CHiME-9 Task 2: ECHI
Enhancing Conversations for the Hearing Impaired

Organisers:

Robert Sutherland[1], Jason Clarke[1], Thomas Kuebert[3], Marko Lugger[3], Stefan Petrausch[3], Juan Azcarreta Ortiz[4], Buye Xu,[4] Stefan Goetze[1,2], **Jon Barker**[1]

[1] The University of Sheffield

[2] South Westphalia University of Applied Sciences

[3] WSAudiology

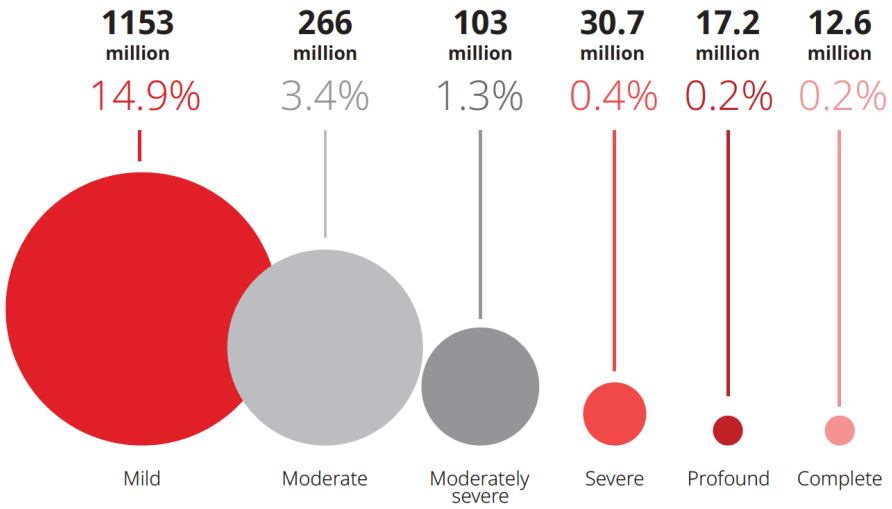
[4] Meta Reality Labs Research

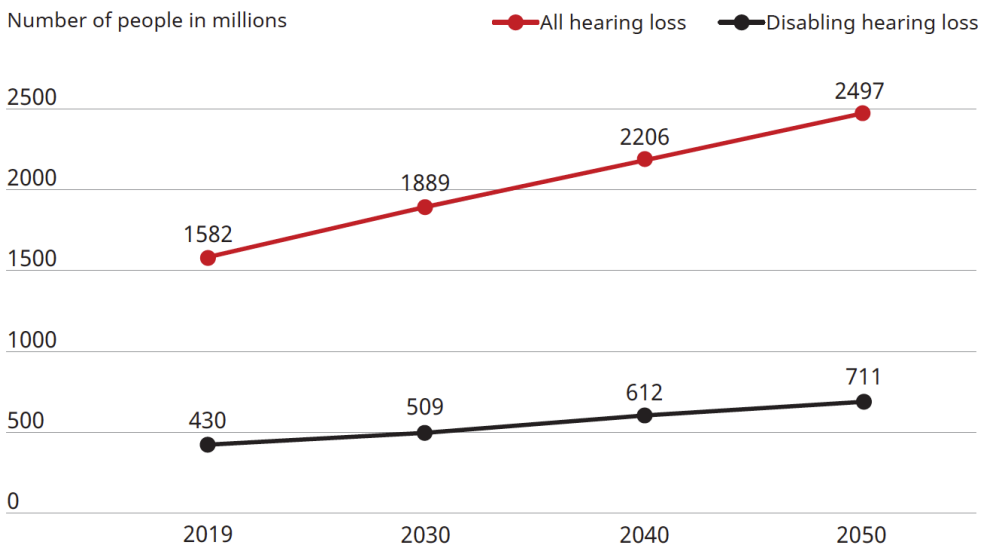
Hearing loss – a growing global problem

WHO Report on Hearing (2021)

- Over 1.5 billion people worldwide live with some degree of hearing loss.
- This number is projected to increase to 2.5 billion by 2050.
- Hearing loss has significant economic and social impacts on individuals and societies.

Hearing loss – a *growing* problem





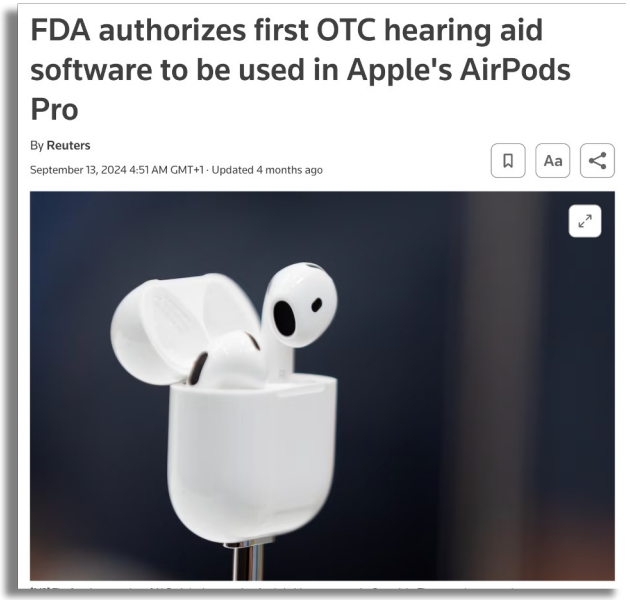
Hearing aids

- Highly sophisticated multi-channel digital signal processing.

- They can successfully restore audibility and speech intelligibility in many critical situations.

- But they often provide little benefit for speech in complex acoustic environments. Crucially, **hearing aid users struggle with conversation in noisy social gatherings**.

Entering a new era of DNN powered hearing devices…



The ECHI concept

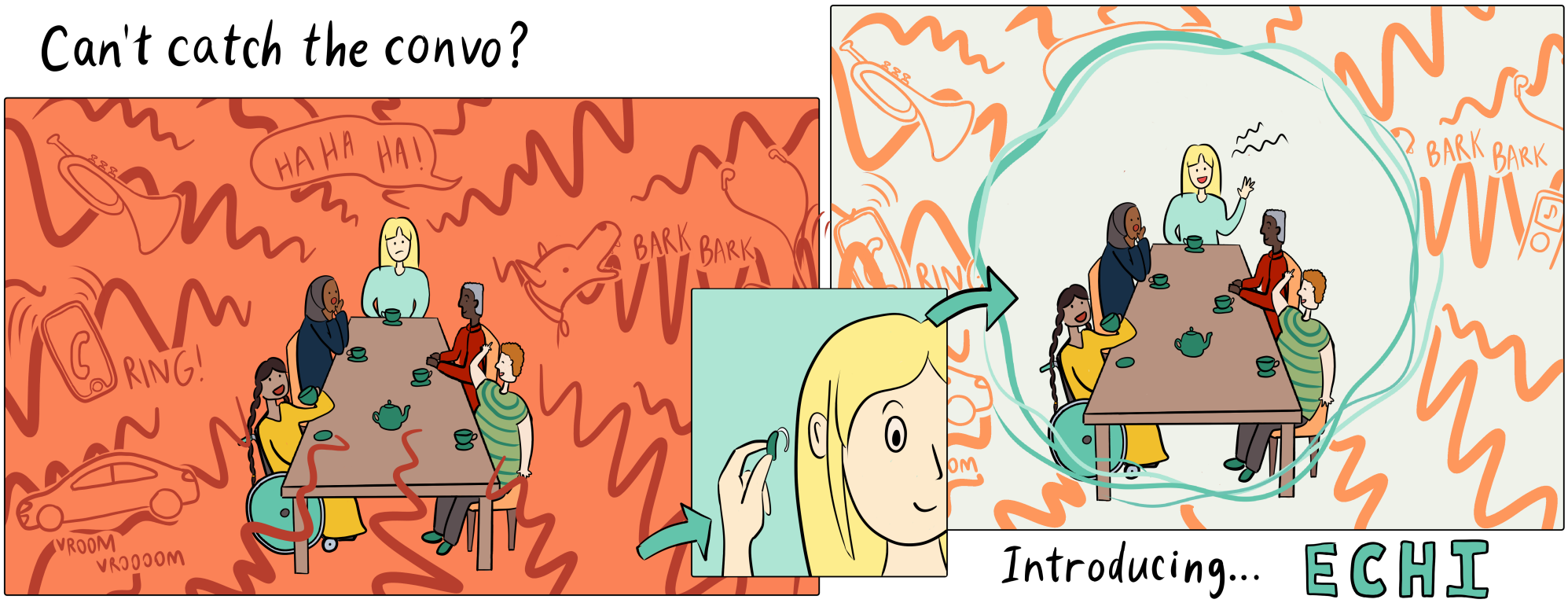
The experience

Hearing impairment makes the background sound overwhelming.

Conversation becomes hard or even impossible to follow.

Can't catch the convo?

Image credit: Tom Kit Barker / (c) 2025

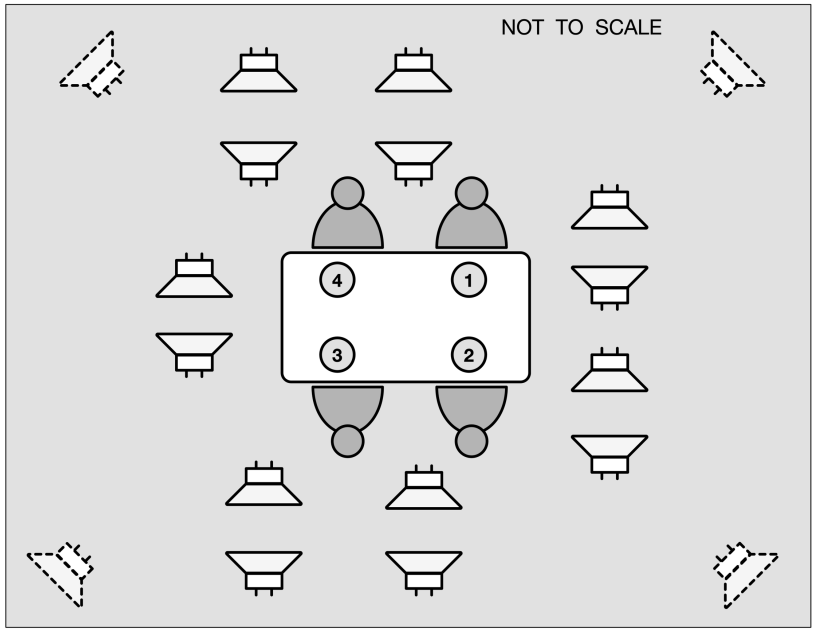


The solution

Real-time ‘conversation enhancement’ that suppresses all background sound sources while preserving the voices making up the conversation.

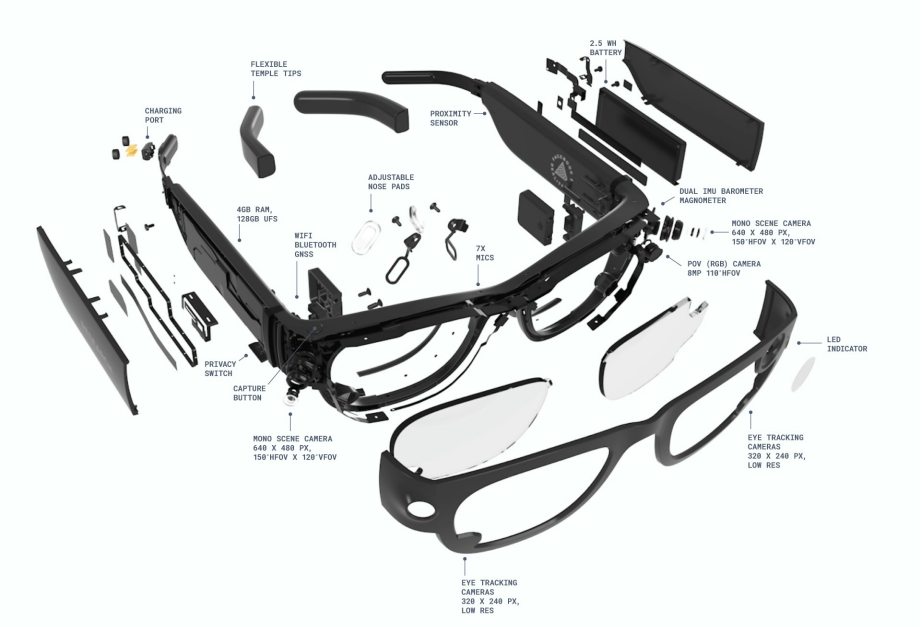
The ECHI Task

Recording Scenario

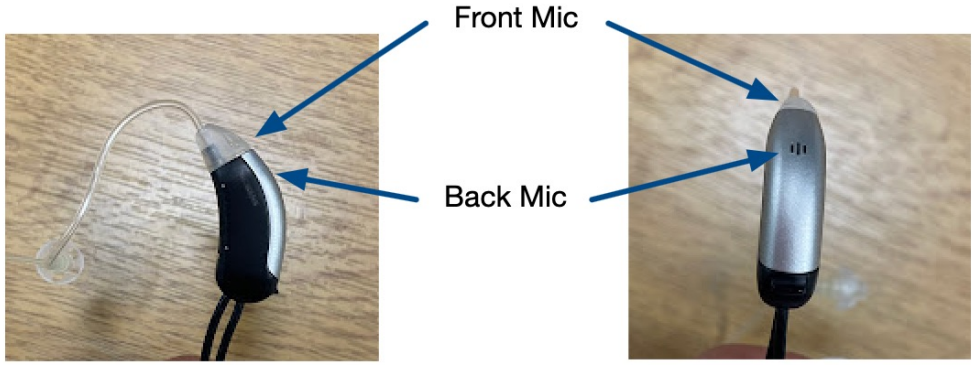


Recording Devices

Meta Aria Smart Glasses
(7 microphone channels)



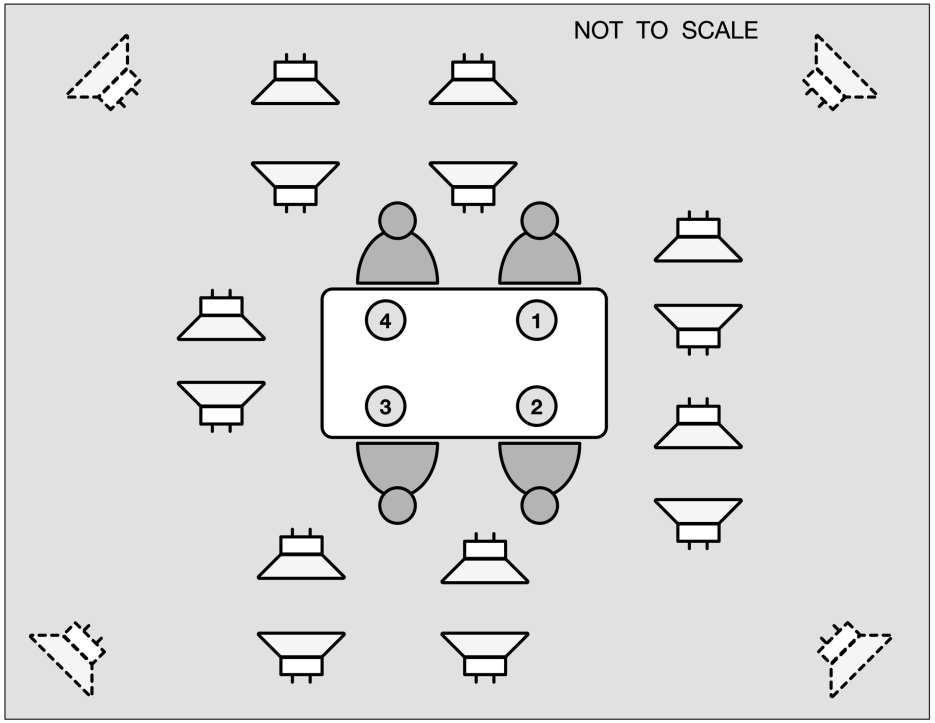
WSA Hearing aid shells
(4 microphone channels)



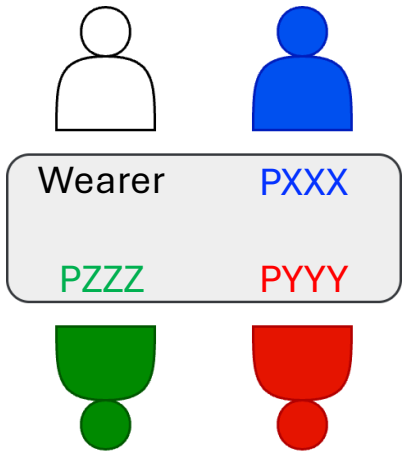
Recording Scenario

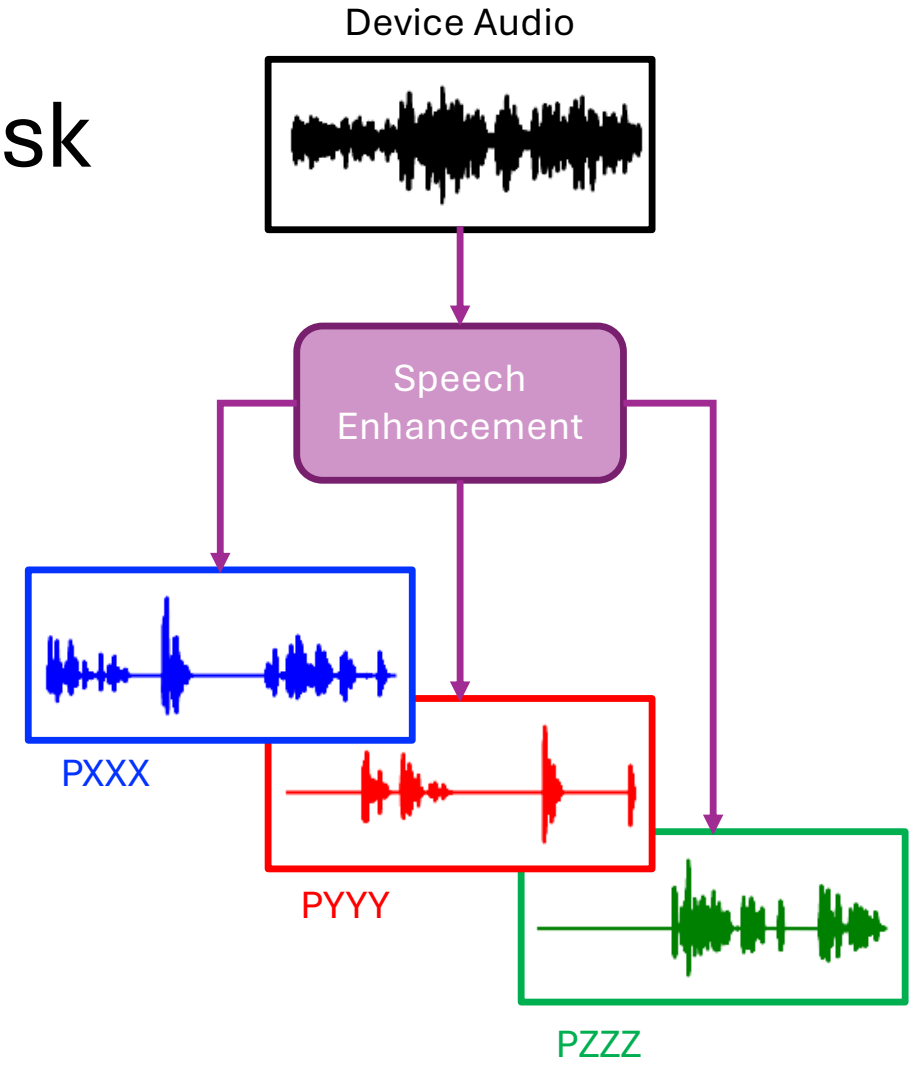
- 14 loudspeakers simulating people in a cafeteria.
  - Interfering speech sources and cafeteria-style sound effects
- 4 loudspeakers in the corners playing ambient WHAMR background noise
- 4 people in the middle having a conversation

Aria HA Ref



The Task





- For each recording session, take the multichannel device audio and extract the 3 conversation partners into separate channels.
- A clean speech sample of each participant is provided, i.e.,allowing target speaker extraction methods.
- Systems must have a maximum algorithmic latency of 20 ms.
- Participants can process the HA data, the Aria data or both. i.e. separate tracks.

The Data

Split	Sessions	Duration	Unique Speakers
Train	30	18 hours	118
Dev	10	6 hours	40
Eval	9	5.5 hours	36
Overall	**49**	**29 hours**	**194**

- Audio from hearing aids and Aria glasses.
- Close-talk microphone (can be used during training, and for objective evaluation).
- Motion tracking data for each participant (can be used during training).
- Clean samples of each participant's voice.

Evaluation

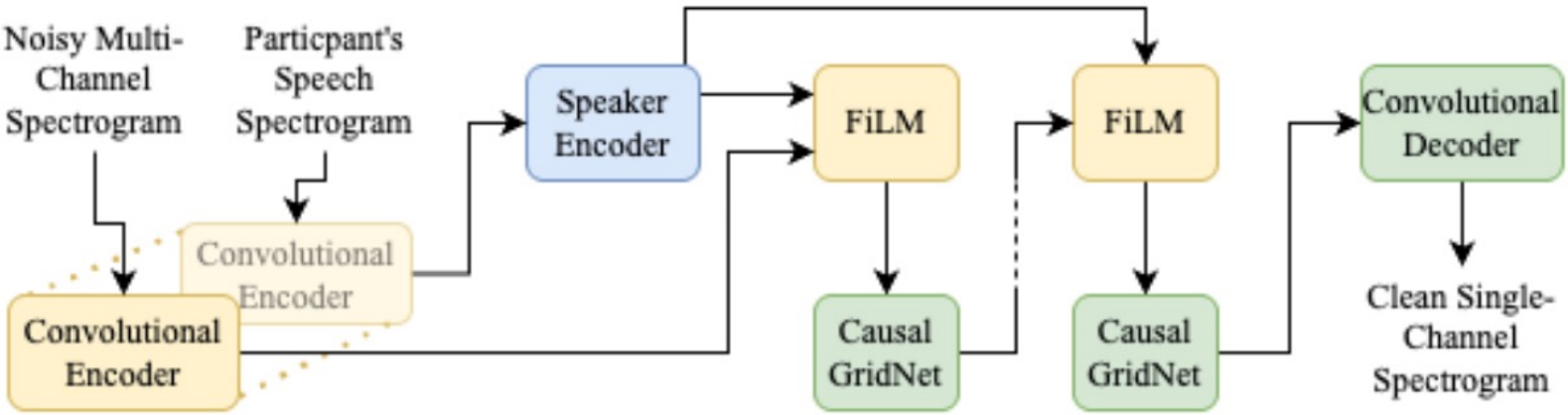
Evaluation

Preliminary Objective Evaluation

Final Subjective Evaluation

Baseline System

Based on TF-GridNet (Wang et al., TASLP, 2023)



The main adaptations from the base TF-GridNet model are:

- Expanding the encoder to accept multiple channels of audio
- The speaker encoder as implemented in Hao, Li and Zheng (2024)
- Feature-wise Linear Modulation (FiLM) to infuse the speaker information into the audio representation, as described in Cornell et al. (2023)
- Modifications to the GridNet blocks to make it a causal system, as described in Cornell et al. (2023).

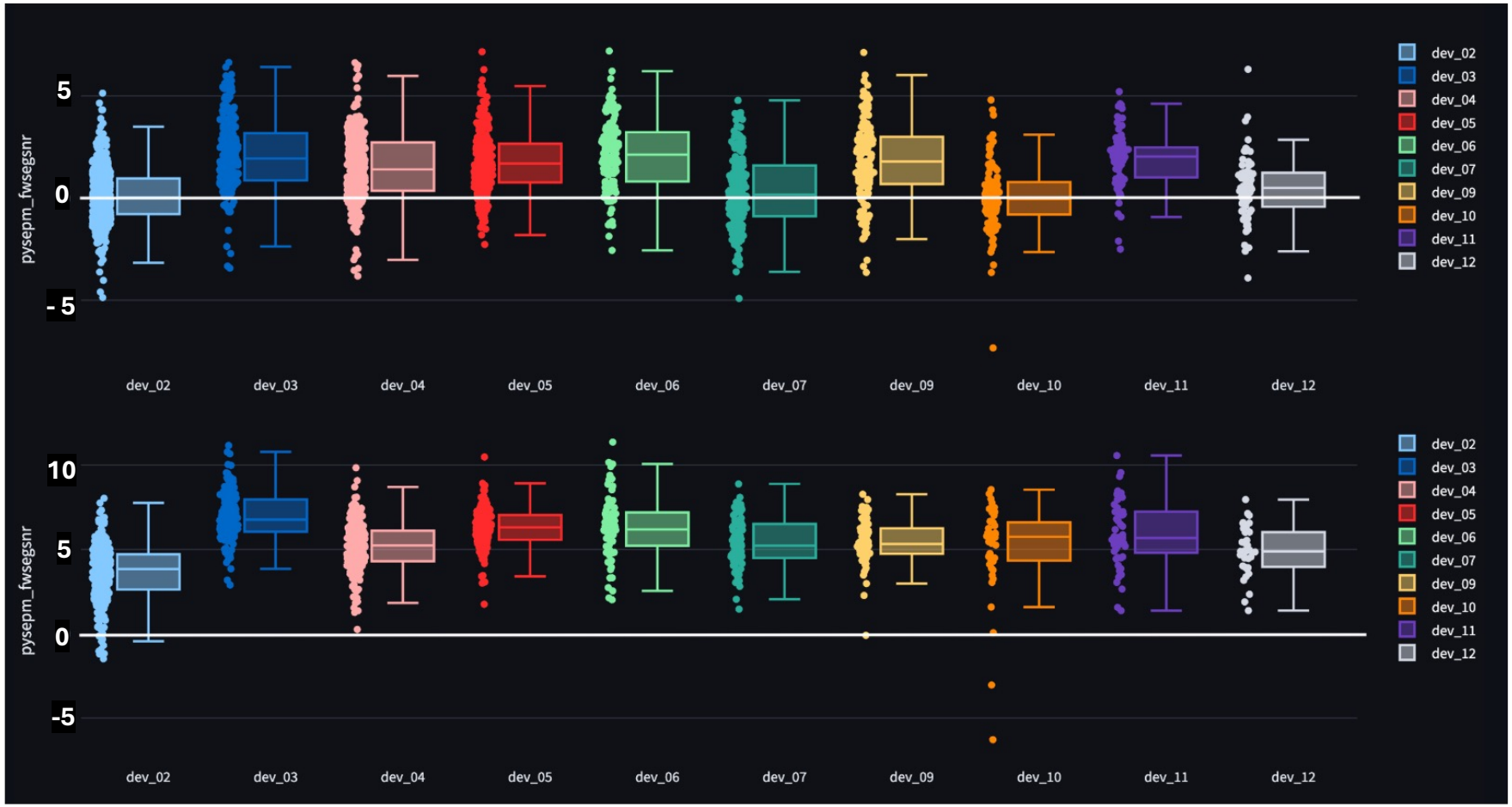
Baseline System

Summed	Device	Frequency-Weighted Segmental-SNR	STOI	PESQ	CSig
Passthrough	Aria	1.40	0.51	1.14	1.74
Baseline	Aria	5.55	0.57	1.19	2.15
Passthrough	HA	1.18	0.48	1.12	1.70
Baseline	HA	5.37	0.50	1.16	1.92

Baseline System

Summed	Device	Frequency-Weighted Segmental-SNR	STOI	PESQ	CSig
Passthrough	Aria	1.40	0.51	1.14	1.74
Baseline	Aria	5.55	0.57	1.19	2.15
Passthrough	HA	1.18	0.48	1.12	1.70
Baseline	HA	5.37	0.50	1.16	1.92

FWSegSNR per dev set session with passthrough (top) and baseline (bottom)



Key Dates

7th February 2026 – Challenge submission deadline

- Enhanced signals and 2-page technical report

4th May 2026 – CHiME workshop joint with HSCMA @ ICASSP 2026

University of Sheffield

CHiME CHALLENGE

SLT Centre for Doctoral Training

UK Research and Innovation

WSAudiology

Meta

Q&A

Visit the website

Get the data