(979) 422-8675
College Station, TX
quamer.waris@gmail.com

Waris Quamer

github.com/warisqr007
linkedin.com/in/warisquamer

Education

Doctor of Philosophy in Computer Science, *Texas A&M University, College Station* — December 2026 (Tentative)
Master of Science in Computer Science, *Texas A&M University, College Station* — August 2023
Bachelor of Technology in Computer Science & Engineering, *Indian Institute of Technology (ISM), Dhanbad* — May 2019

Experience

Research Assistant — **Aug 2021 — Present**
PSI Labs, TAMU — *College Station, TX*

- Led core speech AI research for IARPA ARTS/Honeywell on real-time voice conversion and anonymization, developing streaming systems to modify speaker traits while preserving speech content under strict latency and naturalness constraints.
- Developed low-latency anonymization models including SLT ’24 Best Student Paper, DarkStream (ASRU ’25), and TVTSyn (ICLR ’26), achieving <80 ms latency, 50% EER, and 10× model compression.
- Built PHONOS, a streaming accent-conversion pipeline using golden-speaker generation, DTW alignment, and TVTSyn representations to convert among General American, British, Indian, and Spanish-accented English.
- Designed zero-shot/reference-free accent conversion and VQ-based disentanglement frameworks, reducing accentedness by 67% and showing through TASLP perceptual studies that segmental cues dominate L2 comprehensibility.

Applied Scientist II Intern — **May 2025 — Aug 2025**
Amazon — *Seattle, WA*

- Built PromptDub, a multimodal LLM dubbing-director system that converts video, audio, and script cues into human-editable directions for expressive TTS with emotion and prosody control.
- Developed an LLM-based speech-token generation pipeline for target-speaker dubbing, achieving 0.90 emotion similarity, 54% lower emotion-trajectory DTW, and 91% ABX preference over baselines.

Machine Learning Researcher (Part-time) — **Oct 2023 — Feb 2025**
Sairen — *Paris, France*

- Developed real-time speech denoising and low-latency accent conversion models for conversational call-center audio, improving clarity and intelligibility while preserving speaker identity.
- Built an end-to-end Voice AI agent by integrating ASR, LLM-based dialogue handling, TTS synthesis, and call-orchestration APIs for real-time interaction.

Software Development Engineer — **Jul 2019 — Jul 2021**
Amdocs — *Pune, India*

- Built and shipped full-stack features across 3 enterprise product platforms, spanning frontend UI, backend services, and customer-specific workflows.
- Developed Java/Spring Boot microservices and REST APIs for scalable product customization and system integration.
- Automated dependency verification for production releases, reducing release-preparation time by 50%.

Selected First-Author Publications

- Privacy and Quality trade-off in Real-time Speaker Anonymization via editing of Age and Sex attributes — Interspeech 2026
- TVTSyn: Content-Synchronous Time-Varying Timbre for Streaming Voice Conversion and Anonymization — Proc. ICLR 2026
- DarkStream: Real-time Speech Anonymization with low latency — Proc. ASRU 2025
- Speech Synthesis and Pronunciation Training (book chapter) — The Encyclopedia of Applied Linguistics 2026
- Disentangling Segmental and Prosodic Factors to Non-native Speech Comprehensibility — IEEE/ACM TASLP 2025
- End-to-end Streaming model for low-latency Speech Anonymization — Proc. SLT 2024 (Best student paper)
- Decoupling Segmental and Prosodic cues of Non-native Speech through Vector Quantization — Proc. Interspeech 2023
- Zero-Shot Foreign Accent Conversion without a Native Reference — Proc. Interspeech 2022

Technical Skills

ML/AI	PyTorch, PyTorch Lightning, TensorFlow, scikit-learn, LLMs, multimodal learning, Hydra, Accelerate
Speech/Audio	TTS, ASR, Voice Conversion, Speech Anonymization, Kaldi, Librosa, WebRTC, FFmpeg
Programming/Systems	Python, C++, Java, JavaScript, Docker, Kubernetes, AWS, FastAPI, Flask, REST APIs, Spring Boot