Erica L. Cooper

ecooper@nii.ac.jp

Education	**Columbia University,** New York, NY
	PhD (February 2019)
	Department: Computer Science
	Advisor: Julia Hirschberg
	Thesis: "Text-to-Speech Synthesis Using Found Data for Low-Resource Languages"

Massachusetts Institute of Technology, Cambridge, MA
Master of Engineering (June 2010)
Bachelor of Science (June 2009)
Major: Computer Science

Research and Employment Experience

National Institute of Informatics, Tokyo, Japan   *February 2019-Present*
Yamagishi Lab Postdoctoral Researcher
- Contributing to the JST-ANR CREST VoicePersonae project.

Google UK Ltd., London, UK   *Summer 2012*
Speech Group Software Engineering Intern
- Investigated a universal background model approach to cross-lingual speaker adaptation for speech synthesis.
- Trained text-to-speech models and generated speech samples in order to evaluate this approach for use in a mobile speech-to-speech translation system.

Google Inc., New York, NY   *Summer 2011*
Speech Group Software Engineering Intern
- Investigated bootstrapping acoustic models for new languages for Voice Search.
- Developed tools to compare the effectiveness of different phonesets for ASR in a given language.
- Contributed to the development of pronunciation models for Hungarian and Bulgarian voice search.

Google Inc., New York, NY   *Summer 2010*
Speech Group Software Engineering Intern
- Assisted with the integration of third-party HMM-based speech synthesis tools with the main codebase.
- Trained HMM-based synthesizer voices on large amounts of data from multiple speakers and recording conditions.

Infolab Group, Cambridge, MA   *September 2009-June 2010*
MIT, Master of Engineering Thesis Project
- Created an email notification system for the START question-answering system that notifies developers when web wrapper script or cache failures occur.
- Implemented a recovery system that lets START back off to older, cached versions of webpages when wrapper scripts fail due to webpage format changes.
- Modified START's internal HTML parser to act as a repair module for broken wrapper scripts.

Google Inc., New York, NY   *Summer 2009*
Speech Group Software Engineering Intern
- Helped to build infrastructure for extracting and validating ad-hoc pronunciations from the web. Evaluated different methods of extraction and validation.
- Evaluated the baseline pronunciation lexicon for Voice Search in terms of coverage and word error rate.
- Ran experiments to evaluate the effect of new pronunciations from the web on recognition and word error rate.
- Helped to design and implement a C++ library for generic probabilistic models and mixture models.

Center for Language and Speech Processing, Baltimore, MD   *Summer 2008*
Workshop in Human Language Technology at Johns Hopkins
- Participated in the workshop group "Multilingual Spoken Term Detection: Finding and Testing New Pronunciations."
- Contributed to the development of an OpenFST-based indexing and search tool.
- Investigated the effects of "fuzzy matching" on retrieval of spoken documents by expanding query terms with letter-to-sound and sound-to-sound rules.
- Assisted in the evaluation of new pronunciations extracted from the web by measuring their performance in spoken term detection tasks.

Infolab Group, Cambridge, MA   *June 2007-January 2008*
Undergraduate Research at the MIT Computer Science and Artificial Intelligence Lab
- Assisted in the development of an online question-answering system by implementing tools for extracting information from diverse web sources.
- Implemented, developed new features for, and maintained a web-caching tool to facilitate the storage and organization of information collected from the internet.

Computer Vision Research Group, Cambridge, MA   *Summer 2006*
Undergraduate Research at the MIT Computer Science and Artificial Intelligence Lab
- Implemented usability features for a web-based image annotation tool used in machine vision research.
- Assisted in the design and implementation of new methods for annotating images using the annotation program.

Peer-reviewed Publications

J. Williams, Y. Zhao, E. Cooper, J. Yamagishi, "Learning Disentangled Phone and Speaker Representations in a Semi-Supervised VQ-VAE Paradigm," ICASSP 2021.

S. Kato, Y. Yasuda, X. Wang, E. Cooper, J. Yamagishi, "How Similar or Different Is Rakugo Speech Synthesizer to Professional Performers?" ICASSP 2021.

E. Cooper, C-I. Lai, Y. Yasuda, J. Yamagishi, "Can Speaker Augmentation Improve Multi-Speaker End-to-End TTS?" Interspeech 2020.

Y. Zhao, H. Li, C-I. Lai, J. Williams, E. Cooper, J. Yamagishi, "Improved Prosody from Learned F0 Codebook Representations for VQ-VAE Speech Waveform Reconstruction," Interspeech 2020.

S. Kato, Y. Yasuda, X. Wang, E. Cooper, S. Takaki, J. Yamagishi, "Modeling of Rakugo Speech and Its Limitations: Toward Speech Synthesis That Entertains Audiences," *IEEE Access*, August 2020.

E. Cooper, C-I. Lai, Y. Yasuda, F. Fang, X. Wang, N. Chen, J. Yamagishi, "Zero-Shot Multi-Speaker Text-To-Speech with State-of-the-art Neural Speaker Embeddings," International Conference on Acoustics, Speech, and Signal Processing, May 2020.

E. Tesfaye Biru, Y. Tofik Mohammed, D. Tofu, E. Cooper, J. Hirschberg, "Subset Selection, Adaptation and Gemination for Amharic Text-to-Speech Synthesis," 10th ISCA Speech Synthesis Workshop, September 2019.

S. Kato, Y. Yasuda, X. Wang, E. Cooper, S. Takaki, J. Yamagishi, "Rakugo speech synthesis using segment-to-segment neural transduction and style tokens – toward speech synthesis for entertaining audiences," 10th ISCA Speech Synthesis Workshop, September 2019.

K. Lee, E. Cooper, J. Hirschberg, "A Comparison of Speaker-based and Utterance-based Data Selection for Text-to-Speech Synthesis," Interspeech, September 2018.

E. Cooper, J. Hirschberg, "Adaptation and Frontend Features to Improve Naturalness in Found-Data Synthesis," Speech Prosody, June 2018.

E. Cooper, E. Li, J. Hirschberg, "Characteristics of Text-to-Speech and Other Corpora," Speech Prosody, June 2018.

E. Cooper, X. Wang, A. Chang, Y. Levitan, J. Hirschberg, "Utterance Selection for Optimizing Intelligibility of TTS Voices Trained on ASR Data," Interspeech, August 2017.

E. Cooper, A. Chang, Y. Levitan, J. Hirschberg, "Data Selection and Adaptation for Naturalness in HMM-based Speech Synthesis," Interspeech, September 2016.

G. Mendels, E. Cooper, J. Hirschberg, "Babler - Data Collection from the Web to Support Speech Recognition and Keyword Search," ACL WAC-X, August 2016.

E. Cooper, Y. Levitan, J. Hirschberg, "Data Selection for Naturalness in HMM-based Speech Synthesis," Speech Prosody, May 2016.

G. Mendels, E. Cooper, V. Soto, J. Hirschberg, M. Gales, K. Knill, A. Ragni, H. Wang, "Improving Speech Recognition and Keyword Search for Low Resource Languages Using Web Data," Interspeech, September 2015.

V. Soto, E. Cooper, A. Rosenberg, J. Hirschberg, "Rescoring Confusion Networks for Keyword Search," International Conference on Acoustics, Speech, and Signal Processing, May 2014.

V. Soto, E. Cooper, A. Rosenberg, J. Jirschberg, "Cross-Language Phrase Boundary Detection," International Conference on Acoustics, Speech, and Signal Processing, May 2013.

A. Rosenberg, E. Cooper, R. Levitan, J. Hirschberg, "Cross-Language Prominence Detection," Speech Prosody, May 2012.

D. Can, E. Cooper, A. Sethy, B. Ramabhadran, M. Saraclar, C.M. White, "Effect of Pronunciations on OOV Queries in Spoken Term Detection," International Conference on Acoustics, Speech, and Signal Processing, April 2009.

C.M. White, A. Sethy, B. Ramabhadran, P. Wolfe, E. Cooper, M. Saraclar, J.K. Baker, "Unsupervised Pronunciation Validation," International Conference on Acoustics, Speech, and Signal Processing, April 2009.

A. Sethy, B. Ramabhadran, D. Can, E. Cooper, M. Saraclar, A. Ghoshal, M. Jansche, M. Riley, S. Khudanpur, M. Ulinski, "Web-derived Pronunciations for Spoken Term Detection," Special Interest Group on Information Retrieval, July 2009.

Grants	**Transfer Learning for End-to-End Multi-Instrument MIDI-to-Music Synthesis,** Japan Science and Technology Agency: AIP Challenge Program, PI. June 2020 - March 2021.

Transfer Learning for End-to-End Multi-Instrument MIDI-to-Music Synthesis, JHCPN: TSUBAME Encouragement Program for Young/Female Users, PI. April 2020 - March 2021.

Encoder Factorization for Capturing Dialect and Articulation Level in End-to-End Speech Synthesis, Japan Society for the Promotion of Science: Grant-in-Aid for Research Activity Start-up, PI. August 2019 - March 2021.

Invited Talks	**Found Data and Speaker Modeling for Text-to-Speech Synthesis,** Department of Electrical and Computer Engineering, National University of Singapore. November 6, 2019.

Text-to-Speech Synthesis Using Found Data, IBM Thomas J. Watson Research Center, Yorktown Heights, New York. April 26, 2018.

Research Supervision Experience

PhD Student Interns at NII

- Jennifer Williams (University of Edinburgh)   *March 2020 - present*
  Voice conversion using a vector-quantized variational autoencoder model.

- Antoine Perquin (IRISA)   *November 2019 - March 2020*
  French and multilingual multi-speaker Tacotron.

- Cheng-I Lai (MIT)   *July 2019 - May 2020*
  Speaker and dialect embeddings for speaker verification and for transfer learning to multi-speaker Tacotron-based TTS.

Masters, Undergraduate, and High School Students at Columbia

- Yishak Tofik Mohammed (Columbia University)   *May 2018 - Jan 2019*
  Experiments using found Amharic data for TTS including audiobible and other read speech corpora.

- Elshadai Tesfaye Biru (Columbia University)   *January 2018 - January 2019*
  Experiments using Amharic BABEL and other data for TTS; development of internal infrastructure for listening tests.

- Kai-Zhan Lee (Columbia University)   *May 2017 - May 2018*
  Experiments with MACROPHONE data for TTS, and development of pipeline for large-scale experimentation.

- Emily Li (Columbia University)   *March 2017 - May 2018*
  Development of and experiments with statistical analytical tools for comparing acoustic properties of various speech corpora.

- David Tofu (Columbia University)   *January 2017 - January 2018*
  Experiments using Amharic BABEL and Audiobible data for TTS.

- Mert Ussakli (Columbia University)   *January 2017 - May 2017*
  Experiments using Turkish BABEL data for TTS.

- Xinyue Wang (Columbia University)   *January 2016 - August 2017*
  Experiments using MACROPHONE data for TTS, and development of tools for objective evaluation.

- Olivia Lundelius (Great Neck South High School)   *July 2016 - August 2016*
  Web scraping of low-resource language speech data for training TTS models.

- Luise Valentin Rygaard (University of California, Berkeley) *May 2015 - July 2015*
  Experiments using synthesized speech for data augmentation to improve ASR.

- Gideon Mendels (Columbia University)   *September 2014 - May 2017*
  Web scraping and experimentation with text data in low-resource languages for ASR language modeling.

- Yocheved Levitan (Columbia University)   *May 2014 - January 2017*
  Experiments with BURNC data for TTS, development of crowdsourced naturalness evaluation infrastructure, and experiments with syllabification for TTS.

- Alison Chang (Columbia University) *January 2014 - May 2016*
  Experiments with CALLHOME data for TTS, and development of crowdsourced intelligibility evaluation infrastructure.

Peer Review Service

- NAACL 2021 (area chair)
- ICASSP 2021 (meta-reviewer)
- SLT 2021
- *IEEE Transactions on Audio, Speech and Language Processing,* October 2020
- Blizzard Challenge and Voice Conversion Challenge 2020
- Interspeech 2020
- ICML Self-supervision in Audio and Speech Workshop, 2020
- *Computer Speech and Language,* June 2020
- *Speech Communication,* November 2019 and January 2020
- Speech Prosody 2020
- ICASSP 2020
- Speech Synthesis Workshop 2019

Teaching Experience

Spoken Language Processing *Columbia, Spring 2011 and 2012*
Teaching Assistant
- Helped students in lab with homework assignments and projects.
- Assisted in grading exams and projects.

Computation Structures *MIT, Spring 2010*
Teaching Assistant
- Planned and taught twice-weekly recitations on course content.
- Helped students in lab with assignments.
- Assisted in grading exams and lab assignments.

Introduction to Python *MIT, January 2010*
Laboratory Assistant
- Assisted students with daily Python programming assignments.

Artificial Intelligence *MIT, Fall 2009*
Teaching Assistant
- Planned and taught weekly tutorials on course content.
- Led review sessions to help students prepare for exams.
- Assisted in writing and grading exams.

Activities and Committees

- IEEE SLTC Associate Member, December 2020 - present
- Assisted in reviewing applications for the M.S. in Computer Science program, 2015, 2016, and 2018.
- Student Representative, Columbia University Institutional Review Board, July 2014-December 2018.
- Treasurer, Columbia University Women in Computer Science, September 2010-May 2014.

Honors and Awards

- Best Poster Award, Speech Processing Courses in Crete, 2016.
- Computer Science Service Award, Columbia University, 2012, 2013, 2014, 2016.
- 3rd Prize, CSAW Voice Biometrics and Speech Synthesis Competition, 2012.
- Computer Science Chair's Distinguished Fellowship, Columbia University, 2010.