Aditya Parikh | NLP & Speech AI Researcher

Nijmegen – The Netherlands

✉ aditya.parikh@ru.nl • aditya3107.github.io • aditya-parikh

Experience

Radboud University — **Nijmegen, The Netherlands**

PhD Candidate and ICT Developer — *May 2021 – Present*

- **PhD Candidate | Automatic Speech Assessment for L2 Learners** *Jan 2024 – Present*
  - Investigated automated L2 speech assessment across the full methodological arc, from classical GOP-based mispronunciation detection to zero-shot and fine-tuned large speech-language models, contributing to both phonetic modeling and interpretable feedback generation.
  - Proposed rubric-guided fine-tuning with uncertainty calibration and Direct Preference Optimisation for multi-granular, multi-aspect L2 assessment, achieving strongest alignment with human raters across multiple dimensions of speech.
  - Published 5 first-author papers at leading venues (Interspeech 2025 ×2, SLaTE 2025, LREC 2026, Interspeech 2026), spanning mispronunciation detection, zero-shot SpeechLLM evaluation, and rubric-guided fine-tuning with natural-language rationale generation.
- **ICT Developer | Audio Processing Pipeline (ATRIUM Project EU)** *Sep 2024 – Dec 2025*
  - Architected and deployed a production-ready multilingual audio intelligence pipeline for medical consultations and archaeological survey interviews, processing hours-long conversations daily and reducing manual documentation burden for medical professionals and researchers.
  - Built a modular end-to-end system integrating automatic language detection, Whisper-based ASR, pyannote speaker diarization with timestamps, translation, and LLM based summarisation, deployed via FastAPI and Docker with modular outputs selectable per request across multiple European languages.
  - Solved long-context summarisation for hours long conversations using sliding window chunking, processing the full pipeline with very low latency on a single GPU.
- **Junior Researcher | Inclusive AI (SignON Project EU)** *May 2021 – Dec 2023*
  - Built and owned the ASR component of an EU Horizon 2020 project developing inclusive speech and sign language translation for Deaf and Hard of Hearing (DHH) users, designing systems to handle atypical speech while meeting real-time latency and GDPR privacy requirements critical for accessibility.
  - Evolved the ASR stack from hybrid Kaldi-based systems to self-supervised end-to-end models (wav2vec 2.0, Whisper), improving accuracy and robustness across multiple languages including low-resource settings.
  - Deployed production REST APIs with measured latency, request throughput, and load handling to meet real-time performance requirements of a live mobile application.
  - Improved ASR robustness through ensemble techniques combining hybrid and end-to-end outputs, achieving ∼14% Word Error Rate Reduction (LREC-COLING 2024).

Otto-von-Guericke-University **Magdeburg, Germany**
Master's Thesis *Sep 2020 – Mar 2021*
Proposed a novel codec-based audio augmentation technique, generating acoustically diverse training copies by compressing audio at varying bitrates without altering the original signal, applicable to any speech or audio task. Validated on speech emotion recognition, outperforming SpecAugment with up to 9% improvement in mean F1 score.

Fraunhofer IIS **Erlangen, Germany**
Research Internship *Mar 2020 – Jun 2020*
Designed and integrated neural and statistical language model components into production ASR pipelines, optimising decoding strategies to achieve measurable reductions in word and character error rates.

Tata Consultancy Services **Ahmedabad, India**
System Engineer *Oct 2016 – Oct 2018*
Built and maintained automated test pipelines for large-scale national tax portals, reducing manual testing effort and improving release stability across Agile delivery cycles.

Teaching & Academic Service

Aug 2024 – Dec 2024: Teaching Assistant, Radboud University — Research-Master courses: *Introduction to Language and Speech Technology* and *Transformer-Based Models*.

Sep 2025: Guest Lecture, Radboud University — *Child Speech Recognition* for the Research-Master course Introduction to Language and Speech Technology.

May 2026: Session Chair, LREC 2026, Palma, Spain.

Aug 2025: Session Chair, Interspeech 2025, Rotterdam, The Netherlands.

Education

Otto-von-Guericke-University **Magdeburg, Germany**
M.Sc. Electrical Engineering and Information Technology *Oct 2018 – Mar 2021*

Gujarat Technological University **Ahmedabad, India**
B.Eng. Electrical Engineering *Jun 2012 – May 2016*

Technical Skills

Programming: Python, Bash/Shell, SQL

ML & Training: PyTorch, Hugging Face Transformers, PEFT/LoRA, Scikit-learn, TensorFlow/Keras, NumPy, Pandas

LLM & Alignment: Instruction tuning, supervised fine-tuning, direct preference optimisation (DPO), retrieval-augmented generation (RAG), prompt engineering, conformal calibration

Speech & Audio: Whisper, wav2vec 2.0/XLS-R, pyannote, Kaldi, Librosa, TorchAudio, NVIDIA NeMo

APIs & Deployment: FastAPI, Flask, Streamlit, Docker

MLOps: Weights & Biases, TensorBoard, Git

Cloud & Compute: HPC environments (SURF), GPU-based distributed training

Publications

First-Author

2026: **A Finetuned SpeechLLM for Joint Multi-Granular L2 Assessment and Natural-Language Rationales**. **A. K. Parikh**, C. Tejedor-García, C. Cucchiarini, & H. Strik. arXiv:2606.09470. *Accepted at Interspeech 2026, Sydney, Australia*.

2026: **Rubric-Guided Fine-tuning of SpeechLLMs for Multi-Aspect, Multi-Rater L2 Reading-Speech Assessment**. **A. K. Parikh**, C. Tejedor-García, C. Cucchiarini, & H. Strik. *Proceedings of LREC 2026* (pp. 10255–10265). ELRA. doi: `https://doi.org/10.63317/4dgvijh3226x`.

2025: **Zero-Shot Speech LLMs for Multi-Aspect Evaluation of L2 Speech: Challenges and Opportunities**. **A. K. Parikh**, C. Tejedor-Garcia, C. Cucchiarini, & H. Strik. *Proc. SLaTE 2025* (pp. 11–15). doi: 10.21437/SLaTE.2025-3.

2025: **Enhancing GOP in CTC-Based Mispronunciation Detection with Phonological Knowledge**. **A. K. Parikh**, C. Tejedor-Garcia, C. Cucchiarini, & H. Strik. *Proc. Interspeech 2025* (pp. 5068–5072). doi: 10.21437/Interspeech.2025-829.

2025: **Evaluating Logit-Based GOP Scores for Mispronunciation Detection**. **A. K. Parikh**, C. Tejedor-Garcia, C. Cucchiarini, & H. Strik. *Proc. Interspeech 2025* (pp. 2405–2409). doi: 10.21437/Interspeech.2025-1012.

2024: **Ensembles of Hybrid and End-to-End Speech Recognition**. **A. K. Parikh**, L. ten Bosch, & H. van den Heuvel. *Proceedings of LREC-COLING 2024* (pp. 6199–6205). url: `https://aclanthology.org/2024.lrec-main.547/`.

2023: **Comparing Modular and End-To-End Approaches in ASR for Well-Resourced and Low-Resourced Languages**. **A. Parikh**, L. ten Bosch, H. van den Heuvel, & C. Tejedor-García. *Proceedings of ICNLSP 2023* (pp. 266–273). url: `https://aclanthology.org/2023.icnlsp-1.28/`.

2022: **Design Principles of an Automatic Speech Recognition Functionality in a User-centric Signed and Spoken Language Translation System**. **A. K. Parikh**, L. F. M. ten Bosch, H. van den Heuvel, & C. Tejedor García. *CLIN Journal 2022*. url: `https://www.clinjournal.org/clinj/article/view/145`.

Co-Authored

2026: **Generating High Quality Synthetic Data for Dutch Medical Conversations**. C. Kuan, **A. K. Parikh**, & H. van den Heuvel. *Proceedings of LREC 2026* (pp. 10750–10763). ELRA. doi: `https://doi.org/10.63317/52kv8b8eq52o`.

2023: **SignON: Sign Language Translation. Progress and Challenges**. V. Vandeghinste, D. Shterionov, M. De Sisto, A. Brady, M. De Coster, L. Leeson, J. Blat, F. Picron, M. P. Scipioni, **A. Parikh**, & L. ten Bosch. *Proceedings of EAMT 2023* (pp. 501–502). url: `https://aclanthology.org/2023.eamt-1.53/`.

2022: **Sign Language Translation: Ongoing Development, Challenges and Innovations in the SignON Project**. D. Shterionov, M. De Sisto, V. Vandeghinste, A. Brady, M. De Coster, L. Leeson, J. Blat, F. Picron, M. P. Scipioni, **A. Parikh**, & L. ten Bosch. *Proceedings of EAMT 2022* (pp. 325–326). url: `https://aclanthology.org/2022.eamt-1.52/`.