

GASSER ELBANNA

Doctoral student in the Speech and Hearing Bioscience and Technology program at Harvard University and MIT. I am interested in studying speech and voice processing and perception in the human brain through the lens of deep learning models.

✉ gasser_elbanna@hms.harvard.edu - gelbanna@mit.edu 🌐 gasserelbanna.github.io 📄 github.com/GasserElbanna 🎓 google scholar
in linkedin.com/in/gasser-elbanna 🐦 @gasser_elbanna

EDUCATION

Ph.D. in Speech and Hearing Bioscience and Technology (SHBT)

📅 September 2023 – ongoing

Harvard University and MIT, USA

- Advisor: Prof. Josh H. McDermott (MIT, USA).

MSc. in Neuroscience & Neuro-engineering 📅 September 2020 – April 2023

EPFL, Switzerland

Average Grade: 5.7/6.0 (mention d'Excellence/with High Distinction)

- Thesis Title: Evaluating Speaker Identity Coding in Self-supervised Models and Humans. [🔗](#)
- Thesis Advisors: Dr. Satrajit S. Ghosh (Harvard University/MIT, USA) and Dr. Antoine Bosselut (EPFL, Switzerland).

BSc. (Honors) in Systems and Biomedical Engineering

📅 September 2015 – August 2020

Cairo University, Egypt

Grade: Distinction with Honors

- Thesis Title: Building Analytical Surface EMG Model for ALS Early Detection. [🔗](#)
- Thesis Advisors: Prof. Ayman M. Eldieb (Cairo University, Egypt) and Prof. Sherif Elbasiouny (Wright State University, USA).

WORK EXPERIENCE

Speech ML Research Intern 📅 April 2023 – August 2023

IDIAP Research Institute

📍 Martigny, Switzerland

- Studying the relation between speech signal and heart activity.
- Identifying the salient acoustic features for predicting heart activity.
- Training CNN-based neural networks to predict heart activity (BPM & HRV) from raw speech.
- Benchmarking handcrafted and self-supervised speech features on predicting heart activity features from speech.
- This work yielded a paper in *Interspeech 2024*.

Graduate Research Student | Bertarelli Fellow 📅 March 2022 – February 2023

MIT/Harvard Medical School

📍 Cambridge, MA, USA

- Exploring the invariances and equivariances of self-supervised speech models on speaker identity-related tasks.
- Conducting behavioral experiments using **GORILLA** to evaluate the performance of humans and models on a speaker discrimination task.
- Identifying the brain regions best-predicted by self-supervised models using a **naturalistic fMRI data**.
- This work yielded multiple invited talks at *BCS/MIT*, *CSAIL/MIT* and *SHBT/Harvard* in addition to three poster presentations at *NeurIPS*, *Bridge2AI*, and *OHBM* as well as a journal paper in-prep.

Voice AI Intern 📅 August 2021 – February 2022

Logitech Europe SA

📍 EPFL Innovation Park, Switzerland

- Improving a self-supervised speech model (BYOL-S) by designing a hybrid training protocol to learn from data-driven and handcrafted features simultaneously (**Hybrid BYOL-S**) using **PyTorch Lightning**.
- Benchmarking speech models (e.g. BYOL-A, TRILL, YAMNET, VGGish,...etc) on voice stress detection tasks (Cognitive & Physical Load).
- Exploring hyperbolic representational spaces and its performance on speech emotion recognition tasks.
- This work yielded two papers in *Interspeech* and *PMLR* as well as a model ranked in the top 3 at *NeurIPS HEAR competition*.

ML & Data Visualization Research Assistant 📅 March 2021 – October 2021

Machine Learning and Optimization Laboratory

📍 EPFL, Switzerland

- Detecting and visualising patterns in medical data to guide targeted interventions and medical training (Epidemiology).
- Implementing supervised and unsupervised anomaly detection ML Models for the **Dynamic Project** and using **Tableau** as a web-based dashboard development tool for visualization integrated with **Python** scripts to run ML models.
- This work yielded a paper in *Emerging microbes & infections* with title **Blood virosphere in febrile Tanzanian children**.

Computer Vision Intern 📅 May 2020 – August 2020

Advintic

📍 Cairo, Egypt

- Training a U-Net based architecture to detect and segment main heart coronaries using **Keras with TensorFlow**.

Research Intern 📅 August 2019 – October 2019

Opto-Nano-Electronics Lab

📍 Cairo University, Egypt

- Building a text to speech keyboard for autistic children by installing Linux image on a **Raspberry Pi** and using an open source TTS client (**Festival**) to automate the process of speech generation.

ACHIEVEMENTS & AWARDS

🏆 Awarded Best Poster in the NIH **Bridge2AI** Voice Symposium 2024

📅 May 2024

- Best Student Poster in the category of "Current Technological or Methodological Barriers to Clinical Use".

🏆 **Top 3%** Paper Recognition at ICASSP 2023 📅 June 2023

- Paper with title "Efficient Speech Quality Assessment using Self-supervised Framework Embeddings".

🏆 Nominated for Best Masters Project in Life Sciences Engineering Program at EPFL

📅 May 2023

- Masters project with title "Evaluating Speaker Identity Coding in Self-supervised Models and Humans".

🏆 Awarded Logitech Publication Prize 📅 July 2022

- Received 1,000 CHF to attend and present at Interspeech 2022 Conference. Paper with title "Hybrid Handcrafted and Learnable Audio Representation for Analysis of Speech Under Cognitive and Physical Load".

🏆 HEAR Competition at NeurIPS 2021 📅 December 2021

- **Ranked 1st** on LibriCount task (9% improvement) and Ranked 3rd overall (19 downstream audio tasks).

🏆 Awarded **Bertarelli Fellowship** in Translational Neuroscience and Neuro-engineering

📅 February 2021

- An EPFL-Harvard Medical School one-year fellowship to carry out a Masters thesis in **Sensible Intelligence lab**.

🏆 3D Printed motor neuron registered at ModelDB

📅 April 2020

- 3D printing a cat motor neuron (**vemoto6 Neuron Model**) using **NeuroMorphoVis** as part of bachelor's thesis.

PROJECTS

Me Too Quotes Analysis 📅 September 2021 – December 2021

Course Project at Data Science Lab

- Analyze **Quotebank data** in addition to **twitter dataset** to study the impact of traumatic/non-traumatic incidents on resurrecting the #MeToo movement using NLP in **Python**.
- Build a web **blog** with the data story to illustrate the results.

Predict Breathing Patterns from Speech 📅 July 2021 – September 2021

Semester Project at IDIAP

- Train a CNN-based model using **Pytorch** for estimating breathing patterns from voice samples.
 - Experiment with different model architectures, loss functions and hyper-parameters to optimize performance.
-

Learning Adaptive Behavior Through Competition 📅 July 2021 – September 2021

Semester Project at Mathis Group for Computational Neuroscience and AI

- Design a training procedure which allows an agent to succeed in a progressively larger and more complex set of environments by implementing **PAIRED** algorithm.
 - Changing dynamics due to environmental perturbations and generating unsupervised curriculum for adaptation using **RLlib**.
-

Impact of Motivation on Performance and Neuronal Activity in Mice Engaged in a Sensory Detection Task 📅 February 2021 – June 2021

Semester Project at Laboratory of Sensory Processing

- Analyze behavioral parameters (Engagement, Performance and Cumulative Reward) and Psychometric functions in mice whisker-deflection detection task.
 - Analyze neural parameters (Firing Rate and PCA) recorded from S1, mPFC and tJM1 brain regions.
 - Correlation analysis between neural and behavioral parameters.
-

Applying VoxelMorph Framework to C. Elegans Brain Data for image registration

📅 October 2020 – December 2020

Course Project at Laboratory of Physics of Biological Systems

- Apply image registration on 3D volumes of brain data in **TensorFlow**.
 - Create a deformation field for each 3D volume in a specific time frame relative to first frame.
-

Analytical Surface EMG Model connected to Motoneuron Model for ALS Early Detection

📅 August 2019 – August 2020

BSc. Graduation Project

- Building a motoneuron model using **NEURON** simulating early ALS biophysical features and a sEMG model using **Python**.

Computer Vision GUI 📅 May 2020

- Building user-friendly GUI to implement Hough Transform, Harris Corner Detector, Template Matching and SIFT Algorithms on given images using **openCV** and **PyQt5**.
-

Mini Autonomous Car 📅 October 2019

- Building a self-driving car which detects lanes using **OpenCV** (Hough transform & Contouring).
 - Detecting obstacles using ultrasonic sensor connected with **Arduino** that overrides the steering control in case avoiding obstacles.
-

Wireless Data Transfer 📅 September 2019

- Generating pseudo random numbers that simulate patient data and transfer it wirelessly to a server using **BLE chip**.
 - Visualizing the data acquired from the server in a web app developed using **Django** to simulate real-time vital signal tracking.
-

Volume Rendering Application for Head and Ankle Images

📅 April 2019

- Loading DICOM images for ankle and head then apply Surface Rendering using adjustable ISO value and Ray Cast Rendering using adjustable transfer function using **VTK** and **Qt Designer**.
-

MRI Simulator Software 📅 March 2019

- Implement a generalized MRI simulator with the preparation sequences (IR, T2 Prep. and Tagging) and pulse sequences (GRE, SSFP and SE) in a GUI using **PyQt5**.
- Implement a computational shepp-logan for testing and validation.

PUBLICATIONS AND TALKS

Journal and Conference Publications

- **Elbanna, G.**, Mostaan, Z., & Magimai Doss, M. (2024). Predicting Heart Activity from Speech using Data-driven and Knowledge-based features. In Proc. Interspeech.
- El Hajal, K., Wu, Z., Scheidwasser-Clow, N., **Elbanna, G.**, & Cernak, M. (2023). Efficient Speech Quality Assessment Using Self-Supervised Framewise Embeddings. In ICASSP 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1-5). IEEE.
- **Elbanna, G.**, Scheidwasser-Clow, N., Kegler, M., Beckmann, P., El Hajal, K., & Cernak, M. (2022). Byol-s: Learning self-supervised speech representations by bootstrapping. In HEAR: Holistic Evaluation of Audio Representations (pp. 25-47). PMLR.
- **Elbanna, G.**, Biryukov, A., Scheidwasser-Clow, N., Orlandic, L., Mainar, P., Kegler, M., ... & Cernak, M. (2022). Hybrid handcrafted and learnable audio representation for analysis of speech under cognitive and physical load. In Proc. Interspeech (pp. 386-390).
- Cordey, S., Laubscher, F., Hartley, M. A., Junier, T., Keitel, K., Docquier, M., ... **Elbanna, G.**, Tapparel, C., Zanella, M., Xenarios, I., Fellay, J., D'Acremont, V., & Kaiser, L. (2021). Blood virosphere in febrile Tanzanian children. *Emerging microbes & infections*, 10(1), 982-993.

Pre-prints and In prep Papers

- Dixit, S., Low, D., **Elbanna, G.**, Catania, F., & Ghosh, S. (2024). Explaining Deep Learning Embeddings for Speech Emotion Recognition by Predicting Interpretable Acoustic Features. Under Review at ICASSP 2025.
- *Alavilli, S., *Banerjee, A., ***Elbanna, G.**, *Magaro, A. Incorporating Talker Identity Aids With Improving Speech Recognition in Adversarial Environments. Under Review at ICASSP 2025. **Equal Contribution*
- **Elbanna, G.**, Catania, F., & Ghosh, S. Investigating Speaker Identity Coding in Speech Artificial Neural Networks. In prep.

Conference Abstracts

- **Elbanna, G.**, & McDermott, J. (2025). Human-like Speech Perception Emerges in Models Optimized on Phoneme Recognition. In ARO 2025.
- **Elbanna, G.**, & McDermott, J. (2024). Artificial Neural Networks Explain Continuous Speech Perception in Humans. In NeurIPS 2024 UniReps Workshop.
- **Elbanna, G.**, & McDermott, J. (2024). Modeling Continuous Speech Perception Using Pseudo Supervised Learning. In SANE Meeting 2024.
- **Elbanna, G.**, Catania, F., & Ghosh, S. (2024). Speaker Identity Coding in Speech Artificial Neural Networks. In Voice AI Symposium Bridge2AI 2024.
- Catania, F., **Elbanna, G.**, & Ghosh, S. (2024). The Voice Anonymization Challenge: Achieving Privacy without Compromising Utility. In Voice AI Symposium Bridge2AI 2024.
- **Elbanna, G.** & Ghosh, S. (2024). Predicting Brain Responses in Auditory and Language Regions using Speech Self-supervised Models. In Organization for Human Brain Mapping (OHBM) 2024.
- **Elbanna, G.**, Catania, F., & Ghosh, S. (2023). Towards Understanding Speaker Identity Coding in Data-driven Speech Models. In NeurIPS 2023 MusiML Workshop.

Invited Talks

- *Investigating Speaker Identity Representations in Artificial Neural Network Models*. Contributed Talk at VoiceID Conference, Marburg (2024). [!\[\]\(8d139a66f540002704b5c70b7fe6cc7a_img.jpg\)](#)
- *The Voice as a Window to The Mind: Opportunities and Challenges*. Invited Talk at SLS Group, CSAIL, MIT (2024). [!\[\]\(c209541a4bc5f45e44bd7791f9477320_img.jpg\)](#)
- *Towards Understanding Speaker Identity Coding in Data-driven Speech Models*. Spotlight Talk at MusiML workshop at NeurIPS (2023). [!\[\]\(8fd54d112e752061b5361c5bdf346185_img.jpg\)](#)
- *Learning Self-supervised Speech Representations via Hybrid Training*. Invited Talk at Pindrop Company (2023). [!\[\]\(3525fd0bd3680f905a850c70520e38c7_img.jpg\)](#)
- *Speaker Identity Coding in Self-supervised Models*. CogLunch Talk at BCS MIT (2023). [!\[\]\(3c3fba180f5a473bd1cb3c114e029235_img.jpg\)](#)
- *Speech Processing Lecture*. SHBT-200 graduate course at Harvard (2022). Co-lectured with Dr. Satrajit S. Ghosh. [!\[\]\(e4e94fc8df185dfabcff49964cc4dd39_img.jpg\)](#)
- *What do Machines Hear? Overview of deep learning approaches for representing voice*. Invited Talk at Harvard-MIT Speech Biomarker Group (2022). [!\[\]\(fdd6f407d8c5866c7bd62e139ad13d49_img.jpg\)](#)
- *SERAB BYOL-S Model*. HEAR Competition Submission Talk at NeurIPS (2021). [!\[\]\(76458e872a39caac300a32463fa679ff_img.jpg\)](#)

Blogs

- *Discrimination in Artificial Intelligence for Voice Applications*. [!\[\]\(9bfa69b6b0f097b09744337d04f22d78_img.jpg\)](#)

EXTERNAL SERVICE

Reviewer for **MENA ML Winter School** 📅 November 2024

- Review applicants from the Middle East and North Africa regions for a machine learning winter school sponsored by Google Deepmind, DELL, and QCRI.

Core Organizer of **Muslims in ML Affinity Workshop** at NeurIPS

📅 August 2024-ongoing

- Organizing an affinity workshop dedicated to amplifying the voices of Muslim researchers in the fields of machine learning and artificial intelligence and addressing challenges and research topics that are particularly relevant to Muslims.

B.Sc. Thesis Advisor at Cairo University 📅 August 2024-ongoing

- Advising an undergraduate thesis studying the invariance problem in speech and voice perception.

Mentor at **Fatima Fellowship** 📅 July 2024-ongoing

- Mentoring a Fatima Fellow on studying voice similarity judgments in humans and artificial neural networks.








Co-leading ISWG at Harvard Union 📅 September 2023-ongoing

- Co-leading the International Scholar Working Group (ISWG) at HGSU which involves advocating for issues pertaining international students and workers at Harvard University.

Non-resident Tutor at Harvard Housing 📅 September 2023-ongoing

- NRT at the Leverett House at Harvard which involves tutoring undergraduate students with their graduate applications, organizing scientific activities, and fostering an inclusive environment for all students in the house.

SKILLS

 Technical Development Python and MATLAB	 Graphics and Visualization OpenGL, VTK and Tableau
 Desktop Development C and C++	 Embedded Systems Raspberry Pi, ESP and Arduino
 Deep Learning Frameworks Pytorch Lightning, Pytorch, Tensorflow, Keras, RLib, Ray and VoxelMorph	 Miscellaneous Git, L ^A T _E X, Linux, Qt Designer, fMRIPrep, Prolific, and GORILLA
 Modeling NEURON, NMODL and HOC Language	