Gerardo Roa-Dabike

Machine Learning Researcher: Speech, Audio and Language

Sheffield, UK
(+44) 7713 670375
gerardo.roa@gmail.com
www.gerardoroadabike.com
groadabike *gerardo-roa-dabike*
Google Scholar

Profile

Machine learning researcher specialising in deep learning for speech, singing, and music, with a particular focus on accessibility for listeners with hearing loss. My PhD introduced deep learning methods for sung-speech recognition and automatic lyric transcription, establishing evaluation methodology for the spoken-to-sung transfer. Since completing my PhD in 2022, I have worked as a postdoctoral researcher leading the machine learning challenges on the Cadenza project, an EPSRC programme spanning four UK universities and industry partners including BBC and RNID. Before moving into research, I spent over a decade in industry leading data analytics and BI at a Chilean financial services firm, building platforms and KPIs used across multiple business units, experience that now underpins how I structure collaboration between different stakeholders and groups, and how I translate real-world issues into usable datasets and benchmarks. I am ready to lead an independent research group and to teach deep learning and speech/audio processing at all levels.

Education

2018–2022 **PhD, Computer Science**, *University of Sheffield*, UK.
Thesis: Adapting Spoken Technologies to Sung Speech.
Deep learning for sung-speech recognition and automatic lyric transcription, with evaluation methodology for the spoken-to-sung transfer. Funded by a Chilean Government Scholarship.

2015–2016 **MSc, Advanced Computer Science**, *University of Sheffield*, UK.
Dissertation: Automatic Speech Recognition in Music.

1998–2003 **BEng, Informatics Engineering**, *Universidad Diego Portales*, Chile.

Research Impact

- **Google Scholar**: 124 citations | h-index: 7 | i10-index: 4
- **Impact**: Citation activity has shown a sustained increase since 2022, reflecting increasing visibility and engagement with my research.

Publications

Journal Articles

2026 Scott Bannister, Jennifer Firth, **Gerardo Roa-Dabike**, Rebecca Vos, William Whitmer, Alinka E Greasley, Simone Graetzer, Bruno Fazenda, Trevor Cox, Jon Barker, and Michael A. Akeroyd. The First Cadenza Challenge: Perceptual Evaluation of Machine Learning Systems to Improve Audio Quality of Popular Music for Those with Hearing Loss. ***Trends in Hearing***, volume 30, page 23312165251408761. SAGE Publications Sage CA: Los Angeles, CA, 2026. `https://doi.org/10.1177/23312165251408761`.

2025 **Gerardo Roa-Dabike**, Michael A. Akeroyd, Scott Bannister, Jon P. Barker, Trevor J. Cox, Bruno Fazenda, Jennifer Firth, Simone Graetzer, Alinka Greasley, Rebecca R. Vos, and William M. Whitmer. The First Cadenza Challenges: Using Machine Learning Competitions to Improve Music for Listeners With a Hearing Loss. ***IEEE Open Journal of Signal Processing***, volume 6, pages 722–734. IEEE, 2025. `https://doi.org/10.1109/OJSP.2025.3578299`.

Data Papers

2026 **Gerardo Roa-Dabike**, Trevor J. Cox, Jon P. Barker, Bruno M. Fazenda, Simone Graetzer, Rebecca R. Vos, Michael A. Akeroyd, Jennifer Firth, William M. Whitmer, Scott Bannister, and Alinka Greasley. The Cadenza lyric intelligibility prediction (CLIP) dataset. ***Data in Brief***, volume 65, page 112466, 2026. https://doi.org/10.1016/j.dib.2026.112466.

2026 **Gerardo Roa-Dabike**, Trevor J. Cox, Michael A. Akeroyd, Scott Bannister, Bruno Fazenda, Jennifer Firth, Simone Graetzer, Alinka Greasley, Rebbeca R. Vos, William M. Whitmer, and Jon P. Barker. Descriptor: The Second Cadenza Lyric Intelligibility Prediction Dataset (CLIP2). ***IEEE Data Descriptions***, 2026. Manuscript under review.

2024 **Gerardo Roa-Dabike**, Trevor J. Cox, Alex J. Miller, Bruno M. Fazenda, Simone Graetzer, Rebecca R. Vos, Michael A. Akeroyd, Jennifer Firth, William M. Whitmer, Scott Bannister, Alinka Greasley, and Jon P. Barker. The Cadenza Woodwind dataset: Synthesised quartets for music information retrieval and machine learning. ***Data in Brief***, volume 57, page 111199. Elsevier, 2024. https://doi.org/10.1016/j.dib.2024.111199.

In Conference Proceedings

2026 **Gerardo Roa-Dabike**, Jon P. Barker, Trevor J. Cox, Michael A. Akeroyd, Scott Bannister, Bruno Fazenda, Jennifer Firth, Simone Graetzer, Alinka Greasley, Rebecca R. Vos, and William M. Whitmer. Overview of the icassp 2026 cadenza challenge: Predicting lyric intelligibility. In ***2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2026)***, pages 21757–21759, 2026. https://doi.org/10.1109/ICASSP55912.2026.11463231.

2024 **Gerardo Roa-Dabike**, Michael A. Akeroyd, Scott Bannister, Jon Barker, Trevor J. Cox, Bruno Fazenda, Jennifer Firth, Simone Graetzer, Alinka Greasley, Rebecca R. Vos, and William M. Whitmer. **The ICASSP SP Cadenza Challenge: Music Demixing/Remixing for Hearing Aids**. In *2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*, pages 93–94, 2024. https://doi.org/10.1109/ICASSPW62465.2024.10626340.

2023 **Gerardo Roa-Dabike**, Scott Bannister, Jennifer Firth, Simone Graetzer, Rebecca R. Vos, Michael A. Akeroyd, Jon P. Barker, Trevor J. Cox, Bruno Fazenda, Alinka Greasley, and William M. Whitmer. The first Cadenza signal processing challenge: Improving music for those with a hearing loss, 2023. https://ceur-ws.org/Vol-3528/paper4.pdf.

2021 **Gerardo Roa-Dabike** and Jon Barker. The use of Voice Source Features for Sung Speech Recognition. In ***2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2021)***, pages 6513–6517, 2021. https://doi.org/10.1109/ICASSP39728.2021.9414950.

2019 **Gerardo Roa Dabike** and Jon Barker. Automatic Lyric Transcription from Karaoke Vocal Tracks: Resources and a Baseline System. In ***Interspeech 2019***, pages 579–583, 2019. https://doi.org/10.21437/Interspeech.2019-2378.

Research and Professional Experience

Jun 2022 – present **Research Associate**, *Cadenza Project*, University of Sheffield & University of Salford.
Machine learning challenge lead within the project. UKRI EPSRC research programme across four universities (Sheffield, Salford, Nottingham, Leeds), with industry partners BBC R&D, Google and Logitech. Programme ends December 2026.

- Led the design and delivery of five international ML challenges on music and speech for listeners with hearing loss. Shaped each research question with the team, then owned datasets, baselines, evaluation and results.
- Published five first-author papers from this work, including the ICASSP 2026 challenge overview and a journal article in IEEE Open Journal of Signal Processing. Three more are in progress.
- Had three challenges accepted as ICASSP Grand Challenges: 2024, 2026 and 2027.
- Built deep learning baselines using CNN and transformer architectures, pretrained speech transformers and LLM-based components. Benchmarked them under real-time and latency constraints.
- Designed and built several public datasets where none existed. One example: a lyric-intelligibility corpus derived from the Free Music Archive, with hearing-loss simulation applied to the audio and word-recognition scores collected from normal-hearing listeners. Released as the CLIP dataset in Data in Brief, 2026.
- Worked across an 11-person team spanning computer science, acoustics, medicine and music psychology.
- Released all datasets, baselines and evaluation code openly through the pyclarity toolkit. Challenge entrants and outside research groups reuse them. Each completed ICASSP edition drew more than 20 entrants, from countries including Hong Kong, Spain and the United States.
- Represented the project at ISMIR and HCMIR, presenting the work to the wider research community.

Feb 2022 – Jan 2024 **Research Associate, Data Analytics Lead**, *COVID-OSA Project*, University of Sheffield.

- Led all data collection and analysis for a study on obstructive sleep apnoea and long-COVID outcomes.
- Curated a distributable clinical dataset of 120 patients. Other researchers now use it beyond the original study.

2004 – 2017 **Deputy Manager, Business Intelligence**, *LarrainVial S.A.*, Chile.

- Consolidated 5 to 10 product databases into one analytics platform. Migrated from MS Access to SQL Server.
- Designed KPIs to benchmark 50 to 100 financial advisors, individually and by group.
- Led data projects across Marketing, Sales and Markets Research, working directly with each set of stakeholders.

Teaching and Academic Service

- Co-supervised an undergraduate researcher on a Sheffield Undergraduate Research Experience (SURE) project, summer 2025, on detecting household sounds for people with hearing loss.
- Guided the student from problem framing to baseline models for four common home sounds in background noise, resulting in two workshop posters.
- Worked as a teaching demonstrator on undergraduate and postgraduate modules in Python, Java and speech processing during the PhD.
- Presented a technical tutorial on git and GitHub for a non-expert audience.
- Wrote and presented one-hour tutorials for VCCA conference.
- Wrote baseline and tutorial materials used by international challenge participants from many backgrounds.
- Coordinated multidisciplinary research collaboration across four universities and three industry partners.
- Peer reviewer for EURASIP Journal on Audio, Speech, and Music Processing, Disability and Rehabilitation: Assistive Technology, and Frontiers in Digital Health (Connected Health section), plus conference reviewer for ICASSP and Interspeech.

Core Competencies

Research	Deep learning for speech, audio and language. Benchmark and evaluation design. Dataset construction from scratch. Research programme leadership. National, international and industry collaboration.
Machine learning	CNNs, RNNs, transformers, pretrained speech transformers, LLM-based components.
Teaching	Undergraduate and postgraduate demonstrating in Python, Java and speech processing. Tutorial and baseline materials for international audiences.
Technical	Python, PyTorch, Kaldi, SQL, Linux, signal processing, HRTF processing.

Honours and Awards

2026 **School Citizen Award**, School of Computer Science, University of Sheffield. For contribution to the School community beyond the expectations of the role.