Brain-Inspired Architecture for DNN-Driven Sound Processing

Heike Heuermann, Ph.D., Amit Shahar, Christine Tan, Ph.D.

Key takeaways:

- The human auditory system continuously adjusts to the environment and strategically shifts to optimize clarity, focus and awareness in real time.
- Audibel's multi-modal, brain-inspired DNN sound processing architecture is designed to move beyond traditionally static algorithms and uses the latest breakthroughs in DNN technology to dynamically interpret and respond to complex listening environments.

Introduction

Human hearing is not defined by a single algorithm or fixed strategy. It is a dynamic, context-aware system that continuously adapts to the listener's environment, intent, and cognitive state. In quiet settings, the brain prioritizes fidelity and subtle detail. In noise, it shifts toward selective attention and signal extraction. In complex, multi-speaker scenarios, it integrates speaker memory, spatial, temporal, and linguistic cues to maintain focus. This adaptability reflects a fundamental principle that the brain employs in that effective hearing is not a one-size-fits-all strategy but one that is able to be situationally intelligent and adapt to the environment that it is in.

Audibel's approach to deep neural network (DNN) processing is architected to enable hearing technology to move closer to this biological model. Rather than apply static rules or isolated features, DNN-driven systems for sound processing are architected around a brain-inspired philosophy that mirrors the multiple, parallel strategies the auditory system naturally employs.

DNN represent a fundamental shift in how sound is analyzed and processed. Rather than relying solely on deterministic algorithms, DNNs can learn from vast datasets of real-world audio to identify patterns, distinguish between speech and noise, and make nuanced decisions towards achieving a trained outcome. This data-driven approach enables hearing technology to better approximate the complexity of human auditory perception, by implementing nuanced learned behaviors and respond to dynamically changing environments.

At a high level, a DNN functions as a multilayered computational model composed of interconnected processing nodes or "neurons". Each layer progressively transforms incoming acoustics into higher-level representations. Having more layers within the network allows extraction of more complex acoustic characteristics. Through a process of training and optimization, the network refines its internal parameters to minimize error and improves its ability to generalize across diverse listening scenarios.

A helpful way to understand DNN is to think of it like a team of expert listeners passing a message along a chain (layers). The first "listeners" pick up basic sound details, like pitch and loudness, and then each successive layer refines that information, recognizing patterns such as speech versus noise, and ultimately deciding what matters most.

By the time the message (signal) reaches the final layer, the system has transformed raw sound into meaningful insight, much like how the brain progressively interprets and prioritizes what we hear in every listening situation.

In hearing technology applications, this capability translates into improved outcomes across a broad range of environments. This whitepaper presents Audibel's advancement in multi-modal DNN hearing technologies, which integrate multiple DNN-powered strategies to effectively interpret and respond to complex acoustic scenes. These innovations underpin the Aris AI+ sound processing strategies *(Figure 1)* to dynamically adapt to the listener's needs and deliver a natural listening experience.

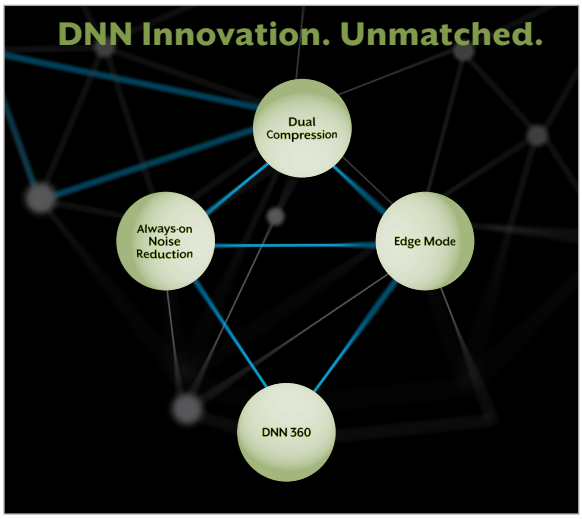

Figure 1: Diagrammatic representation of Audibel's multi-modal DNN systems.

Intelligent Noise Management

For individuals with hearing loss, listening to speech in noisy environments is not only more difficult, but it can also be mentally more demanding, often leading to rapid fatigue. When the auditory signal is degraded, the brain must compensate by allocating greater cognitive resources to decode speech, relying more heavily on attention, working memory and predictive processing. This increased demand is commonly referred to as listening effort, and it has been consistently shown to be elevated in hearing-impaired populations *(Petersen et al., 2015; Alhanbali et al., 2019)*, particularly in complex acoustic environments.

While noise reduction systems in hearing technology have been demonstrated to be more purposeful in improving listening effort than speech reception thresholds *(Sarampalis et al., 2009)*, there are inherent challenges to its design in a hearing aid system.

Noise reduction algorithms must balance suppressing background noise without degrading speech. Over-aggressive processing can inadvertently remove important speech cues, leading to distortion, reduced clarity or compromised sound quality. Many real-world environments also contain competing signals that share acoustic characteristics with speech (e.g. multi-talker babble). Traditional algorithms typically struggle to reliably distinguish target speech from these sources, resulting in limited noise suppression for these situations. Another challenge is that noise reduction systems also need to respond to highly dynamic real-world environments and across various listening situations.

Audibel's DNN-driven noise reduction algorithm *(Betlehem et al., 2024)* is designed to overcome these limitations and is integrated directly into the main audio path of the hearing aid circuit, meaning that it is always available to the hearing aid user.

Directionality reimagined

In complex listening environments, the human auditory system operates in a way that closely resembles a highly adaptive directional microphone system. However, unlike fixed or algorithmic beamforming approaches, the brain achieves directionality through a dynamic integration of binaural cues, neural processing, and attention-driven filtering.

At the most fundamental level, directional hearing begins with binaural input. The brain continuously compares the signal arriving at the two ears, leveraging small differences in timing to determine the direction of incoming sound. These cues are encoded with remarkable precision at early stages of the auditory pathway, beginning in the brainstem and propagating through higher auditory centers. This binaural comparison effectively provides a spatial map of the acoustic environment, forming the foundation of directional selectivity *(Von Kriegstein et al., 2008)*.

Importantly, the brain's "directional processing" extends beyond simple front-focused listening. Depending on the situation, listeners can adopt different strategies such as broad environmental awareness or narrow focus on a specific speaker and demonstrate a level of flexibility that exceeds conventional directional microphone systems. This adaptability is essential for real-world communication where listening goals and environments are constantly changing.

This biological attribute of combining information from multiple inputs, coupled with advanced DNN processing for higher-level interpretation forms the basis for Audibel's DNN 360 feature (comprising both the DNN Directionality and DNN Spatial Awareness features *[Marquardt et al., 2026]*) and balances the nuanced challenge of having listeners engage naturally in conversation while remaining attuned to their surroundings.

The overall utility of the directional processing in conjunction with noise management showed significantly better performance than amplification alone, with the Aris AI hearing aids performing significantly better than other hearing aids previously used *(Halvorsen & Iverson, 2026)*.

Biologically inspired compression

The human auditory system manages an extraordinary challenge in its ability to process sounds that span an immense dynamic range, from faint whispers to loud environmental noise, and all while preserving the clarity of the speech and the quality of the sound heard. To achieve this, the brain operates much like a dual compression system, simultaneously applying both fast- and slow-acting gain control mechanisms across different levels of the auditory pathway.

At the peripheral level, cochlear compression serves as a first stage of dynamic range management. Mechanical and neural processes within the cochlea reduce the effective range of incoming sound pressure and transform a wide range of acoustic inputs into a more manageable range of neural outputs. This nonlinear compression ensures that softer sounds remain audible while louder sounds do not become overwhelming, effectively mirroring the role of wide dynamic range compression (WDRC) in hearing devices.

Beyond the cochlea, the auditory system introduces time-dependent compression through neural adaptation, which functions across multiple timescales. At a very short timescale, fast-acting adaptation mechanisms respond within milliseconds to changes in acoustic input. These rapid responses enhance the neural representation of transient, speech-relevant features such as consonant bursts and amplitude modulations, particularly in fluctuating noise backgrounds.

This fast compression-like process effectively increases access to brief "dips" in background noise, allowing important speech cues to emerge more clearly *(Jennings & Dominguez, 2022)*.

In parallel, the brain applies a slow-acting compression mechanism, driven by longer-term adaptation to the statistical properties of the acoustic environment. Neural responses across the auditory pathway gradually adjust to sustained sound levels, reducing sensitivity to steady-state background noise while maintaining responsiveness to changes or new stimuli. This slower form of gain control supports listening comfort and stability, preventing the auditory system from over-amplifying persistent sounds while preserving overall signal fidelity *(Pérez-González & Malmierca, 2014)*.

Critically, these fast and slow processes operate simultaneously creating a layered compression system analogous to combining fast- and slow-acting compression in hearing technology and forms the basis for Audibel's biologically inspired dual compression architecture.

Audibel's compression architecture utilizes independent, parallel fast and slow compression systems that work together and are designed to optimize both clarity and comfort. Fast compression targets rapidly changing elements of the signal to enhance audibility of soft sounds and improve speech contrast, while slow compression manages overall sound levels based on the patient's dynamic range to maintain comfort and preserve sound quality and spatial cues.

Building on this foundation, Aris AI+'s intelligent compression system incorporates a DNN-driven controller that dynamically balances these processes that help moderate fast compression when needed for clear speech while preserving sound quality *(Jensen et al., 2026)*. The result is a more adaptive compression strategy that delivers consistent audibility, natural sound quality and an overall improved listening experience.

Intent-aware intelligence

The auditory system continuously recalibrates based on environmental conditions and listening intent. As acoustic scenes shift, for instance from quiet one-to-one conversation to more complex, dynamic multi-speaker environments, the brain seamlessly adjusts how it processes sound. This recalibration is driven by an ongoing interplay between sensory input and cognitive control, enabling the listener to prioritize what is most relevant while maintaining awareness of the broader environment. A critical component of this adaptability is *listening intent*. Whether focusing on a single speaker, monitoring surrounding activity or engaging in a group conversation, intent shapes how the brain allocates attention, applies spatial filtering and enhances specific sound features. In essence, the auditory system dynamically "tunes" itself to match the listener's objective, rather than applying a fixed processing strategy.

For hearing technology, this highlights a critical design philosophy: effective sound processing is not defined solely by the environment, but by how the user intends to engage with it. Systems that rely only on environmental classification risk misalignment. These systems typically optimize for the "scene" without understanding the listener's goal.

Audibel's Edge Mode embraces this core philosophy to respect the listening intent and its DNN-driven adjustments help provide listeners with hearing aid settings that are tailored to specific listening intents and listening environments.

With Aris AI+, the Edge Mode feature was further enhanced for better performance in high levels of noise *(Taylor & Marquardt, 2026)*.

Conclusion

The human auditory system remains the gold standard for processing sound in the way it integrates a variety of complex processes to deliver clarity in even the most complex listening environments. For individuals with hearing loss, however, this natural system is compromised. Addressing these challenges often requires more than a single sound processing panacea and demands a fundamentally new approach.

Audibel's biologically inspired DNN sound processing represents a significant step forward in hearing technology. By drawing from principles of how the brain listens, Audibel's multi-modal DNN processing moves beyond traditionally static algorithms towards intelligent, context-aware sound processing. These capabilities are unified within the Aris AI+ product line, where multiple DNN processes work in concert to interpret complex soundscapes and intents to deliver clear, natural sounding listening experiences.

As hearing technology continues to evolve, the convergence of neuroscience and artificial intelligence will play an increasingly important role in shaping innovation. Audibel's commitment to a biologically inspired DNN architecture reflects this vision to bridge the gap between human auditory perception and advanced digital systems. The result is a new generation of hearing solutions that intelligently adapt to the listener's needs and empowers the hearing aid users to engage more effortlessly, more confidently and more fully in the world around them to hear better and live better lives.

References

1. Petersen E.B., Wöstmann M., Obleser J., Stenfelt S. & Lunner T. (2015) Hearing loss impacts neural alpha oscillations under adverse listening conditions. Front. Psychol. 6:177.
2. Alhanbali, S., Dawes, P., Millman, R.E. & Munro, K.J. Measures of Listening Effort Are Multidimensional. Ear and Hearing 40(5):p 1084-1097, September/October 2019.
3. Betlehem, T., Mishra, P., Xu, J., Marquardt, D. & McKinney, M. (2024) Leveraging DNN in Starkey Edge AI. Starkey white paper.
4. Von Kriegstein, K., Griffiths, T. D., Thompson, S. K., & McAlpine, D. (2008). Responses to interaural time delay in human cortex. Journal of neurophysiology, 100(5), 2712-2718
5. Marquardt, D., Xiao, J., Ganeshkumar, A., Xu, J., Taylor, L. & McKinney, M. (2026) Enhancing Directionality with Deep Neural Networks. Starkey white paper.
6. Halvorsen, S.C. & Iverson, S. R. (2026) Clinical Validation of the Omega AI family. Starkey publication.
7. Jennings, S. G., & Dominguez, J. (2022). Firing rate adaptation of the human auditory nerve optimizes neural signal-to-noise ratios. JARO: Journal of the Association for Research in Otolaryngology, 23(3), 365.
8. Pérez-González, D., & Malmierca, M. S. (2014). Adaptation in the auditory system: an overview. Frontiers in integrative neuroscience, 8, 19.
9. Jensen, K., Xu, J., Liu, J., Kalluri, S., Schepker, H., & Waite, B. (2026) Optimizing Hearing Aid Sound Quality Through Deep Learning Control of Compression. Starkey white paper.
10. Taylor, L. & Marquardt, M. (2026) Edge Mode performance that stands out in noise. Starkey white paper.
11. Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. Journal of Speech, Language, and Hearing Research, 52(5), 1230-1240.

Author Biographies

Heike Heuermann, Ph.D. is the Vice President of Product Integration at Starkey, where she bridges complex hearing aid technology with patient-driven goals. Holding a Ph.D. in Cognitive Psychology and a Master's in Physics, she previously spent years in senior leadership at WS Audiology and Siemens/Sivantos, directing global R&D and platform strategies. Today, she leverages her extensive background in audiology and product development to help advance Starkey's established vision of delivering patient-centric, personalized hearing aid features.

Amit Shahar, is the Senior Vice President of Advanced Development at Starkey, where he leads global engineering and established the company's advanced R&D center in Tel Aviv. Drawing on his previous experience as a Senior Director at Intel, Shahar has driven Starkey's transition from traditional audio amplification to AI-powered health tech. Under his leadership, his team pioneered integrating Deep Neural Networks directly onto hearing aid silicon chips for real-time sound processing, while also introducing other groundbreaking hearing aid features.

Christine Tan, Ph.D. is a Senior Market Analyst at Starkey, with a unique, multidisciplinary background that spans a BEng. in Electronics and Electrical Engineering, an MSc. in Audiology, and a Ph.D. in Psychology. As a second-generation audiologist with a deep-rooted passion for industry, she has extensive experience across research and product management. In the current role, she supports the business through streamlining processes, strategic competitive and market insights, and technical storytelling.

Audibel, Audibel logo, Edge Mode, Aris and DNN 360 are trademarks of Starkey Laboratories, Inc.
©2026 Starkey Laboratories, Inc. All Rights Reserved. 8/26 WTPR2869-00-EN-AB