Run Chen

📞 +1-(617)-335-9517
✉ runchen@cs.columbia.edu
🎓 google scholar

▬▬ Summary

- Graduating (Dec 2025) PhD researcher with 10+ years of experience in NLP and computational linguistics looking for Summer 2025 internship and full-time positions in industry
- Published in top conferences (EMNLP, ICASSP, Interspeech); awarded Columbia Presidential Fellowship (top 4 in department); Silver Medalist at the International Linguistic Olympiad
- Expertise in Multimodal AI, Spoken Language Processing, Speech Processing, Emotion and Sentiment Analysis, Dialog Systems

▬▬ Education

Sept. 2020 - present
Columbia University, New York, NY.
Candidate for PhD in *Computer Science*. Advisor: **Julia Hirschberg**
- Awarded Columbia University Presidential Fellowship.

Sept. 2019 - May 2020
Massachusetts Institute of Technology (MIT), Cambridge, MA.
Master of Engineering in *Computer Science and Molecular Biology*.

Sept. 2015 - Jun. 2019
Massachusetts Institute of Technology (MIT), Cambridge, MA.   **GPA: 5.0/5.0**
Bachelor of Science in *Linguistics and Philosophy*, and in *Computer Science and Molecular Biology*.

▬▬ Research and Work Experience

Sept. 2020 - present
Graduate Research Assistant, *Columbia University Spoken Language Processing Group*, New York, NY
- Led multiple DARPA and industry-funded projects, co-authored and secured funding for Amazon CAIT 2022 grant ($74K) and DARPA CCU grant ($5M)
- Conducted dataset collection/curation and generated synthetic datasets for LLMs
- Created multimodal, cross-lingual, and cross-cultural NLP models to detect communication changes, empathetic conversations, sarcasm; built expressive text-to-speech synthesis systems for voice conversion

May 2022 - Aug. 2022
Applied Scientist Intern *Amazon Alexa AI*, remote in New York, NY
- Built models for identifying entrainment behaviors between users and agents
- Analyzed entrainment levels in task-oriented and chit chat dialogs
- Extracted accoutic-prosodic features from the DSTC-10 Task 1 corpus

May 2021 - Sept. 2021
Applied Scientist Intern *Amazon Alexa AI*, remote in New York, NY
- Built dialog act classification BERT models using both text and speech features
- Extracted Parselmouth and OpenSMILE features from the Switchboard corpus
- Designed dialog act schema and annotated dialog acts on a development set for a newly collected speech corpus of task-oriented conversations.

Jun. 2016 - May 2018
Research Assistant, *MIT Computer Science and Artificial Intelligence Lab (CSAIL)*, Cambridge, MA
- Prepared stimuli and ran experiments to collect eye movement data
- Extracted and analyzed data for correlation between native languages and fixation parameters
- Received credits in publication: *Predicting Native Language from Gaze* (Berzak et al., 2017 ACL)

Selected Publications

[1] **Run Chen**, J. Shin, and J. Hirschberg, "Synthempathy: A scalable empathy corpus generated using LLMs without any crowdsourcing," submitted to ACL, 2025.

[2] H. Chen, **Run Chen**, and J. Hirschberg, "EmoKnob: Enhance voice cloning with fine-grained emotion control," in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Y. Al-Onaizan, M. Bansal, and Y.-N. Chen, Eds., Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 8170–8180. [Online]. Available: https://aclanthology.org/2024.emnlp-main.466.

[3] **Run Chen**, H. Chen, A. Kulkarni, E. Lin, L. Pang, D. Tadimeti, J. Shin, and J. Hirschberg, "Detecting empathy in speech," in *Interspeech 2024*, 2024, pp. 1080–1084. DOI: 10.21437/Interspeech.2024-347.

[4] D. Bhattacharya, E. Lin, **Run Chen**, and J. Hirschberg, "Switching tongues, sharing hearts: Identifying the relationship between empathy and code-switching in speech," in *Interspeech 2024*, 2024, pp. 492–496. DOI: 10.21437/Interspeech.2024-1224.

[5] **Run Chen**, E. Lin, S. Hooshmand, M. Mustafa, R. Sloan, R. Nandi, A. Yang, A. Lopez, A. Kothary, I. Suh, C. Lyu, E. Chen, S. Horng, and J. Hirschberg, "Raswda: Re-aligned switchboard dialog act corpus for dialog act prediction in conversations," in *14th International Workshop on Spoken Dialogue Systems Technology (IWSDS 2024)*, 2024.

[6] **Run Chen**, S. Kim, A. Papangelis, J. Hirschberg, Y. Liu, and D. Hakkani-Tür, "Identifying entrainment in task-oriented conversations," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2023, pp. 1–5. DOI: 10.1109/ICASSP49357.2023.10095543.

[7] L. Ai, **Run Chen**, Z. Gong, J. Guo, S. Hooshmand, Z. Yang, and J. Hirschberg, "Exploring new methods for identifying false information and the intent behind it on social media: Covid-19 tweets.," in *Sixth International Workshop on Social Sensing (SocialSens 2021), AAAI-ICWSM Workshop*, 2021.

[8] **Run Chen**, "The superiority effect in albanian multiple wh-movement structures," 1, vol. 5, 2020, pp. 623–630. DOI: 10.3765/plsa.v5i1.4735. [Online]. Available: http://journals.linguisticsociety.org/proceedings/index.php/PLSA/article/view/4735.

[9] **Run Chen**, *Linguistics Olympiad Guide for Beginners, Yu Yan Xue Ru Men Zhi Nan*, Chinese. Hangzhou, China: Zhejiang People's Publishing House, Dec. 2014, ISBN: 9787213064401. [Online]. Available: https://books.google.com/books?id=q2gBjwEACAAJ.

Awards and Honors

2020-2024	Columbia University Presidential Fellowship ($94,467)
2015-2019	MIT Gur Tsoo Yuan Scholarship
2017-2019	Eta Kappa Nu National Honor Society Member
2017-2018	MIT EECS | CS+HASS Undergraduate Research and Innovation Scholar NLP Project: Understanding Reading Comprehension of Nonnatives Using Eye-tracking Method
2016-2018	MIT International Science and Technology Initiatives (MISTI) Hayashi Seed Fund Student Researcher
2014	International Linguistics Olympiad, Silver Medal, Team Rank 3rd

Languages and Technical Skills

Programming	Python, Java, JavaScript, SQL, MATLAB, R, Git, LaTeX, SPSS
Languages	Mandarin Chinese, Cantonese, Wu, Japanese, Russian, American Sign Language