

Machine-Learning Model Could Improve Human Speech Recognition

A tool that predicts how many words per sentence a listener understands could one day allow companies to make bespoke hearing aids with improved capabilities.

By **Rachel Berkowitz**

In 2019, assisted-listening devices brought the gift of hearing to 7.1% of the US population aged 45 and over. But these hearing aids are far from perfect. One way researchers think that they can improve these devices is to integrate them with speech-processing models that predict how individuals with different degrees of hearing loss distinguish words in noisy environments. In a step that could allow for more customized hearing restoration, Jana Roßbach, Bernd Meyer, and their colleagues at Carl von Ossietzky University of Oldenburg in Germany have now developed a machine-learning model that they show can correctly predict speech intelligibility for a variety of auditory conditions [1]. They say that a future version

of their model could be integrated into hearing aids to improve speech intelligibility for the hearing-impaired.

Modern hearing aids convert incoming sound waves into numerical codes and then send amplified versions of those waves into the ear through a speaker. The codes include information about the waves' frequencies and their amplitudes. But audition is more complex than simply detecting sound waves.

The ability to distinguish phonemes—the units of sound that make up words—is a key component of hearing. This ability is often reduced for those with hearing impairments. Hearing aids help in mitigating this loss by using signal-processing algorithms to improve speech recognition. But developing and evaluating these algorithms typically requires time-consuming listening experiments that test the algorithms' capabilities under myriad acoustic conditions.

To solve this problem, Roßbach, Meyer, and colleagues developed a machine-learning model that determines the acoustic conditions experienced by a listener and then estimates just how well that listener can identify words in that environment. To make this estimate, the model uses an automated speech-recognition system based on machine learning.

The researchers trained and tested their model using recordings of sentences that were degraded to mimic how individuals with different types of hearing impairments perceive speech in



A model that predicts how well a hearing-impaired individual understands speech in different acoustic environments could be used to develop the next generation of speech-enhancement algorithms for hearing aids.

Credit: Yuli/stock.adobe.com

different noisy environments. The team then played to normal-hearing and hearing-impaired listeners these same recordings. They asked the participants to write down the words that they heard for each track. From those answers, the team determined the threshold level of noise (in decibels) that resulted in a 50% word-error rate for each listener, for each environment, finding a good correspondence to the model predictions.

Roßbach, Meyer, and the rest of their team hope that a future version of their model might end up in hearing aids. But before that can happen, they need to fix several issues with the current version. One of those issues is that the model “needs information about what’s actually spoken,” Meyer says. But that information does not exist in real-world situations. The team is working on fixing that and other problems with the goal of creating a machine-learning model that can maximize speech

intelligibility for any hearing-impaired person, Meyer says.

Torsten Dau, a researcher in hearing technology at the Technical University of Denmark, says that Roßbach’s model is an important step toward a “nonintrusive” method for improving speech recognition of the hearing impaired. He notes that the model “performs very well” in the acoustic conditions that the team used. “It will be exciting to see how this approach generalizes to [other] acoustic conditions,” he says.

Rachel Berkowitz is a Corresponding Editor for *Physics* based in Vancouver, Canada.

REFERENCES

1. J. Roßbach *et al.*, “A model of speech recognition for hearing-impaired listeners based on deep learning,” *J. Acoust. Soc. Am.* **151**, 1417 (2022).