WIDEX ALLURE™ AI RIC WITH CLARITY BOOST

INTRODUCING 3RD GENERATION AUDIO-SPECIFIC AI: EXPRESSIVE. EFFICIENT. ENGINEERED FOR OPTIMAL BALANCE.

Authors: Federica Bianchi, PhD, Mads Oestergaard, B. Eng. Lead ML Engineer, Jens Brehm Nielsen, PhD

At the heart of the Widex Sound Philosophy is a true dedication to providing users with the most natural hearing experience. The guiding star is an optimal balance between speech clarity and natural awareness of the surroundings, helping users to be fully immersed and present in every moment of life.
Guided by this goal, Widex introduces Allure AI RIC with Clarity Boost, designed to deliver the most natural hearing experience for everyday use, while empowering users to activate a 3rd generation AI co-processor whenever more support in noise is needed. By combining a highly efficient audio-specific AI architecture with perceptually motivated training, Clarity Boost achieves superior denoising while preserving those sound features that are important for human perception.
The superior performance of Clarity Boost was proven in a technical study evaluating output SNR in realistic sound environments, as well as processing delay. The results of the study clearly showed that Clarity Boost outperformed the other manufacturers, by achieving the highest average output SNR of all four competitors with AI-based denoising and the lowest average processing delay among competitors with a dedicated AI co-processor.

Key findings

- Allure AI RIC with Clarity Boost introduces a 3rd generation audio-specific AI, achieving highly efficient denoising through a dedicated AI co-processor.
- The DNN in Clarity Boost is trained using a Large Audio Foundation Model, effectively balancing superior noise attenuation with natural sound quality.
- In a technical study, Clarity Boost delivered:
  - The highest average output SNR of all four premium competitors with AI-based denoising, showing up to 6.2 dB improvement relative to competitors.
  - A 5.1 dB improvement in output SNR relative to the Universal program.
  - The lowest average processing delay among hearing aids with a dedicated AI co-processor.

Introduction

The overarching goal of denoising algorithms in Widex hearing aids is to support users in achieving an optimal balance between speech clarity and natural awareness of their surroundings in real life, while preserving a natural hearing experience. This raises the question: What are the typical situations encountered by hearing-aid users in real life? Wagener et al. (2008) recorded and classified listening situations encountered by experienced hearing-aid users in daily life. Their results showed that these situations vary widely across users, not only in their acoustic characteristics, but also in their perceived importance and in the benefit provided by hearing aids. Further analysis of these recordings by Smeds et al. (2015) estimated that everyday conversations in babble noise typically occur at an average signal-to-noise ratio (SNR) of around 5 dB, while conversations in other complex sound environment may range from SNRs of 0.8 dB (department store) to 7.4 dB (kitchen). These findings highlight the importance of designing hearing-aid features that can provide clear benefits in a variety of listening situations and acoustic conditions commonly experienced in users' everyday life.

Guided by the goal of supporting users in real life, Widex introduces the Allure AI RIC with Clarity Boost, designed to deliver the most natural hearing experience for everyday use, while empowering users to activate a 3rd generation Artificial Intelligence (AI) co-processor whenever needed. Allure AI RIC with Clarity Boost gives users additional support in noise, while preserving a balance between superior noise attenuation and sound quality in everyday life.

This white paper is divided into two parts. In the first part, the innovative technology behind Allure AI RIC with Clarity Boost is presented, with a deep dive into the unique AI architecture of Clarity Boost and the training of the neural network. The second part introduces a technical study comparing the performance of Allure AI RIC with all other premium competitors with AI-based denoising, both in terms of output SNR and processing delay. The study provides evidence that Allure AI RIC achieves competitive denoising with the lowest average group delay of all competitors with an AI co-processor.

Clinical implications

Widex Allure AI RIC with Clarity Boost gives users additional support in noise, while preserving a balance between superior noise attenuation and sound quality.

Machine learning and deep neural networks in hearing aids

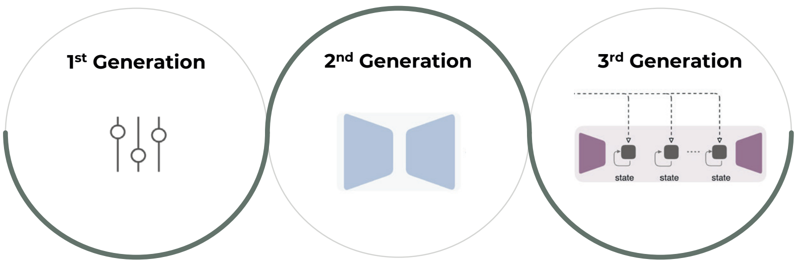
Real-world listening environments are highly complex and continuously changing. Speech, background noise, and environmental sounds vary over time, overlap in frequency, and interact with room acoustics in ways that are difficult to capture with traditional signal processing alone. Machine learning (ML) offers a data-driven approach where algorithms learn directly from large collections of audio data how speech and noise typically pattern in realistic situations.

In the last decade, ML has been introduced in hearing aids with different purposes (e.g., sound scene classification, user-driven personalization of sound, beamformer steering, denoising).
A common class of modern ML models used for audio are deep neural networks (DNNs). DNNs consist of multiple processing layers that gradually transform an input audio signal into an output that is more useful for a given task (e.g., a denoised sentence if the task is noise attenuation). Early layers tend to capture low-level acoustic properties, such as energy balance, spectral and temporal characteristics, while deeper layers represent increasingly complex structures related to sound patterns over time. DNNs models in hearing aids are trained during development to recognize statistical patterns in sound and to apply this knowledge in real time on the device. Importantly, the learning process takes place offline. Once deployed, the hearing aid uses a fixed, optimized model to support the listener without adapting or learning from the individual user's data.

Depending on the manufacturer's goals and philosophy, DNNs have been introduced in hearing aids to address different user needs. Below **three different approaches** are introduced of how DNNs can achieve **denoising in hearing aids** and how **Clarity Boost introduces a totally new way of approaching denoising** that not only operates **directly on the time domain,** in line with Widex' unique time-domain filter bank design, but is also **based on AI architectures that are specific to audio signals** (see Figure 1 for an overview of the three approaches).

The first denoising approach with modern ML in hearing aids is to use DNNs to control parameters in the existing signal processing path. This **first generation** of DNNs for denoising works by defining gains that are applied to the signal using conventional digital signal processing (DSP). While this approach can achieve noise reduction, the DNN model itself does not directly process the audio signal end-to-end (Figure 1). In this approach, the DNN operates as a **controller of existing DSP algorithms.**

A **second generation** of DNNs for denoising includes DNNs in the signal path, directly processing audio end-to-end (**Audio AI**, Figure 1). These systems are built on AI architectures originally developed for image processing, called convolutional neural networks (CNNs). CNNs are neural networks that learn to find patterns in data by looking at fixed windows of the input signal (Figure 2, top panel). A CNN scans across the input signal with a fixed filter (or fixed window). These filters learn to detect simple patterns, like edges in images or short sound patterns in audio. As the signal goes through the network, the model combines these small patterns to recognize bigger and more complex patterns. For example, when processing images, a CNN might first detect edges, then shapes, and finally whole objects like faces or cars. Imagine a CNN that looks at 20 milliseconds of sound at a time to decide what is happening. When it processes the signal, it only "sees" those 20 milliseconds at that moment. Sounds that happened earlier are "not seen" and aren't directly considered by the model. Hence, because **CNNs were originally developed for image processing, they have a limited ability to capture the temporal structure inherent in audio signals.** This means that achieving strong performance requires large models with high computational cost.



	1st Generation	2nd Generation	3rd Generation Widex Allure AI RIC
DNN in the signal path, processing audio end-to-end (**Audio AI**)		✓	✓
AI architecture specific for audio processing (**Audio-specific AI**)			✓
Memory and compute efficient			✓

Figure 1: *Three different approaches (1st, 2nd, and 3rd generation) of how DNNs can achieve denoising in hearing aids. Widex Allure AI RIC with Clarity Boost is the first hearing aid to introduce a 3rd generation audio-specific AI architecture directly in the signal path to achieve highly efficient denoising.*

Widex Allure AI RIC with Clarity Boost launches a **third generation** of DNNs in hearing aids for denoising. While also processing **audio end-to-end,** the DNN in Clarity Boost is built on **AI architectures specific to audio processing** (**audio-specific AI**, Figure 1), called linear recurrent neural networks (L-RNNs). L-RNNs are more naturally suited to capture the temporal structure of audio signals, as they are designed to process sequences step by step while keeping memory of the past ("state" depicted in Figure 2, bottom panel). So, if a L-RNN model processes 20 ms of audio at a time, it will compress and selectively carry information from earlier chunks (e.g., 20ms, 40ms, 100ms ago; "compressed history") while analyzing the current chunk of audio. The current audio and the earlier state are used to update the new state (or memory) of the system. This helps capture long-term patterns, like words in speech or rhythm in music. This audio-specific architecture achieves comparable or superior denoising performance with much fewer parameters than second-generation systems, enabling highly efficient real-time operation on a dedicated AI co-processor. As a result, Allure AI RIC with Clarity Boost is more **memory- and compute efficient** (Figure 1), which helps **reduce power-consumption, hearing-aid size and processing delay.**

To summarize, DNNs have been introduced in hearing aids following different approaches for denoising (see Figure 1 for an overview). Widex Allure AI RIC with Clarity Boost is the first to use a **third generation audio-specific AI architecture** to efficiently achieve additional support in noise. Besides the highly efficient audio-specific AI architecture, the other element of novelty in Clarity Boost lies in the training of the DNN, which allows Clarity Boost to balance noise attenuation with naturalness of sound, speech integrity, and preservation of room acoustics. Both elements of novelty, namely the AI architecture and the DNN training, will be explained in more detail in the next sections.

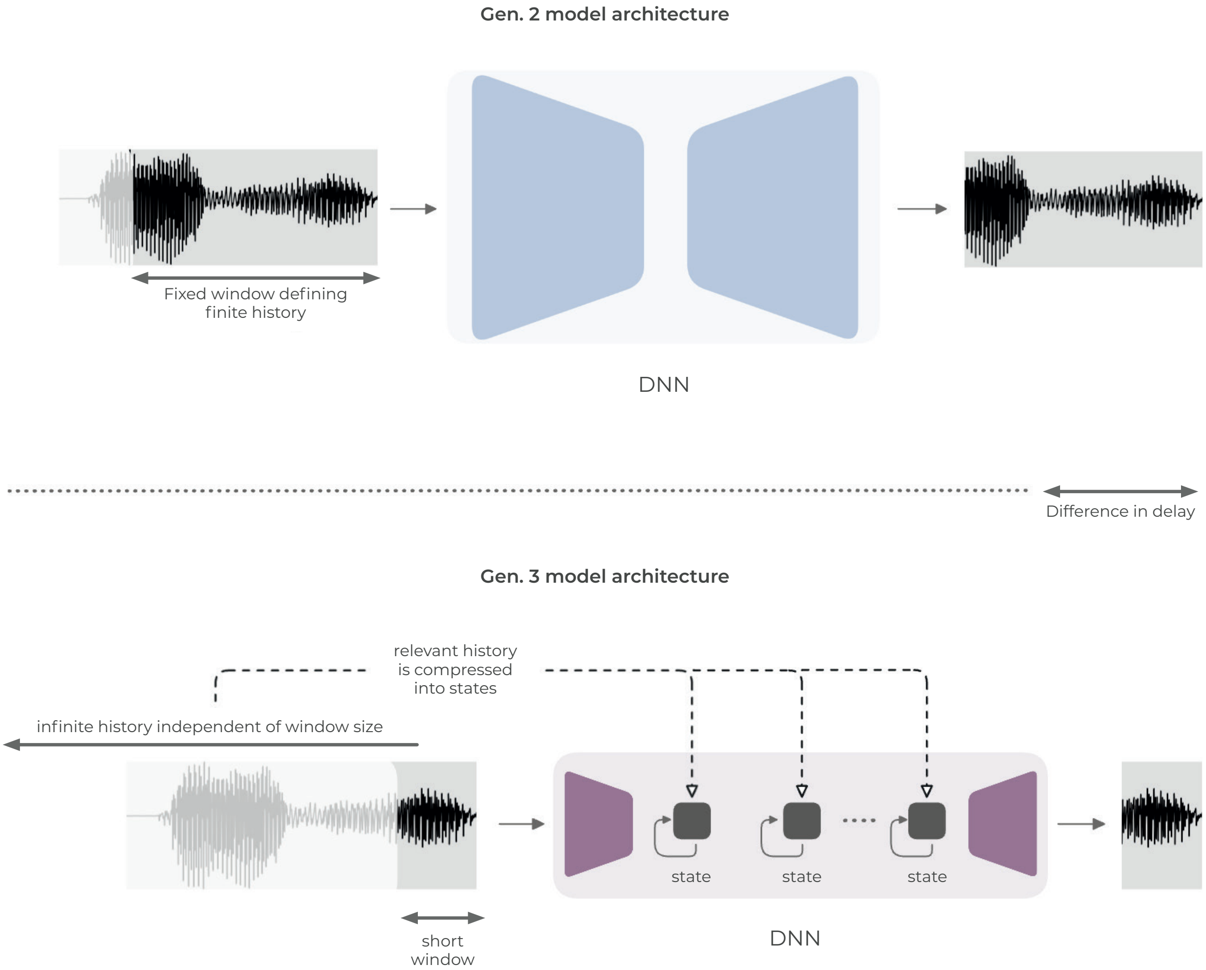


Figure 2: *Comparison of second-generation (top) and third-generation (bottom) AI architectures for denoising audio signals. In the second-generation architecture, the DNN processes a fixed chunk of sound at a time ("fixed window") without having a memory of previous sound windows. In the third-generation architecture (DNN in Clarity Boost), information from previous windows of sound is continuously accumulated and compressed into the system's memory ("state"), enabling efficient capture of the long-term temporal structure of audio signals. As a result, the third-generation approach is far more memory and compute efficient than the second-generation approach, reducing power consumption, hearing aid size, and processing delay.*

Did you know?

Different AI architectures can be used for denoising audio signals in hearing aids. The choice often depends on strict constraints such as processing delay, power consumption, and size. Common approaches include:

- **DNNs controlling DSP algorithms (first generation)**
  This approach uses deep neural networks (DNNs) to estimate gains for each time-frequency bin of a spectrogram. These gains are then applied using conventional digital signal processing (DSP), meaning the DNN controls the processing but does not directly process the audio signal end-to-end.
- **Convolutional neural networks (CNNs)**
  CNNs are designed to process spatial data such as images, videos, and spectrograms. They use convolutional filters that detect local patterns (edges, textures, shapes). Widely used in computer vision and adapted for audio processing. When placed directly in the signal path, they can process audio end-to-end but require longer processing time (second generation).
- **Linear recurrent neural network (L-RNNs)**
  L-RNNs are audio-specific AI architectures, built to efficiently model temporal dependencies in audio. They process sound sequentially while keeping memory of past audio. **Allure AI RIC with Clarity Boost is the first to introduce L-RNNs directly in the signal path (third generation), enabling efficient and effective denoising.**

The technology inside Allure AI RIC

Allure AI RIC features a dual-chip architecture, combining the Allure W1 chip with a dedicated AI co-processor used within the Clarity Boost program ("AI mode", Figure 3). The DNN in Clarity Boost consists of a third-generation audio-specific AI architecture that achieves superior denoising while carefully balancing power consumption and processing delay.

The two Widex classic listening programs, Universal (default) and PureSound, are powered by the Allure W1 chip and rely on traditional noise reduction, most importantly the Speech Enhancer Pro. Based on the Speech Intelligibility Index (SII), the Speech Enhancer Pro reduces noise while preserving speech (Herrlin et al., 2025), effectively balancing focus on speech and awareness of surroundings for everyday listening (Balling et al., 2025c).

Below is a short overview of key points and user benefits for each of the listening programs featured in Figure 3.

- ***Universal:*** Featuring a high-definition adaptive beamformer (HD locator) and Speech Enhancer Pro for denoising, the default program on Allure AI RIC achieves superior sound quality (Balling et al., 2025a) and an improved experience in noise (Balling et al., 2025b). The program is designed to provide a comfortable balance of focus and environmental awareness, for a natural hearing experience in everyday life.

- ***PureSound.*** Featuring directionality and noise reduction systems that are uniquely designed to keep processing delay low, PureSound on the Widex Allure platform retains all the established benefits of ZeroDelayTM (e.g. Balling et al., 2020, 2021; Korhonen et al., 2022; Lelic et al., 2022; Slugocki et al., 2020; Zhou at al., 2024) and delivers significantly improved speech-in-noise performance relative to Moment (Weber and Branda, 2025). With an ultra-low delay around 0.5 ms, the PureSound program is designed to avoid the comb-filter effect and provide natural sound quality (Balling et al., 2020).

- ***Clarity Boost.*** Powered by a dedicated AI co-processor, the Clarity Boost listening program features a broadband beamformer and a DNN for denoising ("AI mode"). The DNN is placed directly on the signal path, receiving the time-domain audio signal as input and returning a time-domain partially denoised audio signal as output. The goal is not to remove all background noise, but to improve speech clarity while preserving environmental awareness and natural sound quality, in line with the Widex Sound Philosophy of helping users remain immersed and present in their surroundings.

The next section provides a deep dive into the development of Clarity Boost, focusing on what makes its AI architecture unique, and how the DNN model was trained and evaluated.

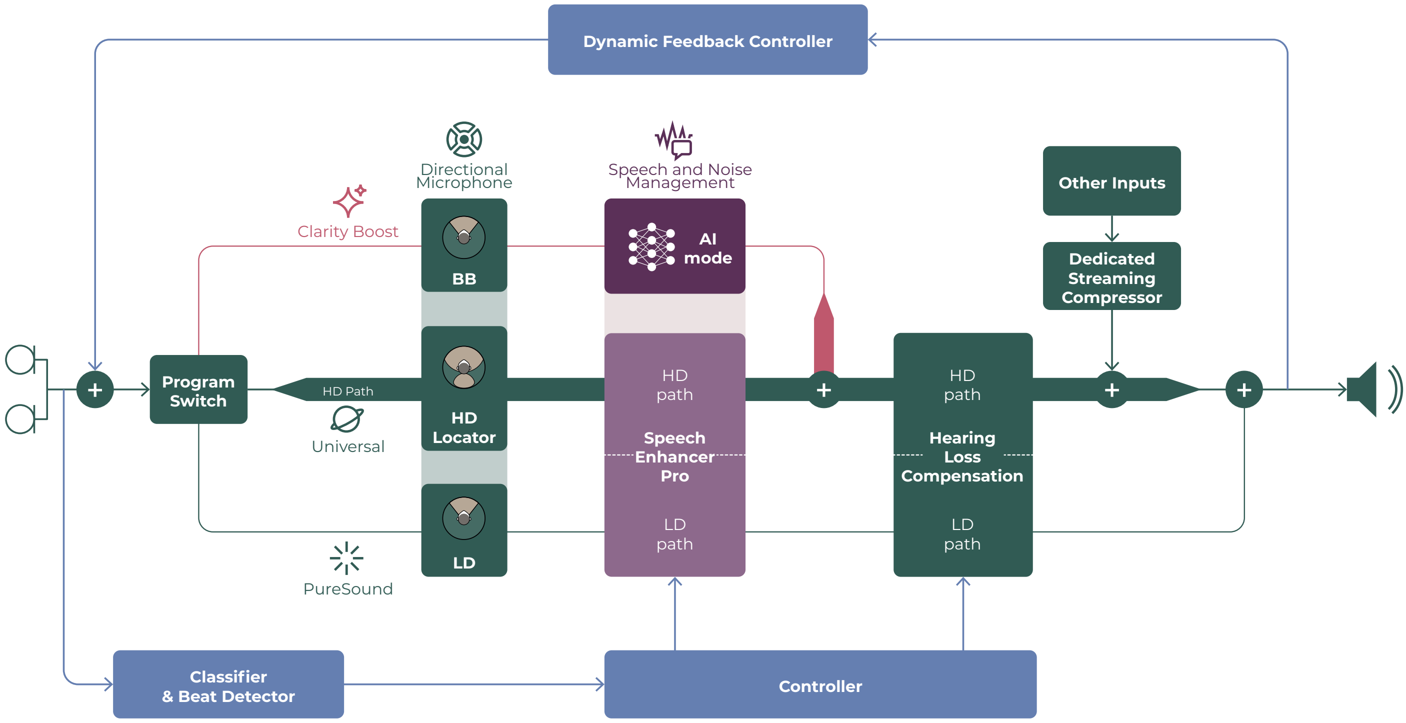


Figure 3: Illustration of the signal processing paths in Widex Allure AI RIC, for three listening programs: Universal (default), Clarity Boost, and PureSound. The "AI mode" depicts the DNN in the Clarity Boost program. Thick lines indicate a multi-channel signal path, thin lines indicate a single-channel signal path (time-domain audio). Abbreviations: BB – broadband beamformer; HD - High Definition; LD - Low delay.

Audio-specific AI in Clarity Boost: development and training

Clarity Boost is designed as a full-audio DNN-based denoising system that operates directly on the time-domain audio signal. While this enables powerful noise attenuation, it also entails strict requirements on sound quality, latency, and battery consumption. The design of Clarity Boost is therefore guided by a small number of core principles that balance denoising performance with the Widex Sound Philosophy.

The following sections focus on the **two main elements of novelty in Clarity Boost:**

1- Audio-specific AI architecture

2- Training guided by human perception to balance noise attenuation with natural sound quality

1) Audio-specific AI architecture: Efficient and expressive

When designing a neural network, the first step is to choose an architecture. This defines the structure of the model: how many layers it has, how information flows through it, and which operations it performs. Hearing aids impose strict constraints on computational complexity, power consumption, and processing delay. To meet these constraints, Clarity Boost is built using a highly efficient DNN architecture that uses about 10 times fewer parameters than many state-of-the-art denoising models, while achieving comparable or superior performance.

Clarity Boost achieves this efficiency by starting with an AI architecture that is specific to audio processing (**audio-specific AI**), using **linear recurrent neural networks** (L-RNNs, Orvieto et al., 2023). These L-RNNs constitute the building blocks of Clarity Boost. L-RNNs are neural networks "with a memory", as they store information from previous time windows of sound into the system's memory (called "state", see Figure 2). The state gets updated every time new audio input arrives using linear updates, which results in **very efficient** audio processing, while avoiding heavy computations that traditional, non-linear RNNs usually require.

> *Think of a person listening to a story. The person's working memory acts like a mental buffer that holds the most recent and relevant information. As a new part of the story arrives, the brain updates the working memory by combining the new story information with what was already in the buffer. In this analogy, the working memory represents the state of the L-RNN that is continuously updated as new chunks of audio arrive. The listening process is the L-RNN updating its state over time to keep track of the whole audio sequence.*

Even though each update is simple, repeating it over time lets the model build up a structure that represents long term information about the signal. Stacking multiple building blocks in depth increases the expressiveness of the network, allowing it to capture complex, time-varying speech and noise patterns without using too many parameters. This combination of efficiency and expressiveness enables real-time operation with low latency on a dedicated AI co-processor.

2) Training of the DNN in Clarity Boost guided by perception

After designing the model architecture, the network needs to be trained. The network is initially just a mathematical structure with many parameters. These parameters start with random values, so the network does not yet know how to perform the task. Training is the process of showing the model many examples and gradually adjusting its parameters so that it performs the task correctly. Training a DNN requires defining a “loss function” that measures how well the model performs the task and guides the optimization of its parameters. Traditionally, the loss function measures the error between the model’s output (“denoised audio” in Figure 4) and the desired target (“target audio” in Figure 4). The training objective is to minimize this error by adjusting the model’s parameters. By repeating this process over many training examples, the network gradually learns the patterns needed to perform the task.

The next section explains in more detail the type of loss function and training objective used during training of Clarity Boost.

Perceptually motivated training using feature matching loss

Traditional DNN-based denoising systems are often trained using a loss function that is based on distance measures, i.e., how different the denoised and target audio are in terms of waveforms or signal-to-noise ratio. While effective at reducing noise energy, these metrics do not reliably reflect perceived sound quality and may encourage overly aggressive processing that removes acoustically and perceptually important information.

> *Think of training a DNN like preparing an ice skater for competition. The skater performs a routine with jumps and spins. After the performance, judges assign a technical score based on how well each element was executed compared to the ideal performance. A loss function that only uses distance measures is like only using the technical score to judge an ice-skater performance. However, the judging system in ice skating (as well as in other sports) does not depend on only one criterion. The final score is a combination of multiple aspects, such as technical execution and the overall quality of the performance. It’s not only about the exact execution of the jumps but also about interpretation, overall impression, balance, flow, expressiveness, and interpretation.*

Similarly in audio denoising, beyond simply removing noise and checking the “correctness” of the process with distance measures (“technical score”), one might also care about how natural speech sounds and whether spatial awareness is preserved. This is why the training of Clarity Boost relies on a **Large Audio Foundation Model (LAM)** - to guide the DNN toward producing a denoised audio signal that **preserves those signal qualities that are important for human perception.** The LAM belongs to a new class of self-supervised Generative Audio AI (e.g., Chen et al., 2022) and is trained on vast amounts of audio to learn how sound and speech are naturally structured over time. Because the LAM has been trained on large amounts of audio, it captures perceptually important structures in sound, which helps guide the DNN in Clarity Boost toward outputs that better align with human perception. This allows Clarity Boost to preserve perceptually important signal characteristics rather than optimizing solely for noise suppression, effectively **balancing noise attenuation with natural sound quality.**

During training of Clarity Boost, both the model’s output (denoised audio) and the desired target audio (speech + some residual background sound) are passed through the LAM (Figure 4). Including a controlled amount of residual background sound in the target is a design choice in line with the Widex Sound Philosophy, ensuring that environmental sounds remain audible to support spatial awareness and a natural listening experience. As the signals go through the model, the LAM produces internal feature representations comparing the features of the model’s output with the features of the target audio at multiple layers (see Figure 4), where early stages represent basic acoustic cues (e.g., signal energy) and higher stages represent more abstract aspects of speech (e.g., sound fidelity, linguistic structure, speech integrity, sound quality). The training objective is to minimize the difference (i.e., the error or “loss”) between the two signals at a feature level (“feature loss”), such that the DNN model in Clarity Boost learns to produce audio whose features are similar to those of the desired target signal inside the LAM. This training approach guided by how sound is perceived is known as **feature matching loss.** By combining feature losses from different layers of the LAM, **the training of Clarity Boost balances noise attenuation against perceptually important signal qualities such as naturalness, speech integrity, and preservation of room acoustics.**

Additionally, during training of Clarity Boost, the LAM teaches the AI architecture (i.e., the L-RNN) what information should be stored in the “state” and what information should be discarded, based on what information is relevant for human perception. This selective behavior allows the DNN in Clarity Boost to efficiently store relevant past information in a compressed state, mimicking how the brain remembers relevant information and may forget irrelevant information. As a result, the **timing, structure, and continuity of the original input sound are preserved, supporting predictable behavior and stable sound quality.**

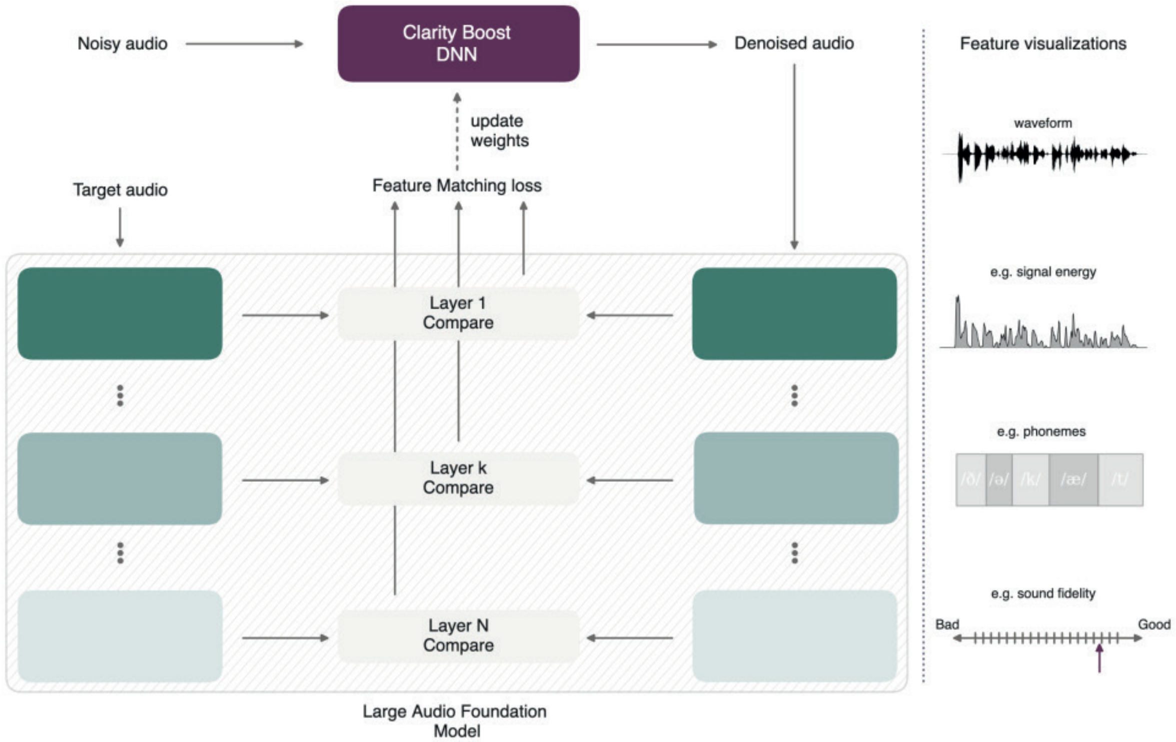
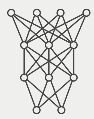


Figure 4: *Illustration of the training procedure of Clarity Boost, guided by how sound is perceived. Clarity Boost is trained by comparing internal representations of the target and denoised audio across multiple layers of a Large Audio Foundation Model. The training objective is to denoise the signal while preserving the signal characteristics that are important for human perception (e.g., naturalness, signal integrity, sound quality). Early layers capture low-level acoustic properties (e.g. signal energy), while deeper layers represent increasingly abstract speech characteristics (e.g. phonemes, sound fidelity).*

Elements of novelty in the DNN of Clarity Boost:



1- Highly efficient audio-specific AI architecture,
leading to:
- Reduction in power consumption
- Same hearing aid size as Allure RIC
- Reduction in processing delay

2- DNN training guided by human perception,
leading to preservation of:
- Spatial awareness
- Naturalness of sound
- Signal quality.

Evaluation metrics

Several evaluation metrics were used during development and model training, guiding design and model choices. These metrics ensure that noise reduction does not come at the expense of sound quality or speech integrity, supporting a balanced listening experience consistent with the Widex Sound Philosophy.

The sound quality of the denoised signal was assessed using neural network-based signal quality predictors, which are sensitive to artifacts and unnatural sound characteristics. Speech preservation was evaluated using automatic speech recognition ML models, by tracking changes in word error rate on denoised signals. In addition, classical objective measures such as signal-to-distortion ratio and time-frequency distance metrics were used to quantify noise attenuation and signal fidelity.

These metrics were used in an overall assessment to guide the model choices during development, complemented with rigorous listening sessions with human listeners, leading to the final Clarity Boost program.

Two-stage training

To ensure robust performance in everyday listening environments, the DNN model in Clarity Boost was trained using a two-stage approach.

1- First, the model was pretrained on a very large and diverse audio dataset to learn general speech and noise characteristics. The pretraining corpus comprised 10 different languages and approximately 35,000 hours (roughly 4 years) of pseudo-unique speech-noise mixtures, covering a wide range of noise types and ecologically valid SNR conditions that reflect real-world listening situations. This stage emphasizes robustness across a wide range of sound scenes.

2- In a second stage, the model was fine-tuned using actual hearing-aid recordings to learn device-specific signal paths. Real-world sound scenes were rendered to a 3D loudspeaker array in a Spatial Audio Laboratory using higher-order ambisonics and recorded with the Allure AI RIC hearing aid placed on a Head and Torso Simulator (HATS). The DNN in Clarity Boost was trained using these recordings that capture device-specific transfer functions and microphone characteristics across multiple spatial configurations. This stage of training ensures robust performance in realistic acoustic conditions, helping to close the gap between laboratory training and real-world performance.

Technical study

Technical measurements of output SNR and group delay were carried out to compare the performance of Allure AI RIC with Clarity Boost relative to all premium competitors with AI, whether they use a dedicated AI co-processor or DNN-based denoising (see Table 1). All hearing aids were fitted using proprietary fitting rationale, N3 standard audiogram (Bisgaard et al., 2010), and closed ear tips. Each hearing aid was programmed with the standard program needed to activate the AI co-processor or DNN-based denoising. Feedback management systems, impulse noise reduction, and wind noise reduction were disabled to avoid interference with the phase-inversion method used for the calculation of output SNR (Hagerman and Olofsson, 2004). All other feature settings were in default for the different programs (see Table 1). The maximum power output was set to maximum for all hearing aids.

Output SNR measurements

Method

In a first set of measurements, five common sound scenarios (canteen, car, party, party babble, train station; Table 2) were reproduced by using eight loudspeakers positioned at angles of 0°, 45°, 90°, 135°, 180°, 225°, 270°, 315° in a circle with a 1.2 m radius in an acoustically treated listening room. A Head and Torso Simulator (HATS) was placed at the center of the circle, facing the front. Target speech was presented from 0° at a fixed level for each sound scene (see Table 2 for details).

Additionally, a second set of measurements were performed in a Spatial Audio Laboratory with a 3D loudspeaker array consisting of 45 loudspeakers on 4 rings. The background noise consisted of a higher-order ambisonics recording of a busy university food court during lunch, taken from the ARTE database (Weisser et al., 2019). The sound scene was rendered to the loudspeaker array using mixed-order ambisonics. The speech material was a continuous monologue recording of an audiobook from a male speaker, recorded as part of the Danish Sentence Test corpus (Kressner et al., 2023). Speech was presented either from 0° or 90° at a fixed level of 73 dB A (Table 2).

For both sets of measurements, the noise level was adapted to achieve realistic input SNRs of 0, +3, +6 dB, in line with the SNRs estimated in real life by Smeds et al. (2015). The phase-inversion method (Hagerman and Olofsson, 2004) was used to calculate the SNR at the output of the hearing aid for each given input SNR. By inverting the phase of the noise signal in one of the recording trials, speech and noise can be separated and calculated at the hearing aid output.

Manufacturer	Model	AI co-processor	Fitting rationale	Program to activate AI
Widex	Allure AI RIC	Yes	Widex rationale	Clarity Boost
Phonak	Audéo Infinio Ultra Sphere	Yes	Adaptive Phonak Digital 3.0	Spheric speech in loud noise
ReSound	Vivia 9 microRIE	Yes	Audiogram+	Hear in noise (incl. Intelligent Focus)
Oticon	Intent 1 miniRITE	No	VAC+	Default (incl. Neural Noise Suppression)
Starkey	Omega AI 24 mRIC R	No	e-STAT 2.0	Default (incl. DNN-speech probability predictor)

Table 1: *List of hearing aids tested in the technical study for output SNR and group delay.*

Scene	Scene description	Target	Target angle	Target level [dB A]
Canteen	Clear distinct speakers in a busy canteen	Female speaker	0°	72.8
Car	Highway noise at 90 km/h, radio playing	Male speaker	0°	65.3
Party	500-people babble, evenly distributed 3-30 m distance	Male speaker	0°	75.3
Party babble	100 people in a reverberant room, scattered within 10-20 m	Male speaker	0°	72.9
Train station	ARTE database (Weisser et al., 2019)	Female speaker	0°	73.9
Food Court (ambisonics)	ARTE database (Weisser et al., 2019)	Male speaker (DAST; Kressner et al., 2023)	0° or 90°	73

Table 2: *Description of the five sound scenarios reproduced in an acoustically treated listening room and the food court scene reproduced in a Spatial Audio Laboratory with a 3D loudspeaker array.*

Results

SII-weighted SNR improvements are depicted in Figure 5 for all six sound scenarios and three input SNRs, when the speech target was presented at 0°. The left panel shows SNR improvements for Clarity Boost, Universal, and PureSound, relative to omnidirectional settings. The right panel shows the SNR improvement Clarity Boost and all four competitors give relative to the unaided condition.

With a median output SNR of 7.8 and a maximum output SNR of 10.1 dB relative to omnidirectional settings (Fig. 5, left panel), **Clarity Boost gives an improvement of up to 5.1 dB in output SNR relative to Universal** and **up to 6.2 dB improvement relative to competitors** when the target is presented at 0° (see Table 3 for improvements in each sound scene). On average, Clarity Boost delivers the **highest output SNR of all four competitors,** at realistic input SNRs across the six sound scenes tested ("Mean SNR improvement", Table 3).

For a side target (90°), Clarity Boost gives up to 7 dB improvement in output SNR relative to Universal and up to 6.3 dB improvement relative to competitors (see Table 4).

Overall, Clarity Boost outperformed all four premium competitors with AI for both frontal and side targets, by achieving the highest average output SNR across six realistic sound environments.

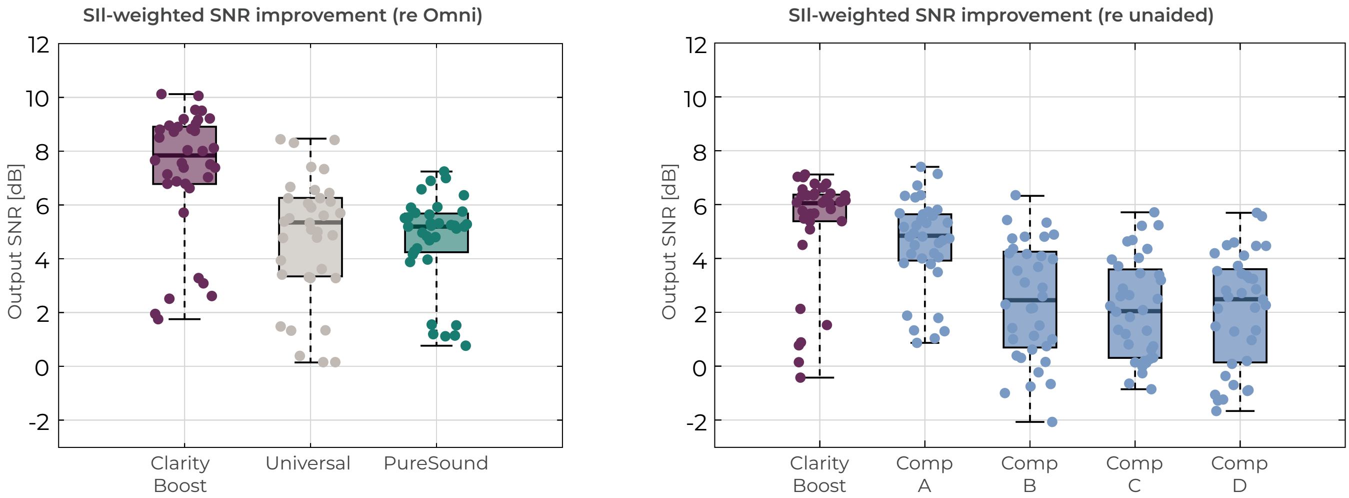


Figure 5: *SII-weighted output SNR (dB) relative to omnidirectional settings (left panel) and relative to unaided (right panel). Each boxplot shows the median output SNR, as well as 25th and 75th percentiles (bottom and top edges of the box, respectively). The whiskers extend to the highest and lowest output SNR. The filled circles depict all 36 measured output SNRs for each hearing aid/program (six sound scenes, three input SNRs, left and right ear).*

Contrast	Mean SNR improvement [dB]	Canteen	Car	Party	Party babble	Train station	Food Court (target at 0°)
Clarity Boost vs. Universal	**2.6**	**2.2**	**3.1**	**0.5**	**1.8**	**5.1**	**2.7**
Clarity Boost vs. Comp A	**0.9**	-0.2	**0.9**	**0.7**	0.3	**1.9**	**1.7**
Clarity Boost vs. Comp B	**4.2**	**0.8**	**4.0**	**5.5**	**5.6**	**6.2**	**3.0**
Clarity Boost vs. Comp C	**3.4**	**1.4**	**5.3**	**1.9**	**3.1**	**5.1**	**3.6**
Clarity Boost vs. Comp D	**3.3**	**2.3**	**5.3**	**2.5**	**2.4**	**4.2**	**3.3**

Table 3: *SII-weighted output SNR improvements (dB) achieved with Clarity Boost relative to Universal and to all four competitors for each tested sound scene and averaged across the tested input SNRs (0, + 3, + 6 dB). Improvements shown for the best ear and frontal target. Improvements equal to or larger than 0.5 are marked in bold. Darker nuances of purple indicate stronger improvements with Clarity Boost (light purple: 0.5-2.4 dB; medium purple: 2.5-4.4 dB; dark purple: ≥4.5 dB).*

Mean SNR improvement [dB]	Food Court (target at 90°-ipsilateral ear)	Food Court (target at 90°-contralateral ear)
Clarity Boost vs. Universal	**7.0**	**4.7**
Clarity Boost vs. Comp A	0.1	**2.4**
Clarity Boost vs. Comp B	**4.5**	**4.0**
Clarity Boost vs. Comp C	**4.1**	**4.2**
Clarity Boost vs. Comp D	**6.3**	**6.3**

Table 4: *SII-weighted output SNR improvements (dB) achieved with Clarity Boost relative to Universal and to all four competitors for the food court scene, averaged across the tested input SNRs (0, + 3, + 6 dB). Improvements shown for a speech target positioned at 90°, for ipsilateral and contralateral ear. Improvements equal to or larger than 0.5 are marked in bold. Darker nuances of purple indicate stronger improvements with Clarity Boost (light purple: 0.5-2.4 dB; medium purple: 2.5-4.4 dB; dark purple: ≥4.5 dB).*

Group delay measurements

Group delay describes how long an input audio signal takes to be processed by the hearing aid. The typical signal processing delays for most premium hearing aids are in the range from 5 to 8 ms (Balling al., 2020), while Widex hearing aids present a much lower delay primarily due to their time-domain approach (Balling et al., 2022). Importantly, the integration of AI co-processors, as well as how the DNN is added to the signal path and the AI architecture (1st, 2nd, 3rd generation), may add substantial processing time to the baseline group delay. Therefore, when evaluating the benefits of an AI co-processor, it is important to consider the impact of added delay, since delays above 10 ms can reduce sound quality and delays around 15 ms can start to affect speech intelligibility (Dillon, 2012).

Method

Group delay was measured for Allure AI RIC with Clarity Boost relative to all premium competitors with AI (listed in Table 1), either with a dedicated AI co-processor (Competitors A and B) or with DNN-based denoising in the main chip (Competitors C and D).

The hearing aids were placed in an acoustic test chamber and connected to a coupler with closed coupling. White noise was played in the test chamber at an overall level of 65 dB SPL. The output was recorded in the reference microphone and the coupler microphone, and a transfer function was computed from which the group delay was derived.

The mean group delay was calculated for each one-third octave band with center frequencies between 500 Hz and 8 kHz. The average delay was then calculated as the overall mean of the individual band means.

Results

Figure 6 depicts the group delay for Widex Allure AI RIC (Clarity Boost, Universal, PureSound) and for all premium competitors with AI-based denoising. Because the addition of an AI co-processor, as well as the denoising approach (1st, 2nd, 3rd generation), may add substantial processing delays, group delays are shown in Fig. 6 (left panel) for all programs/devices with a dedicated AI co-processor for denoising (Clarity Boost, Competitors A and B) and in Fig. 6 (right panel) for all programs/devices that don't use a co-processor for denoising (Universal, PureSound, Competitors C and D).

The average delay for Clarity Boost was 10.0 ms, while higher average delays were obtained for the two competitors with an AI co-processor (Competitors A and B; Fig. 6, left panel). Concerning processing delays without a co-processor (Fig. 6, right panel), Widex PureSound (average delay around 0.5 ms; Balling et al., 2020; Weber and Branda, 2025) and Universal (average delay around 2.6 ms; Balling et al., 2020) achieved much lower delays than competitors C and D. Widex PureSound remains the program with the lowest delay in the industry (Balling et al., 2020).

Overall, Clarity Boost delivered the **lowest average delay among hearing aids with a dedicated AI co-processor** for denoising.

Besides the high efficiency achieved via the third-generation AI architecture, two other factors allow Allure AI RIC with Clarity Boost to keep the delay around 10 ms. First, the flexibility of the W1 chip allows for efficient integration with the AI co-processor, making it possible to limit the additional delay. Second, having a low delay in the default settings (Universal) makes it possible to not exceed the critical 10-ms processing delay even with Clarity Boost.

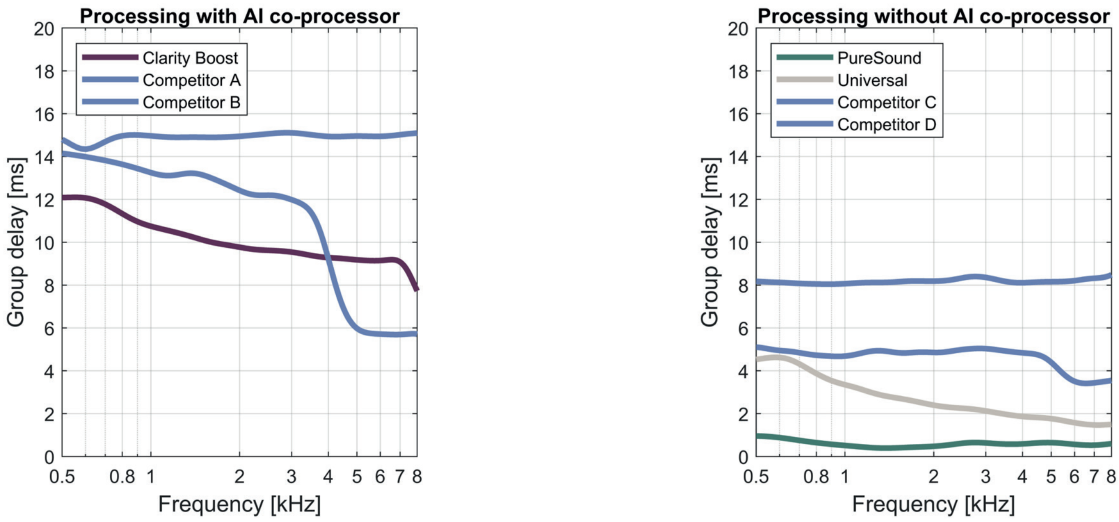


Figure 6: *Signal processing delay [ms] for Widex Allure AI RIC (Clarity Boost, Universal, PureSound) and for all premium competitors with AI, either with a dedicated AI co-processor (Competitors A and B; left panel) or with DNN-based denoising (Competitors C and D; right panel). Curves are smoothed in one-third octave bands.*

Summary and conclusions

Widex Allure AI RIC with Clarity Boost introduces a third generation of DNN-based denoising in hearing aids. This whitepaper has described the main innovations that characterize the DNN in Clarity Boost, focusing on its two primary elements of novelty.

The **first element of novelty** is the **AI architecture.** Based on linear RNNs, this architecture is **specifically built for audio processing,** which allows the DNN to continuously and selectively store information from previous windows of sound into the system's memory ("state"). This enables the DNN in Clarity Boost to very efficiently capture the long-term temporal structure of audio signals. As a result, this third-generation audio-specific AI architecture is far more memory and compute efficient than the second-generation approach, reducing power consumption, hearing aid size, and processing delay.

The **second element of novelty** is the **perceptually motivated training** of the DNN in Clarity Boost. To the best of the authors' knowledge, this is the first time in the hearing aid industry that a denoising DNN model leverages Generative Audio AI during training using a Large Audio Foundation Model to balance noise attenuation against perceptually relevant features of naturalness and sound quality. The result is that Clarity Boost can achieve **superior denoising while preserving what is important for human perception.**

The superior denoising performance of Clarity Boost was proven by a technical study evaluating output SNR in realistic sound environments. The results of the study clearly showed that Clarity Boost outperformed all four premium competitors with AI for both frontal and side targets, by achieving the highest average output SNR across six realistic sound environments.

Some **key highlights of the study** are:

- In realistic sound environments, Clarity Boost achieves a maximum SNR improvement of:
  - 10.1 dB relative to omnidirectional settings
  - 5.1 dB relative to Universal
  - 6.2 dB relative to premium competitors with AI-based denoising.
- Clarity Boost delivers the **highest average output SNR of all four competitors,** at realistic input SNRs across the six realistic sound environments.
- Clarity Boost delivers the **lowest average processing delay among hearing aids with a dedicated AI co-processor,** managing to not exceed the critical 10-ms threshold.

In summary, by combining linear RNNs with perceptually motivated training, the DNN in Clarity Boost unlocks the full potential of AI to boost audio processing to unprecedented levels, whenever the user may need it.

References

Balling, L.W., Townend, O., Stiefenhofer, G., & Switalski, W. (2020). Reducing hearing aid delay for optimal sound quality: a new paradigm in processing. Hearing Review, 27(4), 20–26.

Balling, L.W., Townend, O., Helmink, D. (2021). Sound quality for all: The benefits of ultra-fast signal processing in hearing aids. Hearing Review, 28(9), 32–35.

Balling, L.W., Mosgaard, L.D., Helmink D. (2022) Signal processing and sound quality. Hearing Review, 29(2), 20-23.

Balling, L.W., Vormann, M., Hau, O., and Rose, S. (2025a) The superior sound quality of Widex allure. WidexPress 56.

Balling, L.W., Vormann, M., Mansour, N. (2025b) Widex Allure: a better speech in noise experience with balanced focus and awareness. WidexPress 55.

Balling, L.W., Harris, S., Leung, F., Pineo, K., Helmink, D., Philippon, B. (2025c). Widex Allure: Balanced Focus and Awareness In Real Life. AudiologyOnline, Article 29432.

Bisgaard, N., Vlaming, M. S., & Dahlquist, M. (2010). Standard audiograms for the IEC 60118-15 measurement procedure. *Trends in amplification,* 14(2), 113-120.

Chen, S., Wang, C., Chen, Z., Wu, Y., Liu, S., Chen, Z., ... & Wei, F. (2022). Wavlm: Large-scale self-supervised pre-training for full stack speech processing. IEEE Journal of Selected Topics in Signal Processing, 16(6), 1505-1518.

Hagerman, B., & Olofsson, A. (2004). A Method to Measure the Effect of Noise Reduction Algorithms Using Simultaneous Speech and Noise. Acta Acustica united with Acustica, 90, 356-361.

Herrlin, P., Mansour, N., Nielsen, F., Smeds, K., & Balling, L. W. (2025) The Effect of Speech-To-Noise Ratios On Hearing Aid Noise Reduction Systems in Realistic Environments. Hearing Review, 16-23

Korhonen, P., Kuk, F., Slugocki, C., & Ellis, G. (2022). Low Processing Delay Preserves Natural Cues and Improves Spatial Perception in Hearing Aids. Hearing Review, 29(6), 20-25.

Kressner, A. A., et al. (2023). “The Danish sentence test (DAST) corpus of audio and audio-visual recordings of sentences and monologues”. Proceedings of Forum Acusticum 2023.

Orvieto, A., Smith, S. L., Gu, A., Fernando, A., Gulcehre, C., Pascanu, R., & De, S. (2023). Resurrecting recurrent neural networks for long sequences. Proceedings of the 40th International Conference on Machine Learning (ICML).

Lelic, D., Stiefenhofer, G., Lundorff, E., & Neher, T. (2022). Hearing aid delay in open-fit devices: Preferred sound quality in listeners with normal and impaired hearing. JASA Express Letters, 2(10), 104803.

Slugocki, C., Kuk, F., Korhonen, P., & Ruperto, N. (2020). Neural Encoding of the Stimulus Envelope Facilitated by Widex ZeroDelay Technology. Hearing Review, 27(8), 28-31.

Smeds, K., Wolters, F., & Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *Journal of the American Academy of Audiology,* 26(02), 183-196.

Wagener KC, Hansen M, Ludvigsen C. (2008) Recording and classification of the acoustic environment of hearing aid users. J Am Acad Audiol 19(4), 348–370.

Weisser, A., Buchholz, J. M., Oreinos, C., Badajoz-Davila, J., Galloway, J., Becehey, T., Keidser, G. (2019). “The Ambisonics Recordings of Typical Environments (ARTE) database”. Acta Acustica united with Acustica.

Weber, J. & Branda, E. (2025). Improved speech intelligibility in noise with ultra-low delay. AudiologyOnline, Article 29261.

Zhou, M., Soleimanpour, R., Mahajan, A., & Anderson, S. (2024). Hearing aid delay effects on neural phase locking. Ear and Hearing, 45(1), 142-150.