

# Anuj Jitendra Diwan

Second Year Computer Science PhD at UT Austin

✉ anuj.diwan@utexas.edu • 🌐 ajd12342.github.io • 🐦 anuj\_diwan  
 🌐 ajd12342 • 🎓 Google Scholar

## Education

### University of Texas at Austin

PhD in Computer Science

Co-advised by Prof. David Harwath and Prof. Eunsol Choi. GPA 4.00/4.00

Austin, TX, USA

Aug 2021 - Present

### Indian Institute of Technology Bombay

B.Tech in Computer Science and Engineering with Honors

Minor in Applied Statistics and Informatics

B.Tech (Major) CPI: 9.75/10, Minor CPI: 10/10

Mumbai, India

Jul 2017 - Aug 2021

## Relevant Research Experience

### PhD Student, UT Austin Computer Science

Advisors: Prof. David Harwath, Prof. Eunsol Choi

Austin, TX, USA

Aug 2021 - Present

- o Currently working on building unsupervised textless speech-to-speech translation systems.
- o Mentoring an undergrad project on collecting two-party speech dialogue data for the product price negotiation setting.
- o Investigated failures of multimodal models on visiolinguistic compositionality benchmarks (EMNLP 2022 [Paper](#)).
- o Profiled models from three modalities on several efficiency metrics in a layer-wise fashion. Found that some 'efficient' models are in fact less efficient and that non-self-attention modules are also computationally expensive (Under review).
- o Developed phoneme-based masking strategies for self-supervised speech representation learning ([Report](#), [Slides](#)).

### Research Intern (AI), Meta AI (FAIR)

Host: Dr. Abdelrahman Mohamed

Seattle, WA, USA

May 2022 - Present

- Continual Learning for On-Device Speech Recognition using Disentangled Conformers* (under review)
- o Introduced a new continual learning benchmark called LibriContinual for efficient speaker-specific domain adaptation.
  - o Developed a novel Disentangled Conformer model that is capable of highly parameter-efficient continual learning without suffering from any catastrophic forgetting.
  - o Working on extending the benchmark to the self-supervised setting and improving our model to be more speaker-aware.

### Undergraduate Researcher, IIT Bombay Computer Science

Advisors: Prof. Preethi Jyothi, Prof. Sunita Sarawagi

Mumbai, India

Jan 2020 - Apr 2021

- o Served on the Shared Task Committee for the MUCS (Multilingual and Code-Switching) ASR challenge and released labelled speech data in 7 Indian languages for multilingual and code-switching ASR (Interspeech 2021 [Paper](#)).
- o Proposed novel transliteration-based pretraining for high-to-low resource language transfer (Interspeech 2021 [Paper](#)).
- o Proposed a novel phoneme-inspired reduce-and-reconstruct technique for low-resource ASR (Interspeech 2021 [Paper](#)).
- o Developed a speech-grounded text transliteration system that only requires monolingual speech training data ([Report](#)).

## Publications

(\* indicates equal contribution)

1. **Anuj Diwan**, Ching-Feng Yeh, Wei-Ning Hsu, Paden Tomasello, Eunsol Choi, David Harwath, Abdelrahman Mohamed. "Continual Learning for On-Device Speech Recognition using Disentangled Conformers," (*under review*)
2. **Anuj Diwan\***, Layne Berry\*, Eunsol Choi, David Harwath, Kyle Mahowald. "Why is Winoground Hard? Investigating Failures in Visiolinguistic Compositionality," *EMNLP 2022 Long (Oral)*. [\[paper\]](#)[\[code\]](#)
3. **Anuj Diwan\***, Puyuan Peng\*, Raymond J. Mooney. "Zero-shot Video Moment Retrieval With Off-the-Shelf Models," *Workshop on Transfer Learning for NLP at NeurIPS 2022*. [\[paper\]](#)
4. **Anuj Diwan**, Preethi Jyothi. "Reduce and Reconstruct: ASR for Low-Resource Phonetic Languages," *INTERSPEECH 2021*. 🏆 Shortlisted for the *Best Student Paper Award*. [\[paper\]](#)[\[slides\]](#)
5. **Anuj Diwan**, Rakesh Vaideeswaran, Sanket Shah, Ankita Singh et. al. "MUCS 2021: Multilingual and code-switching ASR challenges for low resource Indian languages," *INTERSPEECH 2021*. [\[paper\]](#)[\[code\]](#)

6. Shreya Khare\*, Ashish Mittal\*, **Anuj Diwan\***, Sunita Sarawagi, Preethi Jyothi, Samarth Bharadwaj. "Low Resource ASR: The surprising effectiveness of High Resource Transliteration," *INTERSPEECH 2021*.  
[\[paper\]](#)[\[slides\]](#)

## Internships

### Research Intern (AI), Meta AI (FAIR)

Host: [Dr. Abdelrahman Mohamed](#)

Continual Learning for On-Device Speech Recognition using Disentangled Conformers

Seattle, WA, USA

May 2022 - present

### Research Intern, Big Data Experience Lab, Adobe Research

Advisors: [Sunav Choudhary](#), [Subrata Mitra](#)

Learning to Learn with Interruptions

Bangalore, India

Apr - Jul 2020

### Research Intern, INMA, ICTEAM, UCLouvain

Advisor: [Prof. Pierre-Antoine Absil](#)

Graph-regularized Matrix Completion using Riemannian manifolds

Louvain-la-Neuve, Belgium

May - Jul 2019

[\[report\]](#)[\[code\]](#)

## Awards

- o Shortlisted for the ISCA Best Student Paper Award at Interspeech 2021 for 'Reduce and Reconstruct'.
- o Awarded the Excellence in Research Award 2021 by the IIT Bombay CS Department.
- o Awarded the Excellence in Teaching Assistantship Award 2021 by the IIT Bombay CS Department.
- o Awarded the IIT Bombay Undergraduate Research Award (URA 01), Spring 2020.
- o Awarded 7 AP grades for outstanding performance (top 1%) in 7 courses.
- o Awarded the Institute Academic Prize for Academic Excellence by IIT Bombay for 2017-18.
- o Among top 35 in India selected for the International Mathematical Olympiad [Training Camp](#) in 2016.
- o Achieved All India Rank 118 in JEE Advanced 2017 and All India Rank 197 in JEE Mains 2017.

## Programming Skills

**Programming Languages & Tools:** Python, C/C++, Java, Bash, MATLAB, Javascript, Git,  $\LaTeX$ , Beamer.

**AI/ML Libraries:** Pytorch, Tensorflow, Numpy, Fairseq, ESPNet, Kaldi, OpenFST, Huggingface Transformers.

## Professional Services

**Reviewer:** IEEE JSTSP. **Teaching Assistant:** Speech Recognition, Logic for CS, Computer Programming.

## Other Research Experience

### Stem2Morph: Low-Resource Morphological Inflection

[\[slides\]](#)[\[code\]](#)

Course Project (Natural Language Processing). Advisor: [Prof. Pushpak Bhattacharya](#)

Jan - May 2021

- o Reimplemented the 'Pushing the Limits of Low-Resource Morphological Inflection' paper's original Dynet implementation in Pytorch. Found that transfer learning using related languages helps. Built a user-friendly demo.

### Negative Interference in Multilingual Models and Code-Switching

[\[slides\]](#)

Course Project (Advanced Machine Learning). Advisor: [Prof. Sunita Sarawagi](#)

Jan - May 2021

- o Showed that pretraining a bilingual Hindi-English model results in worse Hindi-English code-switching performance compared to monolingual pretraining, indicating negative interference and proposed potential fixes.

### Anuvaadya: Instrumental Music Translation

[\[report\]](#)[\[code\]](#)

Course Project (Automatic Speech Recognition). Advisor: [Prof. Preethi Jyothi](#)

Aug - Nov 2019

- o Implemented the 'A Universal Music Translation Network' paper that uses a CNN WaveNet Autoencoder and tested it for Indian instruments. Extended the paper's idea to arbitrary length sequence data by implementing LSTM RNN autoencoders in Pytorch inspired by the novel elements of the paper (attention and a domain confusion network).

### Learning to Learn with Interruptions

Research Intern. Advisors: [Sunav Choudhary](#), [Subrata Mitra](#), Adobe Research

Apr - Jul 2020

- o Designed a novel reinforcement learning-based scheduler that can run ML jobs on interruptible AWS cloud-based VMs.

### Graph-regularized Matrix Completion using Riemannian manifolds

[\[report\]](#)[\[code\]](#)

Research Intern. Advisor: [Prof. Pierre-Antoine Absil](#), UCLouvain

May - Jul 2019

- o Implemented Riemannian optimization for matrix completion using [graph-regularization](#) in MATLAB and LAPACK.