

AI headphones let wearer listen to a single person in a crowd by looking at them just once

May 23 2024, by Stefan Milne and Kiyomi Taguchi



A University of Washington team has developed an artificial intelligence system that lets a user wearing headphones look at a person speaking for three to five seconds and then hear just the enrolled speaker's voice in real time even as the listener moves around in noisy places and no longer faces the speaker. Pictured is a prototype of the headphone system: binaural microphones attached to off-the-shelf noise canceling headphones. Credit: Kiyomi Taguchi/University of Washington

Noise-canceling headphones have gotten very good at creating an auditory blank slate. But allowing certain sounds from a wearer's environment through the erasure still challenges researchers. The latest edition of Apple's AirPods Pro, for instance, automatically adjusts sound levels for wearers—sensing when they're in conversation, for instance—but the user has little control over whom to listen to or when this happens.

A University of Washington team has developed an artificial intelligence system that lets a user wearing [headphones](#) look at a person speaking for three to five seconds to "enroll" them. The system, called "Target Speech Hearing," then cancels all other sounds in the environment and plays just the enrolled speaker's voice in real time even as the listener moves around in noisy places and no longer faces the speaker.

The team presented its findings May 14 in Honolulu at the [ACM CHI Conference on Human Factors in Computing Systems](#). The [code for the proof-of-concept device](#) is available for others to build on. The system is not commercially available.

"We tend to think of AI now as web-based chatbots that answer questions," said senior author Shyam Gollakota, a UW professor in the Paul G. Allen School of Computer Science & Engineering. "But in this project, we develop AI to modify the auditory perception of anyone wearing headphones, given their preferences. With our devices you can now hear a single speaker clearly even if you are in a noisy environment with lots of other people talking."

To use the system, a person wearing off-the-shelf headphones fitted with microphones taps a button while directing their head at someone talking. The [sound waves](#) from that speaker's voice then should reach the microphones on both sides of the headset simultaneously; there's a 16-degree margin of error. The headphones send that signal to an on-

board embedded computer, where the team's machine learning software learns the desired speaker's vocal patterns. The system latches onto that speaker's voice and continues to play it back to the listener, even as the pair moves around. The system's ability to focus on the enrolled voice improves as the speaker keeps talking, giving the system more training data.

The team tested its system on 21 subjects, who rated the clarity of the enrolled speaker's voice nearly twice as high as the unfiltered audio on average.

This work builds on the team's previous "semantic hearing" research, which allowed users to select specific sound classes—such as birds or voices—that they wanted to hear and canceled other sounds in the environment.

Currently the TSH system can enroll only one speaker at a time, and it's only able to enroll a speaker when there is not another loud voice coming from the same direction as the target speaker's voice. If a user isn't happy with the [sound quality](#), they can run another enrollment on the [speaker](#) to improve the clarity.

The team is working to expand the system to earbuds and hearing aids in the future.

Additional co-authors on the paper were Bandhav Veluri, Malek Itani and Tuochao Chen, UW doctoral students in the Allen School, and Takuya Yoshioka, director of research at AssemblyAI.

More information: Bandhav Veluri et al, Look Once to Hear: Target Speech Hearing with Noisy Examples, *Proceedings of the CHI*

Conference on Human Factors in Computing Systems (2024). DOI:
10.1145/3613904.3642057 , dl.acm.org/doi/10.1145/3613904.3642057

Provided by University of Washington

Citation: AI headphones let wearer listen to a single person in a crowd by looking at them just once (2024, May 23) retrieved 16 June 2024 from <https://techxplore.com/news/2024-05-ai-headphones-wearer-person-crowd.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.