Leading the industry in speech intelligibility

Dr.-Ing. Daniel Marquardt and Larissa Taylor, Ph.D.

Key takeaways:

- Omega AI consistently outperforms other brands, delivering up to 6.5 dB advantage (70% better performance) in complex, noisy listening situations
- An innovative, realistic acoustic setup with advanced robust methodologies ensures fair, accurate, and repeatable results that align with real-world listening situations.

Introduction

Conversations in large, noisy groups remain challenging for patients with hearing loss for a variety of reasons. Typically, an impaired auditory system and reduced ability to discern speech from noise content culminates in experiencing increased cognitive load and needing to work harder to "fill in the gaps" during conversation, ultimately leading to fatigue and frustration.

Hearing aids can assist in challenging listening situations by processing the sounds heard in a way that helps speech to be more easily perceived by the listener. Speech intelligibility evaluation is an important measure of how well speech is understood in such a situation and can be assessed objectively or subjectively with human listeners. Although traditional speech intelligibility tests such as the Hearing in Noise Test (HINT) *(Nilsson et al., 1994)* or the American English Matrix Test (AEMT) *(HörTech, 2014)* are highly valuable for gathering direct feedback from hearing-impaired individuals, they come with several limitations. On the one hand, these tests are desirable because they provide insights into how hearing aid users perceive speech intelligibility, allowing a better understanding on how to address specific needs.

On the other hand, these tests typically rely on stationary noise signals like speech-shaped or pink noise, which—while effective at reducing intersubject and intrasubject variability—fail to represent the complexity and dynamics of real-world acoustic environments. Additionally, subjective tests are time-consuming and can be cognitively demanding, potentially affecting test-retest reliability.

To understand the effectiveness of Starkey's sound processing technology in real-world challenging situations and to provide a fair comparison of speech intelligibility between different brands, Starkey pioneered an efficient, reliable, and scalable approach to acoustic analysis, setting a new standard in the field of comparison studies. The cutting-edge artificial intelligence (AI)-based Speech-to-Text (STT) technology initially reported by Betlehem et. al., 2022, was further developed to analyze a multitude of acoustic scenarios to measure the speech intelligibility psychometric function, which is compared across different brands.

The AI-based speech intelligibility evaluation method offers significant advantages. It ensures fairness by processing identical input signals across different comparisons, eliminating variability inherent in subjective testing, such as differences in hearing loss, individual perceived benefits, and listening environments.

The AI-based evaluation system can also be applied to a wide range of acoustic scenarios, including more realistic and dynamic noise environments, enabling better predictions of real-world performance. By leveraging these capabilities, AI-driven approaches provide consistent, scalable, and objective assessments that complement traditional testing methods while addressing their shortcomings.

Methods

Starkey's Omega AI speech intelligibility performance was compared with five of the latest flagship hearing aid brands available at the time of this study. All hearing aid brands were set to their automatic default settings. For 2 out of the 5 brands, the dedicated speech-in-noise settings that used a dedicated DNN chip were evaluated, resulting in 7 total comparisons. Each hearing aid was fitted to an N3 hearing loss *(Bisgaard et al, 2010)* according to the manufacturer's fitting formula and the occluded coupling was used (Power Dome). The Experience Level was set at 100%.

A series of acoustic scenarios *(see Table 1 and Figure 1)* were meticulously crafted within a controlled laboratory environment. These scenarios featured speech and noise signals presented from different locations, with Signal-to-Noise Ratios (SNRs) ranging from -16 dB to 14 dB in 2 dB increments. Each SNR level was measured with 20 sentences from the AEMT with an overall length of 128 seconds and incorporated 6 distinct real-world noise files to ensure a comprehensive performance analysis.

Word recognition performance was evaluated with advanced Speech-to-Text (STT) technology, specifically leveraging the Whisper Automatic Speech Recognition (ASR) system *(Radford et al, 2022)*.The "Tiny" model was used to better account for speech intelligibility deficits of hearing impaired vs. normal hearing.

Similarly, as in a subjective speech intelligibility test, every sentence is processed by the ASR system to accurately measure word recognition rates by comparing the detected words to the ground truth data.

The higher word recognition rate for every sentence between binaural responses was used to account for the better ear effect. For each acoustic scenario and SNR, the average word recognition results were then calculated across the last 10 sentences (leaving the first 10 sentences, i.e., 64 seconds, for algorithm initialization) and all 6 noise files to derive a robust set of data points that were then fitted to a sigmoid function.

Table 1: Acoustic scenarios for comparison across brands. The speech signal was played from a single loudspeaker at the respective speech position. The diffuse real-world signals were generated by playing uncorrelated noise snippets through 8 loudspeakers.

	Acoustic Scenarios				
	1	2	3	4	5
Speech Position	0°	45°	90°	135°	180°
Noise Position	Diffuse				
Speech Signal	20 AEMT sentences				
Noise Signals	Bar Noise, Shopping Mall, Construction, Crowd, Indoor Party, Outdoor				

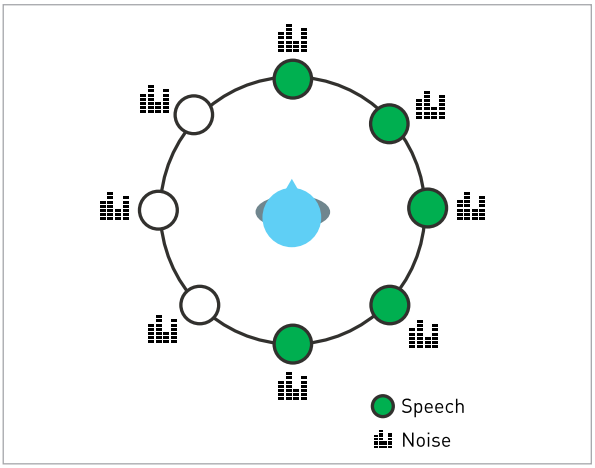

Figure 1: Acoustic scenarios used for evaluation. Location of speech signals used shown as green circles, while diffuse real-world signals (shown as noise signal icon) played through 8 loudspeakers.

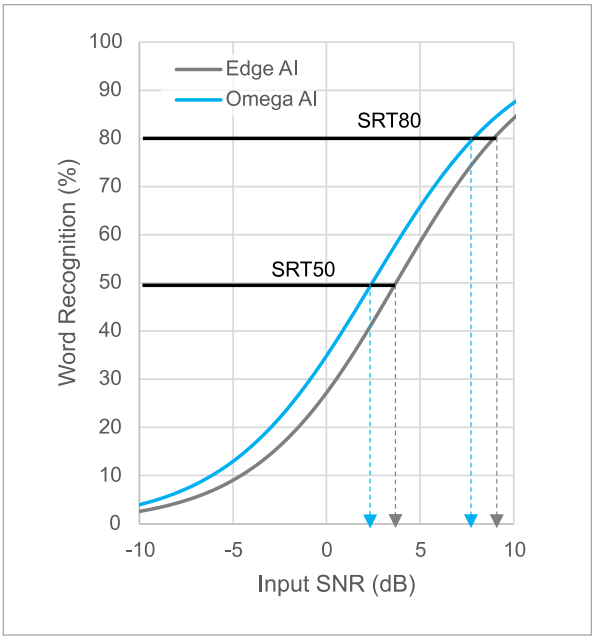

Figure 2: An example of the method used to evaluate speech intelligibility differences between brands. A sigmoid function is fitted to the data acquired (shown in gray for Edge AI, and blue for Omega AI in this example). The SRT50 and SRT80 are derived by extracting the SNR value on the x-axis that corresponds with the 50%- and 80%-word recognition rate respectively. Importantly, this AI-based psychometric function, showing a 1.3dB SRT50 improvement of Omega AI over Edge AI, was calculated for the same acoustic scenario that we used for the subjective intelligibility test with hearing impaired in (Marquardt et al, 2025), where we found a SRT50 improvement of Omega AI over Edge AI of 1.4dB, demonstrating the remarkable accuracy of the AI-based speech intelligibility assessment.

From this fitted function, key performance metrics were calculated, namely the 50% and 80% Speech Reception Thresholds (SRT), SRT50 and SRT80 respectively, as typically measured for subjective listening test *(see Figure 2)*.

Results

For this comparison, the performance across all acoustic scenarios was averaged as a representation of overall performance and shown in Figure 3. Omega AI demonstrated superior performance across all comparisons.

Using the SRT50 values and the slope of the psychometric function of the AEMT, which was found to be 11%/dB for hearing impaired [Harianawala et al. 2019], we can show that

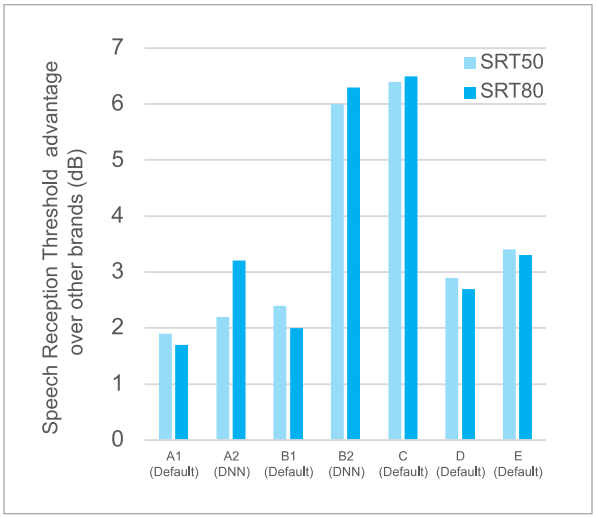

Figure 3: SRT50 and SRT80 advantage of Omega AI over other brands across all the conditions evaluated.

Omega AI provides a minimum of 20% speech intelligibility advantage over Brand A and a maximum advantage of 70% over Brand C, demonstrating the effective real-world performance of Starkey's innovative DNN-powered sound processing technology.

Conclusion

Starkey innovated an AI-based evaluation system to help provide a fair, objective way to assess speech intelligibility in realistic, challenging situations. When evaluated across different flagship brands, Omega AI demonstrated significant advantages for speech intelligibility measures in a variety of acoustic conditions, affirming its strong lead in providing effective sound processing strategies when it matters.

Acknowledgements

The authors would like to express their gratitude to Jinjun Xiao, Al Ganeshkumar, Jinjing Xu, Sian Halvorsen and Martin McKinney for their invaluable contribution to this research and thoughtful feedback throughout the study.

References

1. Betlehem. T., Marquardt, D. and McKinney, M., "Estimating Speech Intelligibility using Google Speech-to-Text", presented at the International Hearing Aid Research Conference (IHCON 2022), Tahoe City, CA, August 10, 2022.
2. Bisgaard, N., Vlaming, M. S., & Dahlquist, M. (2010). Standard audiograms for the IEC 60118-15 measurement procedure. *Trends in Amplification*, 14(2), 113–120. https://doi.org/10.1177/1084713810379609
3. Harianawala J, Galster J, Hornsby B. (2019, April). Psychometric Comparison of the Hearing in Noise Test and the American English Matrix Test. *J Am Acad Audiol.* 30(4), 315-326.
4. HörTech gGmbH. International Matrix Test, reliable speech audiometry in noise. (2014, n.d.).
5. Marquardt, D., Xiao, J., Ganeshkumar, A., Xu, J.J., Halverson, S., Taylor, L., and McKinney, M. (2025) Enhancing Directionality with Deep Neural Networks. Starkey white paper.
6. Nilsson M, Soli SD, Sullivan JA. (1994, Feb.). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am,* 95(2):1085–1099.
7. Radford, Alec; Kim, Jong Wook; Xu, Tao; Brockman, Greg; McLeavey, Christine; Sutskever, Ilya (2022). Robust Speech Recognition via Large-Scale Weak Supervision. (https://arxiv.org/abs/2212.04356)

Author Biographies

Dr.-Ing. Daniel Marquardt completed his studies in Media Technology at the Ilmenau University of Technology, Germany, in 2010, graduating with a Diplom-Ingenieur degree. In 2015, he earned his doctorate in the field of Speech Signal Processing at the University of Oldenburg. He currently works as a Principal Signal Processing Engineer at Starkey.

Larissa Taylor, Ph.D., is an Audio Systems Engineer at Starkey. Her primary focus is ensuring the overall sound quality of hearing aids. Larissa started at Starkey in 2022 upon completing her doctorate on listening effort and sound quality with hearing aids at McMaster University, Hamilton, Ontario. Canada.