b> <i>Henry Li Xinyuan, Z Cai, A Garg, K Duh, Leibny Paola García, S Khudanpur, N Andrews, M Wiesner</i> Best Paper at Voice Privacy Challenge 2024	<b>VPC 2024, SPSC 2024</b> Sep. 2024
<b>Clean Label Attacks against SLU Systems</b> <i>Henry Li Xinyuan, Sonal Joshi, Thomas Thebaud, Jesus Villalba, Najim Dehak, Sanjeev Khudanpur</i>	<b>IEEE SLT 2024</b> Dec. 2024
<b>Privacy versus Emotion Preservation Trade-offs in Emotion-Preserving Speaker Anonymization</b> <i>Zexin Cai, Henry Li Xinyuan, A Garg, Leibny Paola García, K Duh, S Khudanpur, N Andrews, M Wiesner</i>	<b>IEEE SLT 2024</b> Dec. 2024
<b>JHU IWSLT 2024 Dialectal and Low-resource System Description</b> <i>NR Robinson, K Sun, C Xiao, N Bafna, W Tan, H Xu, Henry Li Xinyuan, ...</i>	<b>IWSLT 2024</b> Jul. 2024
<b>2023</b>	
<b>Clustering Unsupervised Representations as Defense Against Poisoning Attacks</b> <i>Thomas Thebaud, Sonal Joshi, Henry Li, Martin Sustek, Jesús Villalba, Sanjeev Khudanpur, Najim Dehak</i>	<b>ASRU 2023</b> Dec. 2023

<b>Joint Energy-Based Model for Robust Speech Classification System Against Poisoning Attacks</b> <i>Martin Sustek, Sonal Joshi, <b>Henry Li</b>, Thomas Thebaud, Jesús Villalba, Sanjeev Khudanpur, Najim Dehak</i>	<b>ASRU 2023</b> <i>Dec. 2023</i>
<b>HK-LegiCoST: Leveraging Non-Verbatim Transcripts for Speech Translation</b> <i>Cihan Xiao, <b>Henry Li Xinyuan</b>, Jinyi Yang, Dongji Gao, Matthew Wiesner, Kevin Duh, Sanjeev Khudanpur</i>	<b>Interspeech 2023</b> <i>Dec. 2023</i>
<b>JHU IWSLT 2023 Multilingual Speech Translation System Description</b> <b>Henry Li Xinyuan</b> , Neha Verma, Bismarck Bamfo Odoom, Ujala Pradeep, Matthew Wiesner, Sanjeev Khudanpur Ranked 1st in the IWSLT ACL 60-60 initiative, unconstrained track	<b>IWSLT 2023</b> <i>Jul. 2023</i>
<b>Learning a Formality-Aware Japanese Sentence Representation</b> <b>Henry Li Xinyuan</b> , Ray Lee, Jerry Chen, Kelly Marchisio	<i>Jan. 2023</i>

## Other Projects

<b>Submission to the Generative Design in Minecraft (GDMC)</b>	<b>GDMC 2022</b> <i>May. 2022</i>
<b>The Phonology of consonant voicing in Shanghainese</b>	<b>JHU, Term Paper</b> <i>May. 2021</i>
<b>Approximate solutions for the Multi-Commodity Flow problem: a survey</b>	<b>JHU, Term Paper</b> <i>May. 2021</i>
<b>Voice Control Rave Goggles</b>	<b>Oxford Hack 2017</b> <i>Nov. 2017</i>

## Contribution to Open Source Projects

<b>K2 Icefall, Lhotse</b> : Added new recipe to speech toolkit.	<i>Jan 2024</i>
<b>Adversarial Robustness Toolbox</b> : Implemented clean label poisoning attacks in the audio domain.	<i>Feb 2024</i>

## Experience

<b>Work Experience</b> .....	
<b>Visiting Research Scholar</b> <i>Human Language Technology Center of Excellence (HLTCOE)</i>	<b>Baltimore, Maryland</b> <i>Jun. 2023–Aug 2023</i>
<b>Graduate Research Assistant</b> <i>Johns Hopkins University Center for Language and Speech Processing</i>	<b>Baltimore, Maryland</b> <i>Sep. 2021–Present</i>
<b>Course Assistant : CS 601.645 Natural Language Processing</b> <i>Johns Hopkins University</i>	<b>Baltimore, Maryland</b> <i>Sep. 2022–Dec. 2022, Sep. 2023–Dec. 2023</i>
<b>Course Assistant : EN 520.666 Information Extraction from Speech and Text</b> <i>Johns Hopkins University</i>	<b>Baltimore, Maryland</b> <i>Sep. 2022–Dec. 2022, Sep. 2023–Dec. 2023</i>
<b>Architecture Intern</b> <i>NVIDIA</i>	<b>Shanghai, China</b> <i>Oct. 2019–Jan. 2021</i>
<b>Other</b> .....	
<b>Conference Paper Reviewer</b> <i>ICASSP, LREC-COLING, INTERSPEECH</i>	<i>Jan. 2024–Present</i>
<b>Student Researcher</b> <i>Frederick Jelinek Memorial Summer Workshop (JSALT)</i>	<b>Baltimore, Maryland</b> <i>Jun. 2022–Aug. 2022</i>

**Director General**

*Oxford Global Education Development Ltd.*

**Oxford, United Kingdom**

*Feb. 2019–Dec. 2019*

**Topology Seminar TA**

*Harvard Summit for Young Leaders in China*

**Shanghai, China**

*Aug. 2018*

## Languages

---

**Chinese(Mandarin), Shanghainese, English, French:** Native or Fluent

**Italian:** Basic