Modeling language from raw speech with GANs

Gasper Begus, PhD
Assistant Professor
Dept. Linguistics,
UC Berkeley, USA

In this talk, I propose that language can be modeled from raw speech data in a fully unsupervised manner with Generative Adversarial Networks (GANs) and that such modeling has implications both for the understanding of language acquisition and for the understanding of how deep neural networks learn internal representations. I propose an extension of the GAN architecture (fiwGAN) in which meaningful linguistic properties emerge from two networks learning to encode and decode information. FiwGAN captures the perception-production loop of human speech and, unlike most other deep learning architectures, has traces of communicative intent. I further propose a technique to identify latent variables in deep convolutional networks that represent linguistically meaningful units in a causal, disentangled, and interpretable way. We can thus uncover symbolic-like representations at the phonetic, phonological, syntactic and lexical semantic levels, analyze how learning biases in GANs match human learning biases in behavioral experiments, how speech processing in the brain compares to intermediate representations in deep neural networks, and what GANs' innovative outputs can teach us about productivity in human language.

the CHAI talk series

Venue: Online*
Time: 10 AM IST
13th Sep 2023

About the Speaker

Gasper Begus is an Assistant Professor at the Department of Linguistics at UC Berkeley where he directs the Berkeley Speech and Computation Lab. He is also the Linguistics Lead at Project CETI and a Member of Berkeley's Institute of Cognitive and Brain Sciences. Previously, he was an Assistant Professor at the University of Washington. Before that, he graduated with a Ph.D. from Harvard. His research focuses on developing deep learning models for speech data. Gasper combines machine learning and statistical models with neuroimaging and behavioral experiments to better understand how deep neural networks learn internal representations and how humans learn to speak.

Organized by

Mehta Family **School of Data Science & Artificial Intelligence**
Indian Institute of Technology Guwahati

*for link e-mail: mfsdsai_off@iitg.ac.in