Via Cantarana 30
Ferrara, Italy
[Website]

Sankar Mukherjee

+39.329.661.9542
sankar1535@gmail.com
[Linkedin] [Github]

Education

- **PhD. in Bioengineering and Robotics**, CTNSC, Istituto Italiano di Tecnologia, Università degli studi di Genova, Italy
  Thesis Title: Sensorimotor Processes in Speech Listening and Speech-based Interaction **Nov 2015 - March 2019**
- **M.S. in Speech Technology**, Indian Institute of Technology Kharagpur, India
  Thesis Title: Modification of F0 contour in Bengali HMM based Speech Synthesis System **Aug 2012 - May 2014**
- **B.Tech** Electronics and Communication Engineering, Jalpaiguri Govt. Engg. College, India **Aug 2005 – May 2009**

Work Experience

Speech Scientist **DefinedCrowd, Lisbon** **September 2019 - Present**

- Research and prototype new approaches and develop new algorithms using ML techniques that can be exploited to clean noise from speech signals.
- Work with other data scientists, engineers, PMs to provide valuable insights from system telemetry.
- Mine data sets with Spark, SQL, Cosmos from disparate sources in complex data pipelines.
- Transform data into innovative features/signals that can improve a machine-learning task.
- Communicate final recommendations and drive decision making.
- Implement and deploy features into production.

Postdoc Researcher **Istituto Italiano di Tecnologia, Italy** **April 2019 - August 2019**

- Grant writing for a research project titled 'Neuro-behavioural aware conversational agent' which win 10th Christian Benoît Award.
- Collaborate with senior researchers and engineers to build a mobile app which can detect coordination between speakers.
- Design an experiment to explore behavioral-underpinnings between two speakers during conversation.
- Research novel data fusion techniques which combines acoustic (e.g. back-channels) and visual cues (e.g. head nodding) during conversation.

Research Assistant **Istituto Italiano di Tecnologia, Italy** **June - Oct 2015**

- Collaborate with Senior researchers and Professors between Aix Marseille université (AMU) and Istituto Italiano di Tecnologia (IIT) for a project entitled SPIC.
- One of the aim of SPIC is to measure speaker similarity during a conversation also known as phonetic convergence.
- Created an algorithm with GMM-UBM which captures speakers phonetic convergence during conversation based on bi-syllabic words.
- Created a Skype plugin which triggers via speech onset and used as communication software between AMU and IIT.

Research Assistant **Laboratoire Parole et Langage, France** **Sept 2014 - May 2015**

- Performance analysis and comparison of various ML pipelines (PCA, LDA, SVM, SGD, CART, RandomForest) for classification of conversation feedback (linguistic and acoustic) using two french dialog databases.
- Worked with senior researchers and technicians to measure trajectories of 3D sensor coils inside Electromagnetic Field.

Software Engineer **Yantra Software, India** **May - July 2014**

- Built text-to-speech synthesis system (English, Hindi, Telegu) for Indian Banks in their customer service module.
- Text-to-speech synthesis system was Speech Application Programming Interface (SAPI) compatible and built using HMM based speech synthesis system (HTS) with Asterisk with UniMRCP integration.

Project Engineer **Indian Institute of Technology Kharagpur, India** **April 2011 - April 2012**

- Created a web-toolkit which generates VoiceXML formatted Pronunciation Lexicon Specification (PLS) for Indian languages. PLS can be referenced from other markup languages, such as the Speech Recognition Grammar Specification (SRGS) and the Speech Synthesis Markup Language (SSML) which can be used by TTS or ASR systems.
- Lead a team of two students and one linguist and presented the results in Ministry of Human Resource Development, New Delhi, India. This work won the best paper award in Oriental COCOSDA 2013.

Skills

- **Currently using:** Python, PySpark, MySQL, PyTorch, Scikit-learn, HTK, Azure, Docker.
- **Previously used:** Java, Matlab, JavaScript, JSP, C.

Research Projects

BrainDecode [Link] **Istituto Italiano di Tecnologia, Italy** **Sept 2018 - July 2019**

Leveraging deep neural network architecture to find complex speech representation in brain.

- Acoustic-to-articulatory inversion mapping was done using auto-encoder and LSTM recurrent neural network.
- Each hidden layers complex information were used to find its relation to brain.
- **Keywords:** Auto-encoder, LSTM, Acoustic-to-articulatory inversion.
- **Coordinator:** Prof. Alessandro D'ausilio

NeuroLip [Link] **Istituto Italiano di Tecnologia, Italy** **Oct 2017 - Sept 2018**

Disentangling auditory (bottom-up) and motor (top-down) contributions to speech processing during speech listening.

- 25 subjects listened to speech stimulus (from MSPKA corpus) with varying speech rate while their EEG was recorded.
- Speech envelop and multiple vocal tract features (i.e. lip opening, tongue vertical movement etc.) were extracted from stimulus and used to find relation between speech and vocal tract kinematics with EEG.
- **Keywords:** Speech processing, Kinematic signal processing, Phase connectivity.
- **Coordinator:** Prof. Alessandro D'ausilio

SPIC: [Link] **Laboratoire Parole et Langage, France** **Nov 2016 - August 2017**

Understanding cerebral, articulatory and acoustic convergence between speakers in conversational interaction.

- 5 couples participated in verbal interaction task while their EEG, Speech and Articulatory (vocal tract kinematics) data were recorded.
- During joint coordination in conversation speakers articulatory and neural changes were captured and characterized.
- **Keywords:** Brain-to-brain coupling, Speaker style shifting.
- **Coordinator:** Prof. Noël Nguyen

Hyperscanning [Link] **Istituto Italiano di Tecnologia, Italy** **Sept 2015 - Oct 2016**

Exploration of neural markers during two persons verbal interaction.

- Created a verbal interaction task which was engaging as well as flexible towards natural conversation.
- 10 couples played a verbal domino game while their EEG was recorded.
- The time points of speech acoustics coordination during conversation was captured and corresponding neural markers were identified.
- **Keywords:** Dual EEG, Mutual adaptation, Verbal interaction.
- **Coordinator:** Prof. Alessandro D'ausilio

ProsodyGen [Link] **Indian Institute of Technology Kharagpur, India** **Aug 2012 - May 2014**

Statistical modeling of HMM synthesized speech prosody.

- Modelled intonation of text-to-speech synthesis system with supervised and semi-supervised modelling using:
  - Fujisaki F0 Generation Process model.
  - Deep Boltzmann Machine (DBM) and Twin Gaussian Process (TGP) hybrid model.
- Improved Part-of-Speech tagging in front-end of text-to-speech synthesis system using Global Linear Model (GLM).
- **Keywords:** Intonation modelling, Speech synthesis, Part-of-Speech tagging.
- **Coordinator:** Prof. Shyamal kumar das mandal

Selected Publications

- **S. Mukherjee**, A. Tomassini, L. Badino, A. Pastore, L. Fadiga, A. D'Ausilio, *"Cortical tracking of speech reveals top-down reconstructive processes"* **Submitted eLife**
- **S. Mukherjee**, L. Badino, P. Hilt, A. Tomassini, A. Inuggi, L. Fadiga, N. Nguyen, A. D'Ausilio, *"The neural oscillatory markers of phonetic convergence during verbal interaction"* **Human Brain Mapping**, 2018

- **S. Mukherjee**, T. Legou, L. Lancia, P. Hilt, A. Tomassini, L. Fadiga, A. D'Ausilio, L. Badino, N. Nguyen, *"Analyzing vocal tract movements during speech accommodation"* **Interspeech**, 2018
- **S. Mukherjee**, A. D'Ausilio, N. Nguyen, L. Fadiga, L. Badino, *"The Relationship Between F0 Synchrony and Speech Convergence in Dyadic Interaction"* **Interspeech**, 2017
- Prévot, L., J. Gorisch, **S. Mukherjee** *"Annotation and Classification of French Feedback Communicative Functions"*, **PACLIC**, 2015
- **S. Mukherjee**, S. K. Das Mandal, *"Generation of F0 Contour using Deep Boltzmann Machine and Twin Gaussian Process Hybrid Model for Bengali Language"* **Interspeech**, 2014
- **S. Mukherjee**, S. K. Das Mandal, *"PL-ILT: A Web Tool for Creation of Pronunciation Lexicon in Indian Languages"* **OCOCOSDA**, 2013 (**Best Student Paper**)

Additional Experience and Awards

- **Awards:**
  2019: Winner of 10th Christian Benoît Award for research project on "Neuro-behavioral Aware Conversational Agent"
  2014: Winner of ISCA Grant recipient for Interspeech
  2013: Best Student Paper Award in Oriental COCOSDA
- **Kaggel:** DecMeg 2014 (Top 29%), Home Credit Default Risk (Top 11%), Freesound General-Purpose Audio Tagging Challenge
- **Languages:** English, Hindi, Bengali (Native), Italian (intermediate), French (beginner).

References

- **Prof. Alessandro D'ausilio**, Associate Professor, Human Physiology @ Università di Ferrara.
  Email: Alessandro.dausilio@iit.it
- **Dr. Leonardo Badino**, Senior Researcher at PerVoice SpA.
  Email: leobad08@gmail.com
- **Prof. Noël Nguyen**, Professor at Aix-Marseille University
  Email: noel.nguyen-trong@univ-amu.fr
- **Prof. Shyamal Kumar Das Mandal**, Assistant Professor at Indian Institute of Technology Kharagpur
  Email: sdasmandal@cet.iitkgp.ernet.in