

## Columbia Research Exploring How AI Can Hear One Voice Among Many Receives Honor

*Engineering society gives award to impactful, pioneering work that could help lead to brain wave-reading hearing aids*

NEW YORK, 25 January, 2022 – NEW YORK – For their groundbreaking research that opened new directions on how artificial intelligence software can better mimic the human ability to hear one voice among many in a crowd, [Nima Mesgarani](#), PhD, a principal investigator at Columbia's Zuckerman Institute and scientist [Yi Luo](#), his former doctoral student, have received the 2021 IEEE Signal Processing Society Best Paper Award.

This award honors the authors of a paper of exceptional merit published in the last six years dealing with a subject related to the IEEE [Signal Processing Society](#)'s technical scope and is judged based on general quality, originality, subject matter and timeliness. The prize awards Mesgarani and Luo each \$500 and a certificate.

"The IEEE Signal Processing Society Best Paper Award is given to a paper that has passed the test of time and has had a proven impact in the field," Dr. Mesgarani said. "The recognition for this paper shows that our study was able to open new research directions for the community and pave a new path forward in the difficult task of automatic speech separation."

Dr. Mesgarani, who is also an associate professor of electrical engineering at Columbia, and Dr. Luo, currently a senior research scientist at Tencent, received this award for their 2019 study in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, titled "[Conv-TasNet: Surpassing Ideal Time–Frequency Magnitude Masking for Speech Separation](#)."

This highly cited paper focused on the challenging task of automatic speech separation, also known as [the cocktail party problem](#). People without hearing impairments are usually extremely good at focusing on who and what we want to hear at cocktail parties and crowded scenes while filtering out background noise. However, this task remains difficult for machines.

The paper described a novel solution for the cocktail party problem, a kind of AI known as a neural network, which mimics the structure of neurons in the biological brain. The AI, called the convolutional time-domain audio separation network, or Conv-TasNet, encoded recordings of more than one person speaking at once into a digital form of data that was optimally designed to help the neural network distinguish between multiple speakers. It then separated out each individual and converted each data stream back into audio form.

# COLUMBIA | Zuckerman Institute

## MORTIMER B. ZUCKERMAN MIND BRAIN BEHAVIOR INSTITUTE

Conv-TasNet significantly outperformed what was previously thought to be the upper possible accuracy limit for neural networks on automatic speech separation, using an entirely different strategy than conventional methods. Conv-TasNet was also substantially smaller and faster than previous neural networks, making it useful for real-time automatic speech separation applications in wearable devices.

"Our study represents a conceptual advance in the field," Dr. Mesgarani said. "Instead of using traditional ways of representing an audio signal, we proposed a different way that removes many inherent problems of the old method."

One potential application of this work is a better [cognitive hearing aid](#), a device designed specifically to help solve the cocktail party problem. Dr. Mesgarani published details about this invention in [2019](#) and was also among an interdisciplinary group of researchers that received [the 2021 Misha Mahowald Prize for Neuromorphic Engineering](#).

Conventional hearing aids work by augmenting all sounds at once, producing a cacophony of noise for device wearers in crowded scenes that can make it extraordinarily difficult for the hearing-impaired to follow or participate in conversations. In contrast, cognitive hearing aids automatically separate out the voices of multiple speakers in a group and then compare each voice to the brain waves of the person wearing the hearing aid. The speaker whose voice pattern most closely matches the listener's brain waves, a sign that this is the person that the listener is most interested in, is amplified over the others. Dr. Mesgarani first detailed this method to deduce which sounds the brain listens to in a study in [2012](#).

"Receiving this prestigious award is very humbling and exciting at the same time, and it reaffirms the insights we have had on how this problem should be approached," Dr. Mesgarani said.

***Columbia University's Mortimer B. Zuckerman Mind Brain Behavior Institute brings together a group of world-class scientists and scholars to pursue the most urgent and exciting challenge of our time: understanding the brain and mind. A deeper understanding of the brain promises to transform human health and society. From effective treatments for disorders like Alzheimer's, Parkinson's, depression and autism to advances in fields as fundamental as computer science, economics, law, the arts and social policy, the potential for humanity is staggering. To learn more, visit: [zuckermaninstitute.columbia.edu](https://zuckermaninstitute.columbia.edu).***