Karen Rosero

5000 Forbes Avenue, Pittsburgh, PA 15213

kroseroj@cmu.edu, karenrosero14@gmail.com

Website LinkedIn Google Scholar

My research develops AI tools for the assessment of speech disorders. I model atypical speech and speech-related facial movement across developmental and acquired communication disorders, including cleft lip and palate, speech sound disorders, and dysarthria. My work combines multimodal modeling, generative methods that adapt to clinical populations with limited data while protecting patient identity, and interpretable assessment, in collaboration with surgeons and speech-language pathologists.

Education

- **Ph.D. in Language and Information Technology** — **2023 – Expected Summer 2027**
  Carnegie Mellon University - GPA: 4.0 — *Pittsburgh, USA*

  Advisor: Prof. Carlos Busso
  Research focus: Audiovisual modeling for atypical speech.
  Transferred from The University of Texas at Dallas (2023–2024) with advisor.

- **Master of Language Technologies** — **2025 – Expected Summer 2027**
  Carnegie Mellon University - GPA: 4.0 — *Pittsburgh, USA*

  Advisor: Prof. Carlos Busso
  Earned en route to the Ph.D.

- **Master of Science in Electrical Engineering** — **2020 – 2022**
  University of Campinas - GPA: 4.0 — *Campinas, Brazil*

  Funded by a FAPESP Research Excellence Fellowship
  Advisor: Prof. Bruno S. Masiero
  Research focus: Bio-inspired AI for sound event localization in spatial audio.

- **Electronics and Telecommunications Engineering** — **2014 – 2019**
  Universidad de las Fuerzas Armadas ESPE - GPA: 3.5 — *Quito, Ecuador*

  Best Graduate Award
  Advisor: Prof. Felipe Grijalva
  Research focus: Speech-inspired machine learning for detecting volcano-seismic events.

Research Interests

- AI for communication disorders
- Multimodal assessment of speech production
- Articulatory modeling of speech
- Personalized speech generation
- Generative AI for limited clinical data
- Speech and vision large language models
- Accessible voice technology
- Interpretable and responsible AI in healthcare

Research Experience

- ***Graduate Research Assistant***, *Carnegie Mellon University* — *Jan 2025 – Present*
- ***Graduate Research Assistant***, *The University of Texas at Dallas* — *Jan 2023 – Dec 2024*

Facial analysis and generation to support clinical assessment
- Improved facial landmark detection for individuals with craniofacial differences without patient training data.
- Developed AI tools for objective assessment of lip symmetry and lip abnormality after cleft lip repair.
- Developed text-guided synthesis and anonymization of facial images that preserve clinically relevant anatomy.

Speech generation for children with speech sound disorders
- Developed generative models that reconstruct a speaker's intended utterance in their own voice.
- Adapted text-to-speech and voice conversion to generate personalized synthetic disordered speech.

- Showed that synthetic disordered speech improves mispronunciation detection and automatic speech recognition.

Multimodal assessment of motor speech disorders
- Improved dysarthria severity estimation by combining speech, text, and video of orofacial movement.
- Proposed a severity-conditioned contrastive objective that treats modality mismatch as a measure of impairment.

Clinical research and grant development
- Collaborated with surgeons and speech-language pathologists at UT Southwestern and Children's Health.
- Prepared IRB protocols for studies involving clinical speech and facial data.
- Co-developed an NIH proposal on AI-based assessment of children's speech.

Agentic evolutionary search for multimodal data
- Lead two master's students applying LLM-guided evolutionary search to fuse multimodal sensor data.
- Use a record of past failures to guide the search and reduce computational cost, extending methods mostly tested on toy problems to real-world tasks.

Collaborative research
- Contributed to a benchmark for automatic phoneme recognition across multiple speech disorders.
- Contributed to a multilingual and code-switched speech dataset.
- Applied mixture-of-experts models to facial expression recognition in video.

○ ***Master's Research Fellow*** (***FAPESP***), *University of Campinas* *Mar 2020 – Feb 2022*

- Developed auditory-inspired AI methods for 3D sound localization, using models of cochlear frequency analysis.
- Applied self-supervised learning to sound localization in spatial audio.
- Developed methods for multichannel spatial audio generation.
- Mentored one undergraduate student working on song emotion recognition, resulting in a journal article.

○ ***Research Assistant***, *Universidad de las Fuerzas Armadas ESPE* *Mar 2019 – Jan 2020*

- Adapted voice activity detection methods to volcano-seismic signals.
- Applied Gaussian mixture models and unsupervised learning to detect and classify volcano-seismic events.

Industry Experience

○ ***AI/ML PhD Intern***, *Apple Inc.* *May 2026 – Aug 2026*

- Designed a controllable speech generation system for speakers with atypical speech.
- Showed that synthetic atypical speech used for ASR training substantially reduces the amount of real data required.
- Built data foundations for the Speech Accessibility Project dataset (500K+ utterances).

○ ***Telecommunications Engineer***, *Puntonet S.A.* *Sep 2018 – Mar 2019*

- Supported enterprise network setup, including firewalls, servers, and wireless, fiber-optic, and satellite links.

Teaching Experience

○ ***Adjunct Professor***, *Universidad San Francisco de Quito* *May 2024 – Dec 2025*

- Designed and taught *Deep Learning*, *Natural Language Processing*, and *Ethics for AI* in the master's programs in Artificial Intelligence and Data Science (remote).
- Supervised four master's theses, resulting in two conference papers and three journal articles.

○ ***Teaching Assistant***, *Carnegie Mellon University, Multimodal Machine Learning* *Jan 2026 – May 2026*

○ ***Teaching Assistant***, *The University of Texas at Dallas, Linear Algebra* *Jan 2023 – May 2023*

Professional Service

○ Program committee member, OneVoice-MSD Workshop on Multilingual Speech Disorders, SLT 2026.
○ Reviewer, LTI PhD admissions committee, Carnegie Mellon University (Fall 2026 and Fall 2027 cycles).
○ Student organizer, 2026 LTI Student Research Symposium, Carnegie Mellon University.
○ Peer reviewer for Interspeech, ACM ICMI, IEEE ICASSP, and IEEE/ACM TASLP.

Honors and Awards

- Travel Award, Doctoral Consortium, IEEE Face and Gesture Recognition Conference (FG), 2025.
- FAPESP Research Excellence Fellowship, São Paulo Research Foundation, 2020 – 2022.
- Best Graduate Award, Universidad de las Fuerzas Armadas ESPE, 2019.

Selected Publications

- **K. Rosero**, E. Yeo, D. R. Mortensen, C. Van't Slot, R. R. Hallac, and C. Busso, "Generative reconstruction of pediatric disordered speech for automated clinical evaluation," in *Proc. IEEE Spoken Language Technology Workshop (SLT)*, Italy, Dec. 2026.

- **K. Rosero**, C.-C. Lee, R. R. Hallac, and C. Busso, "Multimodal biomarkers of dysarthria: Severity-conditioned contrastive alignment of speech, text, and facial asymmetry," in *Proc. 28th Int. Conf. Multimodal Interaction (ICMI)*, Italy, Oct. 2026.

- **K. Rosero**, A. N. Salman, L. M. Harrison, A. A. Kane, C. Busso, and R. R. Hallac, "Deep learning-based assessment of lip symmetry for patients with repaired cleft lip," *The Cleft Palate Craniofacial Journal*, 2025.

- **K. Rosero**, L. M. Harrison, A. A. Kane, C. Busso, and R. R. Hallac, "Anonymizing faces, preserving anomalies: Deep-learning based privacy protection for data of patients with cleft lip," in *International Congress on Cleft Lip/Palate and Related Craniofacial Anomalies*, Japan, Oct. 2025.

- **K. Rosero**, L. M. Harrison, A. A. Kane, R. R. Hallac, and C. Busso, "DiffusionCleft: Facial anomaly synthesis guided by text," in *Proc. 27th Int. Conf. Multimodal Interaction (ICMI)*, Australia, Oct. 2025.

- **K. Rosero**, A. N. Salman, S. Chandra, B. Sisman, C. Van't Slot, A. A. Kane, R. R. Hallac, and C. Busso, "Advancing pediatric ASR: The role of voice generation in disordered speech," in *Interspeech 2025*, The Netherlands, Aug. 2025.

- **K. Rosero**, A. N. Salman, R. R. Hallac, and C. Busso, "Lip abnormality detection for patients with repaired cleft lip and palate: A lip normalization approach," in *Proc. 26th Int. Conf. Multimodal Interaction (ICMI)*, Costa Rica, Nov. 2024.

- **K. Rosero**, A. N. Salman, B. Sisman, R. R. Hallac, and C. Busso, "Enhanced facial landmarks detection for patients with repaired cleft lip and palate," in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition (FG)*, Turkey, May 2024.

A complete list is available on Google Scholar.

Skills and Languages

- **Computing:** Python, MATLAB, PyTorch, TensorFlow, Git, LaTeX, HPC (Slurm), AWS.
- **Speech processing:** Praat, Montreal Forced Aligner, IPA transcription, electromagnetic articulography.
- **Facial and video analysis:** MediaPipe, OpenFace, 3D facial imaging (3dMD), optical flow.
- **Acoustics and hardware:** acoustic and 3D spatial sound recording, Arduino prototyping.
- **Languages:** Spanish (native), English (advanced), Portuguese (advanced).

References

- **Prof. Carlos Busso**, Language Technologies Institute, Carnegie Mellon University, busso@cmu.edu
- **Dr. Rami R. Hallac**, UT Southwestern Medical Center and Children's Health, rami.hallac@childrens.com
- **Prof. Berrak Sisman**, Johns Hopkins University, sisman@jhu.edu