

A digital twin for your hearing

Voice assistants from various brands are now found in many households and accompany us everywhere on our smartphones. Cluster of Excellence researcher Bernd T. Meyer and his team use the artificial intelligence (AI) on which these apps are based in their hearing research – and make good use of the voice assistants' deficiencies. Because very often these apps encounter the same problems as humans do.

By Sonja Niemann



The mincemeat for the bolognese is sizzling in the pan, the radio is blaring and the extractor hood is whirring loudly as the spaghetti is removed from its packaging and dropped into the pot with a splash. Now we just have to make sure we don't miss that perfect "al dente moment".

"Computer, set the timer for nine minutes."

"Sorry, I don't know that."

"Computer! Set the timer for nine minutes!"

"The timer has not been set."

There are certain situations in which home voice assistants seem incapable

of understanding commands. Despite, or perhaps because of this flaw, these small devices and the AI on which they rely have a special place in hearing research. After all, when it comes to understanding speech amid background noise, they have similar problems as people with impaired hearing. Making use of this similarity is just one of several approaches adopted by the researchers of the Oldenburg Cluster of Excellence Hearing4all in their quest to use speech recognition software to improve human hearing.

Professor Dr Bernd T. Meyer and his team from the Communication Acoustics Division are key players in this endeavour. Together with five PhD students he conducts research at the intersection of speech and hearing

research. Meyer has been fascinated by the possibilities of modern speech recognition ever since he did his degree in physics. At the time, Professor Dr Birger Kollmeier, who now leads the Hearing4all Cluster of Excellence, suggested the topic for a presentation. "I was interested in why the software's speech recognition was so much worse than that of people with healthy hearing," says Meyer. He began to develop methods to improve the systems – and also found ways to transfer his findings to hearing research.

One of his research approaches is based on using speech recognition software to diagnose hearing impairments. Thanks to research carried out in Oldenburg, German speaking persons can carry out a preliminary hear-

The human ear picks up acoustic signals, but it is the brain that decodes sounds, noises and speech. Voice assistants also break down sound waves into smaller units and then reconfigure them into words – and sometimes encounter the same difficulties as people with impaired hearing.

ing test in the comfort of their own living room using the Alexa Smart Home system. You simply give the specific command and the skill from Oldenburg will start playing short sentences, which you are asked to repeat.

500.000 utterances taught the AI to understand speech

The answers reach the voice assistant in the form of sound waves, which then divides them into short acoustic units. "The AI has been trained to match the acoustic signal units with the smallest units of human speech,

phonemes," Meyer explains. But this only solves half of the speech recognition puzzle. In the next step, the AI calculates the most probable sequence of the phonemes it has recognised and then strings them together, ideally forming the word that the test person said.

Whereas the home hearing test uses the Alexa AI, Meyer's team usually uses its own purpose-built AIs to conduct research. For this, the researchers use methods from machine learning, which means they teach a computer to recognise patterns in data and thus learn to transcribe spoken language into text. They train their AI by supplying it with speech samples – most recently, more than 500,000 utterances by more than 1,000 people. An artificial

neural network learns from this data and uses it to generate an output (in this case, a written word) from an input (in this case, a sound wave).

Unlike the Alexa hearing test, which provides feedback about possible hearing impairments using a rather limited traffic light system, the researchers want their AI to deliver sufficient accuracy to provide a clinical diagnosis. The team has already come very close to this goal, and the underlying algorithm is well advanced. "We have shown that our neural network can test how well a person hears with a similar degree of accuracy to a test performed by a medical professional," says Meyer.

Despite constant improvements in diagnosis and hearing aids, science is

still not able to restore all facets of hearing for hearing-impaired people. This is because it is not enough to simply amplify the soundscape. In everyday life, the main objective of hearing is to understand speech. And in this respect humans have very similar problems to the voice assistant in the kitchen: the spoken word easily gets lost amid all the background noise.

Meyer's research team is seeking to turn this shared weakness into an advantage using another approach. Instead of programming a neural network to make it particularly proficient in understanding a speaker, as with the hearing test, the physicist and his PhD student Jana Roßbach aim to create a speech recognition system that hears as well or as poorly as a real person. The Oldenburg researchers recently attracted a lot of attention with a study in which they were able to calibrate their neural network to mimic the hearing impairments of test persons with such accuracy that the performance of the humans and the computer in hearing tests was almost identical – like acoustic doppelgangers. "The idea behind this is that if an AI is able to predict that a hearing aid user will not be able to understand a certain word spoken in their presence, it will also be able to optimise the hearing aid settings to ensure that the person can follow the conversation. In this way we turn a disadvantage into an advantage," explains Meyer. "Up to now we have focused on hear-

ing with one ear, but we know that binaural hearing has positive effects on speech comprehension," he adds. Therefore, they now plan to train the AI to make the same predictions in the test as a human using both ears.

Usable not just in hearing aids, but in all hearables

In the experiments conducted so far, the AI was fed speech samples which were overlaid with a strong signal noise that simulated the hearing impairment of the test person. The AI transcribed this speech sample and compared the result with the information about what the speaker had actually said. If the two were identical, the AI knew that the test person would understand the speech sample. "In real life, of course, an AI doesn't know what is actually being said," says Meyer. So he is working with Roßbach to make predictions possible without this information – based on the sound quality of the speech sample.

The benefits of AI applications like this go beyond hearing aids. "They are relevant for the entire spectrum of communication with hearables," says Meyer. In addition to hearing aids this includes smart headphones, which have long been used not just for playing

music but also to block out unwanted background (noise-cancelling) or amplify speech in a particular environment. Tailoring these devices to the hearing ability – or preferences – of the wearer could be the next logical step in their development. However, the power supply required for complicated algorithms as well as the limited performance of the built-in processors still pose hurdles at this stage.

But an example from Meyer's department shows that these hurdles are already being overcome in Oldenburg's hearing research. In 2017 Meyer attended a scientific lecture in Baltimore and was very impressed by a demonstration in which an AI was able to separate the acoustic signals of two people speaking into a microphone at the same time. "That really blew me away," he says. But although this seemed to offer a promising approach to a fundamental problem in hearing research – separating desired sounds from unwanted noise – it was initially useless for that purpose. "The app was too resource-heavy and too slow. Hearing aids can only work with a maximum delay of ten milliseconds, otherwise the natural undelayed sounds and the delayed sound from the hearing aid result in distortions," Meyer explains.

Together with PhD student Nils Westhausen he has now found a way to reduce the delay to two milliseconds making it sufficiently fast and resource-efficient for use in hearing aids. This technology could enable people to actively select the acoustic signal they want to listen to – and, for example, block out whirring extractor hoods and sizzling pans. That would be another major breakthrough.

Communication acoustics expert Bernd T. Meyer uses machine learning for speech and hearing research. One of his projects uses voice assistants to perform high-precision hearing tests.



Outlooks

How can we reduce social inequalities?



Prof. Dr. Gundula Zoch

Sociology of Social Inequalities

"We are facing enormous social challenges such as the transformation of the labour market and new technologies. To meet these challenges, comprehensive access to education for children and adults is essential. They all need the skills required to seize opportunities and cope with social change.

Education is also key for social cohesion. It plays a role not only in achieving material prosperity and wellbeing, but also in social, political and societal participation. However, in Germany we urgently need reforms to ensure more equal opportunities in the education system.

Research has shown that good education in early childhood is crucial to reducing educational inequalities. It benefits children, including those from disadvantaged groups, throughout their lives. It also helps older generations, because the costs of government benefits will be borne by young people who embrace technological developments in the labour market and use or complement them with their skills.

My vision for the future is that fifty years from now we will have created an education system in Germany which benefits everybody, regardless of their family circumstances and background. And that education is seen as a lifelong process which paves the way for people to participate successfully in society even in times of major change. Investing money and expertise here is worthwhile, because it will enable us to reduce social inequalities across the board."