OBJECTIVE AND PERCEPTUAL EVALUATION OF DEEP LEARNING-BASED NOISE REDUCTION IN A COMMERCIAL HEARING AID

KHEJLA FATHIMA P. K.

Registration No: P01II24S023016

A Dissertation Submitted in Part of Fulfilment of the

Degree of Master of Science

(Audiology)

University of Mysore

ALL INDIA INSTITUTE OF SPEECH AND HEARING,

MANASAGANGOTHRI, MYSURU -570006

JUNE, 2026

CERTIFICATE

This is to certify that this dissertation entitled **"Objective and Perceptual Evaluation of Deep Learning-Based Noise Reduction in a Commercial Hearing Aid "** is a bonafide work submitted as part of fulfilment of the degree of Master of Science **Audiology** of the student with Registration Number **P01II24S023016**. This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other university for the award of any other Diploma or Degree.

Mysuru
June, 2026

Dr. M. Pushpavathi

Director
All India Institute of Speech and Hearing
Manasagangothri, Mysuru-570 006

CERTIFICATE

This is to certify that this dissertation entitled **"Objective and Perceptual Evaluation of Deep Learning-Based Noise Reduction in a Commercial Hearing Aid "** has been prepared under my supervision and guidance. It is also being certified that this dissertation has not been submitted earlier to any other university for the award of any other Diploma or Degree.

Mysuru
June, 2026

Dr. P. ARIVUDAI NAMBI
Guide
Associate Professor
Department of Audiology
All India Institute of Speech and Hearing
Manasagangothri
Mysuru-570 006

DECLARATION

This is to certify that this dissertation entitled **" Objective and Perceptual Evaluation of Deep Learning-Based Noise Reduction in a Commercial Hearing Aid "** is the result of my own study under the guidance of **Dr. P. ARIVUDAI NAMBI, Associate Professor of Audiology**, All India Institute of Speech and Hearing, Mysuru and has not been submitted earlier to any other university for the award of any other Diploma or Degree.

Mysuru
June, 2026

Registration no: P01II24S023016

ACKNOWLEDGEMENT

"In the name of Allah, the most gracious, the most merciful."

*I am deeply grateful to have **Dr P. Arivudai Nambi** as my guide and mentor. The insightful suggestions, motivating words and dedicated mentorship have greatly contributed to the successful completion of this work. His encouragement and critical insights truly shaped the direction and quality of this work.*

*I sincerely thank **Dr M. Pushpavathi**, Director, AIISH, for giving me the opportunity to carry out this dissertation.*

*I extend my sincere thanks to **Dr Ajith Kumar. U**, Chairperson of the Centre for Hearing Sciences, and the entire team, including the faculty members and JRFs alike, for creating a conducive environment and enabling the completion of this work.*

*This work would not have been possible without the invaluable assistance of several individuals. I am particularly thankful to **Miss Ananya** for her dedicated support throughout this process. I would also like to acknowledge **Mr Kamalakannan, Miss Varsha, Mr Rakesh, Mr Suraj, Miss Nayana** and **Mr Chinnarasu** from the Centre for Hearing Sciences for their help and encouragement during my research journey.*

*I want to thank my dissertation partner, **Sharath Aradhya**, for being an excellent collaborator and source of support throughout this research journey.*

I also extend my heartfelt appreciation to all the participants of the study who generously gave their time and trust to this research.

*To my anchors, **Umma and Uppa** who always had been there for me through my highs and lows with silent prayers and constant motivation. Thank you for the boundlesss love, countless sacrifices and trust in me. To my sister **Khenza** for checking up on me even when I go silent and I am forever grateful to have you. I would also like to thank my **Ummachi** for her endless love and duas which had kept me going. To **Vaseem Ikka** — you came into this family not too long ago, but it already feels like you've always belonged here.*

*I would also like to extend my regards to **Anagha** for your heartfelt words and motivation whenever I needed. Thank you **Hiba** for always being there just a call away whenever I needed and all the late night preparations. I would also like to thank my posting partners **Aishwarya, Abu and Karunya** for making postings memorable and the companionship. I would also like to thank all my batchmates for the memories and learnings I had in the past two years.*

*I am truly blessed to have people like **Uma, Bhavya** and **Yanam** in my life. **Uma** - you never needed me to explain myself to understand me. You just knew. Thank you for every moment, every quiet presence, every time you simply showed up. **Bhavya** - my go-to person, the one I reached for when I needed to feel grounded again, when I needed perspective, or when I just needed someone to be there. **Yanam** - not every bond needs words. The silence between us was always the comfortable kind, the kind that feels like home. **Floriyechi, Amalettan, and Ameenikka** - Your care reached me in ways I never knew I needed, and I held onto it on every difficult day of this journey. **Ayshatha, Niranjanechi, Sneha chechi, Naziratha,** and **Aleenechi** - your guidance meant more than you know, and having you ahead of me on this path made it much better.*

I would also like to acknowledge each and everyone who have helped me, directly or indirectly for the completion of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Individuals with sensorineural hearing loss experience significant difficulty understanding speech in noise due to reduced audibility and degraded auditory processing. While deep learning-based noise reduction (DNN-NR) has been integrated into commercially available hearing aids, independent evaluations using ecologically valid and linguistically relevant materials remain limited. The present study investigated the efficacy of an AI-based deep learning algorithm in enhancing speech perception in noise using the Phonak Audéo Sphere Infinio i90 hearing aid among twenty-five native Kannada-speaking adults (M = 59.32 years) with bilateral moderate-to-moderately severe sensorineural hearing loss and no prior hearing aid experience.

Speech reception thresholds in noise (SRTn) were measured adaptively using standardised Kannada sentences presented against four-talker Kannada babble and locally recorded traffic noise under co-located and spatially separated conditions, comparing the DNN-NR program against a traditional noise reduction program. Objective SNR improvement was additionally quantified using the modified phase-inversion technique. Generalised linear mixed model analysis revealed significant main effects of spatial condition, noise type, and noise reduction program on SRTn.

DNN-NR yielded significantly better SRTn than Traditional-NR only under babble noise in the co-located condition (difference = 1.69 dB, $p = 0.03$), with no significant differences observed under spatially separated babble or traffic noise in either spatial condition. Spatial separation consistently improved performance across all conditions, with greater benefit in babble (~4–5 dB) than traffic noise (~2 dB), and 76% of participants subjectively preferred the DNN-NR program. These findings indicate that DNN-based noise

reduction provides context-dependent perceptual benefits, with the greatest advantage in co-located multi-talker conditions involving informational masking, underscoring the need for condition-specific evaluation of AI-based hearing aid technologies across diverse clinical and linguistic populations.

Keywords: deep neural network, noise reduction, sensorineural hearing loss, speech perception in noise, speech reception threshold, informational masking, spatial separation, Kannada

CHAPTER 1

INTRODUCTION

Individuals with sensorineural hearing loss (SNHL) experience difficulty understanding speech in background noise not only because of reduced audibility, but also due to degraded auditory processing abilities such as poorer frequency resolution and temporal processing (Dong et al., 2024; Schum, 2000; Voll, 2000). As a result, they often require substantially higher signal-to-noise ratios (SNRs) than individuals with normal hearing to achieve comparable speech recognition performance (Deniz et al., 2024; Aghasoleimani et al., 2020). Although amplification improves audibility, it does not fully compensate for the reduced ability to segregate speech from competing background noise. Even with sufficient amplification, many hearing aid users continue to face severe communication difficulties in noisy situations (Dong et al., 2024; Deniz et al., 2024; Miller et al., 2017; Aghasoleimani et al., 2020). Poor SNR also increases listening effort, cognitive load and fatigue in everyday communication, thus influencing the overall communication efficiency and quality of life (Chong & Jenstad, 2018; Baboukani et al., 2022). Therefore, a key to improving speech understanding in listeners with SNHL is likely to be increasing the effective SNR available to them. Hearing aid signal processing has incorporated several enhancement strategies to compensate for SNR deficits in subjects with SNHL. Among the most well-established are directional microphone systems and classical single-channel digital noise reduction algorithms, namely spectral subtraction and Wiener filtering.

Directional microphones exploit spatial filtering by attenuating acoustic signals arriving off-axis while preserving frontal incidence, with adaptive and beamforming configurations dynamically steering the polar sensitivity pattern toward the target source. This mechanism yields SNR improvements of approximately 1–4 dB or greater, with

corresponding gains in speech recognition, reduced listening effort, and improved device usage time in both adult and pediatric populations (Chung, 2004; Magnusson et al., 2013; Wolfe et al., 2022; Wu et al., 2018; Jürgens et al., 2025). However, performance is highly dependent on acoustic coupling and fitting configuration. Open-fit implementations demonstrated reduced directional benefit, with gains as low as 1.6 dB SNR (Magnusson et al., 2013; Jürgens et al., 2025). Degradation also happens when the target source is off the frontal axis, or when automated mode-switching algorithms fail to adapt to dynamic real-world acoustic scenes (Wolfe et al., 2022; Wu et al., 2018). Additionally, ecologically valid outcomes demonstrate a minimal variation in performance between premium and basic directional systems, despite benefits measured in the lab (Wu et al., 2018).

Single channel noise reduction techniques such as spectral subtraction and Wiener filtering rely on an estimation of the noise power spectral density, typically during non-speech segments, and a frequency dependent gain attenuation to suppress the noise dominated spectral components.Hearing-aid adapted Wiener filter implementations with filter bank decomposition and dynamic range compression have shown perceptual quality metric improvements such as MOS and PESQ scores in addition to word recognition improvements of up to 27% under low-SNR conditions in moderate SNHL (Pujar, 2019). Comparative evaluations reveal that both algorithms achieve favourable quality-intelligibility trade-offs and low computational latency, thereby making them optimal for real-time processing constraints (Trifosa & Wulandari, 2025; Rohith & Bhandarkar, 2021). Extended variants, including smoothed Wiener and NNSC-based Wiener formulations, show further reductions of residual noise under non-stationary conditions (Kang & Gong, 2024; Kim, 2020). Despite these attributes, classical implementations in commercial devices have consistently resulted in improved listening comfort without corresponding gains in speech intelligibility, a fundamental perceptual trade-off (Chung, 2004; Jürgens et al., 2025). Gain attenuation

applied to noise-dominant frequency bands inadvertently reduces speech audibility within those same bands, constraining recognition benefit (McCreery et al., 2012; Chung, 2004). Furthermore, both algorithms exhibit degraded performance under non-stationary and multi-talker babble conditions, and are susceptible to the introduction of musical noise artifacts arising from imprecise spectral noise estimation (Chung, 2004; Sheikhzadeh et al., 1994; Jürgens et al., 2025).

To mitigate the constraints of traditional hearing aid processing, artificial intelligence (AI) has emerged as a viable paradigm through adaptive, context-sensitive signal processing. Compared to the traditional methods, which rely on manually designed algorithms, AI and its subset, machine learning (ML), utilise adaptable models trained on representative datasets, with deep learning specifically employing multilayered neural networks to map complex input-output relationships (Andersen et al., 2021; Bishop, 2006; Copeland, 2015). In case of complex multi-talker and highly non-stationary auditory environments, where the traditional methods are inadequate, the application of deep neural networks (DNN) yielded significant enhancements in intelligibility (Andersen et al., 2021; Diehl et al., 2022; Klein et al., 2025; Park et al., 2020). According to algorithmic research, these neural networks can produce efficient postfilter gains or use U-Net encoder-decoder architectures to minimize loss functions and restore degraded inputs when trained iteratively on paired clean and noisy signals using stochastic gradient descent and backpropagation (Anderson et al., 2021; Diehl et al., 2023). By classifying inputs and learning rich spectro-temporal patterns, architectures such as convolutional neural networks (CNNs) effectively isolate target speech from heterogeneous noise mixtures in real-time, thereby lowering cognitive burden, improving listening comfort, and mitigating the classic trade-off between aggressive noise suppression and speech distortion (Ashkanichenarlogh et al., 2025; Bhat et al., 2019; Essaid et al., 2025; Park et al., 2020; Zhang et al., 2025). This AI-driven processing

yields profound clinical implications, demonstrating sufficient improvements in objective speech quality and real-world intelligibility to alter cochlear implant candidacy criteria and significantly enhance bimodal hearing outcomes (Kolberg et al., 2025; Saoji et al., 2024). While these results collectively highlight deep learning's ability to overcome traditional algorithmic constraints, research indicates that the strong performance of such cutting-edge technologies in controlled laboratory settings frequently falls short of user expectations in dynamic, real-world applications (Mustafa et al., 2025). This gap between technology development and clinical validation is an important research gap. The aim of this study is therefore to systematically evaluate the real-world efficacy of AI-based algorithms in commercially available hearing aids and its impact on speech perception in ecologically valid listening conditions.

1.1 **Need for the study**

Deep learning based noise reduction (DNN-NR) has been a major innovation in hearing aid technology and is now part of several commercial hearing aids. Improvements in SNR, speech intelligibility, listening effort and user preference over conventional noise reduction algorithms have been reported by various manufacturer-sponsored studies and companies' own investigations (Andersen et al., 2021; Healy et al., 2023; Fitzgerald et al., 2025; Ashkanichenarlogh et al., 2025). These benefits have been proven by objective bench testing, controlled speech-in-noise testing and by real-world user testing. For instance, DNN-based processing in commercial hearing aids has been shown to improve speech intelligibility, selective attention and listening effort (Andersen et al., 2021), while other studies have reported improvements in speech recognition scores and objective SNR measures on the order of about 4-5 dB(Healy et al., 2023). However, a significant proportion of the available evidence arises from studies that were conducted, funded or co-authored by hearing aid manufacturers like as Sonova, Starkey, and Sivantos (Healy et al., 2023; Fabry & Bhowmik,

2021; Schröter et al., 2020). Consequently, independent validation is essential to establish the generalizability, transparency, and clinical relevance of these findings.

Independent evaluations are especially important because gains observed in objective acoustic measures do not always correlate with perceptually meaningful benefits for hearing aid users. This concern has been raised by recent independent investigations. In a laboratory evaluation of the deep learning algorithm implemented in Oticon More, there was a modest objective SNR advantage, and perceptual benefits were confined to listening situations with spatially separated maskers, and not observed in co-located speech and noise or dual-task listening (Raj et al., 2025). The findings suggest that the benefits of DNN-based processing could be highly dependent on the listening environment and the task at hand. Thereby, highlighting the importance of systematic assessment of the individual contexts in which these algorithms provide clinically meaningful benefits.

Furthermore, most commercially available DNN algorithms are developed and trained using predominantly Western speech corpora and environmental noise databases (Fabry & Bhowmik, 2021; Khan et al., 2025). These systems are subsequently fitted to users from diverse linguistic and cultural backgrounds, including India, with limited evidence regarding their effectiveness in local contexts. Indian languages possess distinct phonetic, phonological, and prosodic characteristics that may influence the performance of speech enhancement algorithms. In addition, the acoustic characteristics of everyday listening environments in India, including transportation noise, crowded public spaces, and occupational noise exposures, differ considerably from those represented in the datasets used to train many commercial DNN systems. Consequently, the real-world benefits of these technologies for Indian hearing aid users remain largely unknown.

Given the increasing adoption of AI-driven hearing aid technologies in clinical practice, there is a pressing need for independent and scientifically rigorous evaluations using

culturally and linguistically relevant speech materials and ecologically valid noise conditions. Such investigations will provide unbiased evidence regarding the benefits and limitations of commercially available DNN-based noise reduction systems, identify the listening situations in which they are most effective, and generate clinically meaningful data to support evidence-based hearing aid fitting and rehabilitation practices. Ultimately, independent validation will help bridge the gap between manufacturer-reported performance and real-world clinical outcomes, ensuring that the adoption of AI-based hearing aid technologies is guided by robust scientific evidence rather than industry claims alone.

1.2 Aim

To investigate the efficacy of artificial intelligence–based deep learning algorithm in enhancing speech perception in noise in commercially available hearing aid.

1.3 Objectives of the study

- To objectively quantify the signal-to-noise ratio (SNR) improvement provided by the DNN algorithm using the modified phase-inversion technique.
- To evaluate the perceptual benefit of the deep learning algorithm on speech reception thresholds in noise (SRTn) under co-located and spatially separated noise condition.

CHAPTER 2

REVIEW OF LITERATURE

Noise reduction in hearing aids is essential for effective interaction among individuals with hearing impairments, particularly in the presence of background noise, and for speech perception. These individuals would face various difficulties, including misunderstanding information, difficulty understanding speech, and reduced ability to participate in conversation. A variety of deficits in the impaired hearing system have been identified through psychophysical experiments, which are not associated with a simple loss of sensitivity. These consist of a diminished capacity to use spatial cues, frequency spread of masking, and temporal spread of masking (Moore, 2007).

Historically, hearing aids have been designed to amplify incoming acoustic signals, leading to indiscriminate amplification of background noise. As a result, the signal-to-noise ratio (SNR) is reduced, making it more difficult for the hearing-impaired user to concentrate on desired sounds, such as conversations. To achieve comparable performance in challenging listening conditions, individuals with sensorineural hearing loss often require a higher SNR than those with normal hearing (Deniz et al., 2024). To address this, various technological interventions have been developed and evaluated. Methods to enhance SNR and overall speech intelligibility can be systematically categorised into three main areas: wireless technologies, beamformers, and digital noise reduction and adaptive processing.

2.1 Wireless Technologies

The wireless technologies, especially remote microphones by utilizing frequency modulation, digital platforms, or proprietary protocols, consistently produce significant signal-to-noise ratio and speech-in-noise gains compared to the use of hearing aids alone (Wang et al., 2021; Ciorba et al., 2014; Thibodeau, 2019; Rodemerk & Galster, 2015;

Ibrahim et al., 2012; Mehrkian et al., 2019; Zanin et al., 2024; Vroegop et al., 2018; Thibodeau, 2014). These significant improvements have been documented across diverse populations, including both adult and pediatric users outfitted with either hearing aids or cochlear implants.The magnitude of these improvements is large with research showing improvements of the order of 10 to 20 dB in the signal-to-noise ratio as well as large percentage improvements in word and sentence recognition scores (Wang et al., 2021; Ciorba et al., 2014; Ibrahim et al., 2012; Mehrkian et al., 2019; Zanin et al., 2024; Thibodeau, 2014). In the most challenging and difficult noise conditions, these wireless interventions sometimes allow users to achieve performance that meets or even exceeds the capabilities of normal-hearing listeners. Furthermore, the acoustic benefits provided by remote microphones are quite robust to physical distance and environmental reverberation, preserving stable performance levels even when the sound source is located 1.5 to 6 meters away from the listener (Wang et al., 2021; Rodemerk & Galster, 2015; Zanin et al., 2024).

The everyday connectivity to external multimedia devices, in addition to remote microphones, further enhances the listening experience. The direct wireless connection to televisions, mobile phones and other audio sources significantly improves the signal-to-noise ratio and everyday usability (Ciorba et al., 2014; Thibodeau & Johnson, 2014; Lazzerini et al., 2023; Fabry & Bhowmik, 2021). Binaural wireless streaming is particularly effective at improving speech perception of bilateral hearing aid users during telephone calls and in the presence of environmental wind noise. Wireless applications are also relevant for unilateral hearing loss. Wireless auditory training with a remote microphone has been shown to further improve speech-in-noise results and self-reported hearing abilities (Lovato et al., 2024) . Likewise, wireless contralateral routing of signal devices and bone-anchored solutions provide significant speech-in-noise benefits to unilateral listeners, but with no corresponding localisation benefit (Snapp et al., 2016).

Despite these profound objective benefits, wireless technologies are subject to practical limitations. Their effectiveness is heavily dependent on proper setup and user behavior. To achieve optimal results, users must ensure the correct placement of the remote microphones and select the appropriate operational modes, such as deliberately turning off the localized hearing aid microphones in extremely noisy environments (Ciorba et al., 2014; Thibodeau, 2019; Rodemerk & Galster, 2015; Zanin et al., 2024). Additionally, while these systems improve speech clarity, they offer a highly limited impact on natural sound localization. Some wireless synchronization technologies provide a slight improvement in front-to-back localization, but they fail to improve left-to-right localization, and contralateral routing of signal devices do not restore any localization abilities for monaural listeners (Ibrahim et al., 2012; Snapp et al., 2016).

2.2 Beamformers

In beamforming technology, spatial filtering is used to improve the signal to noise ratio by focusing on sounds coming from the front of the listener and rejecting noise around them. In wireless integration, wireless binaural beamformers are objectively superior to bilateral omnidirectional microphones in speech-in-noise tasks, but there is evidence that they are not clearly superior to other directional strategies available in the market (Kumar et al., 2023). Directional and adaptive digital remote microphones consistently outperform their simple omnidirectional remote devices in multi-talker or complex environments such as restaurants (Thibodeau, 2019; Thibodeau, 2014). The spatial advantage is also very useful for bimodal users, who use a cochlear implant in one ear and a hearing aid in the other. For these individuals, the addition of a wireless directional remote microphone improves speech-in-noise perception, with dedicated directional modes providing an additional layer of measurable benefit in comparison to standard omnidirectional modes (Vroegop et al., 2018).

One of the main limitations of beamforming technologies is the permanent gap between laboratory objective results and the subjective user perception. Although, even if the wireless beamforming clearly improves clinical speech-in-noise test scores, often the self-reported advantage over other microphone modes is not always large, and the overall quality of evidence for subjective improvements is low (Kumar et al., 2023).

2.3 Digital Noise Reduction and Adaptive Processing

The physical microphones capture the sound and the dynamic nature of the auditory environment is handled through digital noise reduction and adaptive pre-processing algorithms. A lot of the high end remote mics use digital algorithms that adapt their processing depending on the noise floor of the space. Compared to traditional frequency modulation remote microphones or basic omnidirectional models, adaptive digital remote microphones have been shown to provide much better performance at the highest noise levels (Thibodeau, 2014; Ibrahim et al., 2012). The researchers are continuously assessing the efficacy of computational onboard methods, and are actively investigating whether physical remote microphones still outperform these sophisticated pre-processing algorithms under a wide range of listening conditions (Lazzerini et al., 2023).

The principal difficulty in evaluating digital noise reduction and complex adaptive processing is simulating real-world complexity. Most of the lab based studies on these algorithms uses controlled, stationary noise sources. Social environments and restaurants in the real world, by contrast, provide quite diffuse and variable noise environments. Hence the unique advantages reported in laboratory environments tend to exaggerate the real improvements experienced by users in everyday life (Thibodeau, 2019; Zanin et al., 2024; Mustafa & Krishnamurthy, 2025). Systematic reviews of literature underscore these evidence gaps, thereby stating that even with continuous improvements in pre-processing and digital algorithms, many users still experience significant difficulties with noise, and the real-life

benefits of these complex technologies are uncertain and highly dependent on the individual user (Alexander, 2021; Mustafa & Krishnamurthy, 2025).

2.4 Evolution of Artificial Intelligence Applications in Hearing Devices

Conventional hearing aids amplify incoming acoustic signals but exhibit significant limitations in complex acoustic environments, which often results in reduced user satisfaction. Today's hearing aids use artificial intelligence (AI) to solve these problems. The term AI was created by John McCarthy in 1955 and is defined as "the science and engineering of making intelligent machines." AI refers to the ability of machines to carry out tasks that are normally done by humans, such as learning, problem solving and decision making.

The machine learning (ML) in this technology enables systems to learn from data and improve their performance over time, rather than being programmed with fixed instructions. The auditory technology utilises an even more sophisticated subset of machine learning, Deep Neural Networks (DNNs), or deep learning. These systems equip machines with intelligence through algorithms that are explicitly inspired by the biological structure and functioning of the human brain. DNNs apply many layers of processing to recognise very complex patterns, for example to distinguish speech from background noise in real-world situations with high accuracy.

2.5 Enhancing the Auditory Experience through Intelligent Processing

Hearing aids have evolved from simple sound amplifiers to intelligent listening systems that can generate very adaptive and personalised sound experiences with the use of AI, ML and DNNs. Such technologies enable the devices to constantly monitor the surrounding acoustic environment and to dynamically adapt to the specific auditory needs of the user. The devices can monitor the environment in real time and detect different sound scenes, such as speech, background noise or music, and then automatically adjust themselves to optimise overall clarity and comfort. By executing precise voice augmentation and

adaptive noise reduction, the hearing aid responds intelligently in real time to provide enhanced speech intelligibility and superior noise suppression. Thus, users no longer need to switch manually between listening programs, but get automatically adapted and personalised listening experience.

2.6 Manufacturer Innovations and Integrated Sensor Technology

Major hearing aid manufacturers, including Starkey, Phonak, Oticon, ReSound, Signia, and Widex, have successfully incorporated these AI-driven innovations into their devices. Some of the most advanced models now feature dedicated onboard DNN chips that have been extensively trained on massive datasets of real-world sounds to ensure immediate and precise acoustic classification.

In addition to advanced sound processing, researchers and developers have studied the effect of integrating sophisticated sensors into hearing aids to monitor a wider range of environmental factors and user metrics. These devices now utilize motion sensors and user activity data to automatically adjust microphone directionality and amplification based on how and where the user is moving. Beyond their audiological benefits, these integrated sensors are also capable of tracking critical health parameters and lifestyle factors, including monitoring social engagement, measuring physical activity, and providing life-saving fall detection (Fabry & Bhowmik, 2021).

2.7 Machine Learning Based Techniques Employed in Commercial Hearing Aids

The Widex Evoke program, known as "SoundSense Adapt", can learn users' preferences in various listening environments. Thus facilitating the hearing aid user's access to and control of personalised choices. Information regarding the user's preferences is stored in WidexCloud (Townend et al., 2018).

Through acoustic environmental classification (AEC), modern hearing aids classify the listening environment (e.g., quiet, speech, noise, and music) and automatically enable sound

management features suited to that environment. This computational process is derived from auditory scene analysis and imitate auditory system's ability to separate individual sounds in real world listening environments. AEC systems mainly combine two processing strategies: feature extraction and feature/pattern classification, followed by post-processing and environmental sound classification. The Hearing Reality Sound AEC of Starkey comprises eight automated sound classes: music, speech in quiet, speech in loud noise, speech in noise, machine, wind, noise, and quiet. By making specific adjustments in gain, compression, directionality, noise management and other parameters this system prioritises speech intelligibility in noise (Fabry & Bhowmik, 2021).

The voice priority processing (VPP) in Oticon Syncro hearing aids has sequential processing and optimises speech output and noise reduction.It was introduced in 2004 and has three signal processing approaches: Multiband Adaptive Directionality, TriState Noise Management and Voice-Aligned Compression(Flynn et al 2005). Multi-band adaptive directionality offers a polar pattern tailored to each frequency band; TriState Noise Management classifies noise scenarios into clearly defined listening modes; and voice-aligned compression delivers compression across a broader bandwidth utilising eight independent channels (Umashankar et al., 2021).

2.8 Deep Neural Network Based Techniques Employed in Commercial Hearing Devices

The integration of DNNs into commercial hearing aids marks a major milestone in clinical application. Oticon Inc. in 2021 made the first commercially available hearing aid with "on-the-chip" DNN technology. The Oticon MoreTM provides better speech in noise ability, recall/memory and enhances selective attention. It was developed based on the idea that by enhancing neural code, the brain's ability to understand sound would improve (Beck 2021). Starkey's Edge AI makes use of a DNN to distinguish "non-stationary source noise, and in turn better reduce noise levels for these noise types". It also "give more relief in

environments where a mix of stationary and non-stationary noise is present" (Betlehem et al. 2024).

The results from the recent study by Kolberg et al. (2025) indicated that DNN-based noise reduction in HA (Phonak Audéo Sphere Infinio 90) significantly improved speech understanding in noise for bimodal CI users.

Schulz, et al (2026) in their study investigated the impact of low-latency deep-learning noise reduction on both measured and anticipated speech intelligibility in noise among normal-hearing (NH) listeners, hearing-impaired (HI) individuals, and cochlear implant (CI) users. The noise reduction algorithm reduced speech clarity for normal-hearing listeners (SRT worsened by 0.9 dB) but improved it for hearing-impaired listeners and cochlear implant users (SRT improved by 0.8 dB and 5.7 dB, respectively).

Battery consumption, computational limitations, and the requirement for real-time processing are among the cons of applying AI in hearing aids. According to Khan et al., (2025), implementation of AI enhanced processing consumed 15 -150 times more power in comparison with that of traditional hearing aids. As noted by Jovanovic et al. (2022), there is an increasing demand for the development of AI models that are both effective and efficient for real-time applications, especially in assistive technology.

Added on to this, most of the AI models are trained on clean laboratory datasets. Thus generalizing to real world conditions are limited (Khan et al., 2025). Recent research by Schulz et al., 2026 reported that at there was no significant differences from the baseline using DNN when measured at -5dB SNR and medium reverberation condition. Furthermore, objective measurements frequently fail to accurately predict subjective enhancements in intelligibility for DNN-enhanced speech (Gelderblom et al., 2023). Recent study by Ashkanichenarlogh et al. (2025) revealed a mere 47.9% connection between intrusive Hearing-Aid Speech Quality Index(HASQI) and non-intrusive (pMOS) quality indicators.

Meeting these specifications poses significant challenges, as optimizing performance while maintaining device usability necessitates ongoing innovation in both hardware and software. Unless these challenges are addressed, AI-powered hearing aids will not be practical for everyday use or accessible to a wide range of consumers (Ismail, 2025).Thus, incorporating AI into hearing aids could enhance the user experience and provide personalized management options for hearing-impaired individuals.

CHAPTER 3

METHODS

3.1 Study Design

The present study utilized an observational research design.

3.2 Participants

A total of twenty-five individuals (Males = 12, Females = 13) with hearing impairment ranging from 50 -70 years of age(*M=59.32 years , SD =4.57)* participated in the study. The mean PTA was 49.35 dBHL (*SD=7.3*) and 50.19 dBHL (*SD=6.38)* for the right and left ears, respectively. The mean speech identification scores (SIS) were *92% (SD = 6.35)* and *92.64% (SD = 7.18)* for the right and left ears, respectively. All participants provided their informed consent before data collection. The study was conducted in accordance with the ethical guidelines specified in the *Ethical Principles for Bio-Behavioural Research Involving Human Subjects* (Venkatesan & Basavaraj, 2023). Ethical clearance was obtained from the Institutional Review Board of the All India Institute of Speech and Hearing (AIISH), Mysuru (Ref: SH/IRB/M.4/AUD.01/2025-26). Annexure I provides the ethical approval letter.

Inclusion Criteria

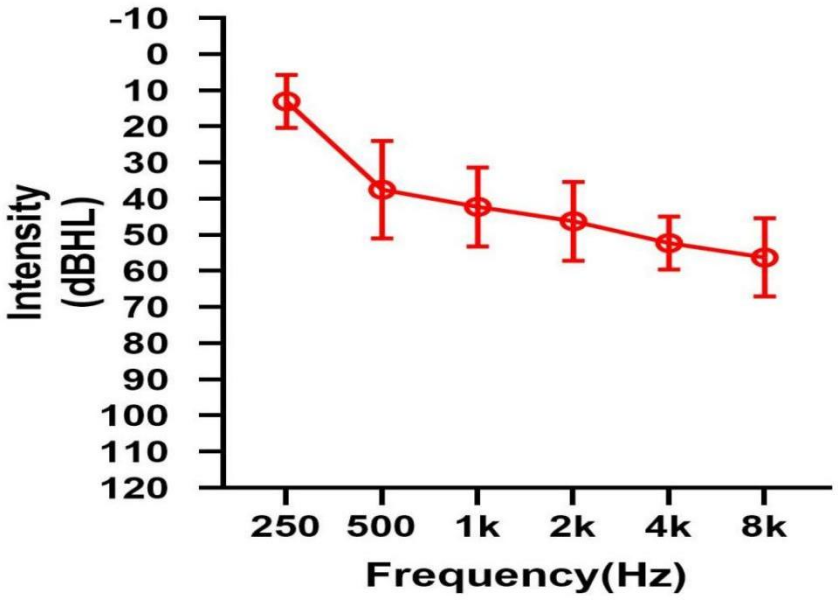
Participants who met the following criteria were included in the current study

1. Individuals with bilateral moderate to moderately severe sensorineural hearing loss. Figures 1 and 2 represent the mean and SD of the average hearing thresholds of the right and left ear, respectively.
2. Native speakers of the Kannada language.
3. Individuals with speech identification scores(SIS) that were proportionate to the degree of hearing loss.

4. Individuals who had "A" or "As" type tympanogram.
5. Individuals who passed in Mini Mental State Examination.

Figure 1

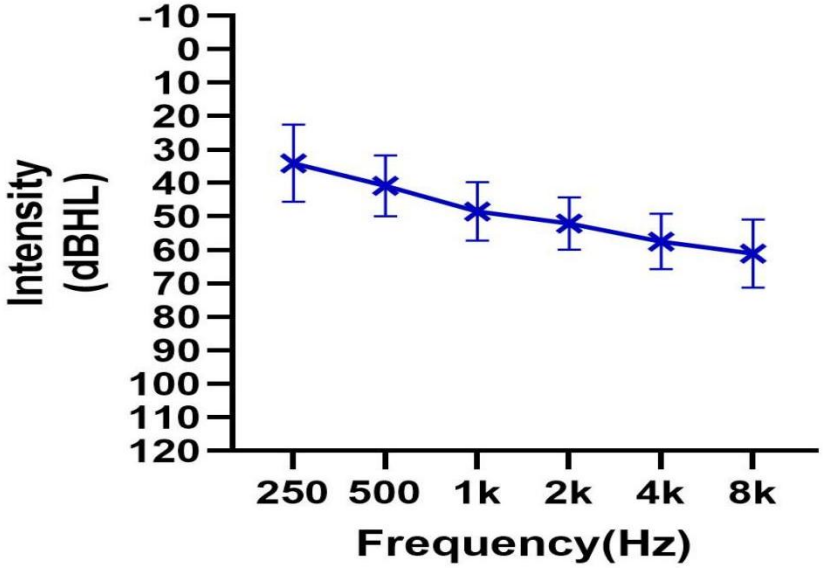
Line plot illustrating the average hearing thresholds for the right ear across different frequencies.



Note. Here, the x-axis shows frequency in Hertz (Hz), ranging from 250 to 8000 Hz. The y-axis shows hearing level in decibels hearing level (dB HL), ranging from −10 dB HL (better hearing) at the top to 120 dB HL (poorer hearing) at the bottom.

Figure 2

Line plot illustrating the average hearing thresholds for the left ear across different frequencies.



Note. Here, the x-axis shows frequency in Hertz (Hz), ranging from 250 to 8000 Hz. The y-axis shows hearing level in decibels hearing level (dB HL), ranging from −10 dB HL (better hearing) at the top to 120 dB HL (poorer hearing) at the bottom.

Exclusion Criteria

Participants were excluded if they had

1. Any recent history or complaint of middle ear pathology and neurological conditions.
2. Prior experience with hearing aid usage.

3.3 Test Environment and Equipment

All experiments were carried out in acoustically treated room as per ANSI S3.1 standards (1991) at the Centre for Hearing Sciences at AIISH, Mysuru.

The following instruments were used for the study:

- A personal computer equipped with a 13th Generation Intel® Core™ i5-13400F processor
- To ensure high-fidelity sound output, the MOTU M2 audio interface was used.
- The auditory stimuli were delivered using Genelec 1020A active loudspeakers.
- The MedRx REM system was used to verify real ear insertion gain (REIG)
- Spyder version 6 was used for stimulus presentation
- GRAS 45BB KEMAR was used for recording the output of HA for objective analysis.

Stimuli

The target stimuli consisted of Kannada sentences from the standardized "Kannada Sentence Identification Test" (Geetha et al. 2014) which was digitized at 44.1kHz. A four-talker Kannada babble (Geetha et al, 2018) served as the primary masker. In order to simulate ecologically valid conditions, traffic noise was also used as a masker. The traffic noise was recorded using sound professional ASMR binaural microphones from a heavy traffic location in Mysore. The recorded noise was then digitized at 44.1kHz sampling rate. The signals were calibrated using sound level meter (SLM) to ensure that the intensity levels of presentation.

Hearing Aid Fitting and Verification

The participants were fitted bilaterally with Phonak Audeo Sphere Infinio i90 RIC hearing aids, coupled with a power receiver and a power dome. The manufacturer was corresponded for optimal processing of deep neural network (DNN) based algorithm. The programming of the hearing aids was done according to the suggestions provided by the manufacturer. To isolate the contribution of the DNN, all other features (soundrelax, noiseblock, whistleblock) not linked to its operation was held constant across conditions in the two programs. For perceptual evaluation, the HA fitting was based on manufacturer's fitting formula (Adaptive Phonak Digital 3.0) as per the individual's hearing loss, using

Phonak target 11.1.1 fitting software. Incase of objective evaluation, the HA was programmed according to the standard N3 audiogram using the same fitting formula. The prescription gain was kept at 100%. Two programs were added to the device : a speech in noise program as Program 1(traditional-NR) and a Spheric Speech in Loud Noise as Program 2(.DNN-NR). The 'spheric speech clarity' strength was set to '7'. The gain was adjusted until the Real - Ear Insertion Gain(REIG) for the International Speech Test Signal (ISTS) presented at 65dBSL matched the target REIG prescribed by the proprietary fitting formula within a tolerance of ±5 dB in the real ear.

Test Conditions

Two hearing aid conditions was tested: traditional-NR and DNN-NR programs. Two spatial configurations was also assessed: co-located (target and masker at 0° azimuth) and spatially separated (target at 0° azimuth, masker at 60° azimuth).

Objective Analysis

Objective quantification of SNR improvement due to the DLA was carried out using the modified phase-inversion technique (Souza et al., 2006). The masker intensity was held constant at 70 dBSPL, and the intensity level of speech was 65dBSPL. The following variations of the input stimulus was generated: (1) speech plus masker (+Sp +N), (2) speech plus phase-inverted masker (+Sp –N), (3) phase-inverted speech plus masker (–Sp +N). These stimulus were provided in both co-located and spatially separated conditions. Inversions was implemented through Spyder software using python code by multiplying signal's amplitude value by -1. Stimuli was mixed at five input SNRs (–10, –5, 0, +5, +10 dB). Each trial began with chirp of 2sec followed by 0.05sec of masker and later speech–noise mixture was be played. A delay of 50msec were provided between the two channels in spatially separated condition.. The initial chirp stimulus was used for time alignment, which was verified manually.

The HA output was recorded for each SNR under the six input conditions under both spatial conditions: HA(+Sp +N), HA(+Sp –N), HA(–Sp +N). Speech and noise components was then extracted as:

Extracted speech = 0.5 × [HA(+Sp+N) + HA(+Sp–N)] ……(1)

Extracted noise = 0.5 × [HA(+Sp+N) + HA(–Sp+N)] ……. (2)

SNR = 20 × log10 (rms(Extracted speech) / rms(Extracted noise)) …… (3)

Speech Perception in Noise (SPiN) Procedure

The participants were instructed to listen carefully to speech sentences that were presented along with the background masker and to repeat whatever they have heard, by focusing on the target speech. In order to familiarize the participants with the task and environment, a practice session was conducted using a sentence list comprising 10 sentences in the spatially separated condition.

The main testing phase was then initiated and the speech-in-noise stimuli were presented under eight randomized conditions as per the following combinations:

- Two spatial configurations: Spatially collocated (0° azimuth) and separated (60° azimuth) configurations
- Two hearing aid program conditions: Traditional-NR and, DNN-NR
- Two noise types: Traffic noise and four talker speech babble

The background masker was presented continuously at a fixed level of 70 dB SPL, while the speech sentence intensity was dynamically varied to achieve the threshold SNR .

Prior to formal data collection, participants underwent a familiarization and practice session using a dedicated list of 10 practice sentences presented in the spatially separated condition. This ensured complete comprehension of the task requirements and adaptation to the laboratory environment.

The main experimental phase commenced with an initial ascending search phase for the first sentence. The first sentence was presented at an unfavorable speech level of 65 dB (-5 dBSNR). If the participant failed to meet the intelligibility criteria, the sentence was repeated with progressive 2 dB increments until a positive response was elicited.

Once the initial threshold region was identified, a 1-down, 1-up adaptive staircase tracking procedure was initiated for sentences 2 through 10. The speech level varied by a fixed step size of 2 dB based on the accuracy of the preceding trial:

- Positive Response: The SNR for the subsequent sentence was decreased by 2 dB.
- Negative Response: The SNR for the subsequent sentence was increased by 2 dB.

Each sentence contained exactly four target keywords. A lenient “2 out of 4” keyword criterion was strictly adopted to define a successful sentence-level trial. A response was scored as “correct” (1) if the participant accurately repeated two or more keywords (50% word accuracy), and "incorrect" (0) if they repeated one or fewer keywords.

This specific operational threshold was chosen to account for the reduced speech recognition capabilities and the prominent floor/ceiling effects observed in hearing-impaired cohorts during pilot trials when strict (e.g., 4/4) scoring criteria are enforced. By lowering the required keyword threshold, the paradigm mitigated participant fatigue and cognitive load, focusing instead on functional comprehension.

Mathematically, this procedure was modeled as a cumulative binomial distribution where the probability of a positive sentence response, P(Y), is defined as $P(k \geq 2)$ for $n = 4$. Under this binomial model, the 50% speech reception threshold (SRT_n) tracks the exact point where the underlying probability (p) of identifying an individual keyword is approximately 37.8%.

The final SRT_n for each condition was computed using the mid-run averaging method. Reversal points, defined as individual trials where the staircase shifted trajectory from ascending to descending or vice versa, were tracked. To eliminate any initial convergence or search-phase artifacts, the first reversal point was excluded from analysis. The SRT_n was calculated by averaging the midpoints of all subsequent peak and trough intensity levels.

CHAPTER 4

RESULTS

The aim of study was to analyse the benefit of DLA based noise reduction program in enhancing the speech perception in noise in hearing-impaired individuals. This was evaluated both perceptually as well as objectively.

4.1 Perceptual Analysis

The SRTn scores, which indicates the lowest value to perceive speech in the presence of background noise, were obtained across the following conditions: co-located (0^0 azimuth) and spatially separated (60^0 azimuth) using traffic noise and four talker Kannada babble. The hearing aids were programmed as having traditional speech in noise condition (program 1) and DNN based spheric speech in loud noise (program 2). These programs were switched across conditions. In all the conditions, noise level were set to a constant level of 70dBSPL and the initial speech level was set at 65dBSPL. The speech intensity was varied in 2dB steps adaptively as per the subject's response, to obtain the SRT value. The values obtained were statistically analyzed to determine main and interaction effects.

All statistical analyses were carried out using RStudio (2026.04.0). The 'glmmTMB' package was used for fitting generalized linear mixed models (GLMMs). For residual diagnostics and assumption checks, the diagnostics for hierarchical modeling (*DHARMa)* package was used. The type-II Wald Chi-Square test was used to investigate the main effects, which was performed using '*car'* package. The '*emmeans'* package was used to compute estimated marginal means (EMMs) and to conduct pairwise comparisons, which were adjusted using Bonferroni correction. The '*ggplot2'* package was used to plot probability density function.

SRT_n was analysed using a generalized linear mixed model (GLMM) with a Gaussian distribution and identity link function. The model included Spatial condition (co-located vs. separated), noise reduction program (Traditional-NR vs. DNN-NR), and noise type (babble vs. traffic) as fixed effects, along with all possible two-way and three-way interaction terms. Participant ID was included as a random intercept to account for repeated measures within individuals. The random intercept for ParticipantID accounted for meaningful between-subject variability ($\sigma^2 = 6.68$), while residual within-subject variance was 4.37.

The visual inspection of the plot revealed an approximately normal distribution and hence using a Gaussian GLMM with identity link SRTn ~ Spatial condition * Noise Reduction programs * Noise + (1 | PatientID), the dependent variable was analyzed. In addition, DHARMa diagnostics confirmed that the residuals were uniformly distributed (*D = 0.06, p = 0.45*) as ascertained by the Kolmogrov-Smirnov test. The dispersion test revealed no evidence of over- or under-dispersion (dispersion = 1.03, *p = 0.82*). Moreover, Levene's test confirmed homogeneity of variance across all the conditions, (*F(7,192) = 0.61*, *p = 0.75*).

Type II Wald chi-square tests were used to investigate the main effects of spatial condition, noise reduction programs conditions and types of noise, as well as their interaction on SRTn. The analysis revealed significant main effects of spatial condition ($\chi^2(1) = 122.88$, $p < .001$), noise reduction programs ($\chi^2(1) = 20.88$, $p < .001$) and noise type ($\chi^2(1) = 130.22$, $p < .001$). A significant interaction between spatial condition and noise type was also revealed by the analysis ($\chi^2(1) = 17.97$, $p < .001$). There was no significant two-way interaction between spatial conditions and noise reduction programs ($\chi^2(1) = 0.01$, $p = 0.91$), as well as noise reduction programs and noise type ($\chi^2(1) = 0.02$, $p = 0.89$). The three-way interaction, spatial condition x noise reduction programs x noise type ($\chi^2(1) = 0.74$, $p = 0.39$), was also not significant.

Estimated marginal means are presented in Table 1 and Table 2. Overall, higher SRT_n values, indicating poorer speech recognition performance and a greater SNR requirement for understanding speech, were observed in the co-located condition compared to the separated condition and were generally higher under babble noise than traffic noise. To further examine the significant condition × noise interaction, Bonferroni-corrected pairwise comparisons were conducted separately for babble and traffic noise conditions.

4.1.1 Babble Noise

Under babble noise, the traditional-NR program yielded significantly higher SRTn than the DNN-NR program in the co-located condition (difference = 1.69 dB, z = 2.86, p = 0.03), indicating poorer speech recognition performance with traditional-NR when speech and noise originated from the same spatial location. However, in the separated condition, the difference between traditional-NR and DNN-NR did not reach significance (difference = 1.09 dB, z = 1.85, p = 0.39), suggesting that spatial separation reduced the performance difference between the two noise reduction programs.

Within the traditional-NR program, the co-located condition resulted in significantly higher SRTn than the separated condition (difference = 4.83 dB, z = 8.17, p < .001), demonstrating a substantial benefit of spatial separation for speech recognition. Similarly, within the DNN-NR program, the co-located condition also produced significantly higher SRTn than the separated condition (difference = 4.23 dB, z = 7.16, p < .001) As depicted in Table 1, the spatial separation reduced the SRT_n for both traditional-NR(76.5 to 71.7 dB) and DNN-NR (74.9 to 70.6) across the both spatial listening conditions These findings indicate that spatial cues substantially improved speech recognition in babble noise irrespective of the noise reduction program, although the DNN-NR program provided additional benefit in the co-located listening condition. The figure 2 illustrates the distribution of speech reception

threshold(SRT_n) values across the co-located and separated spatial conditions for the two NR programs.

Table 1

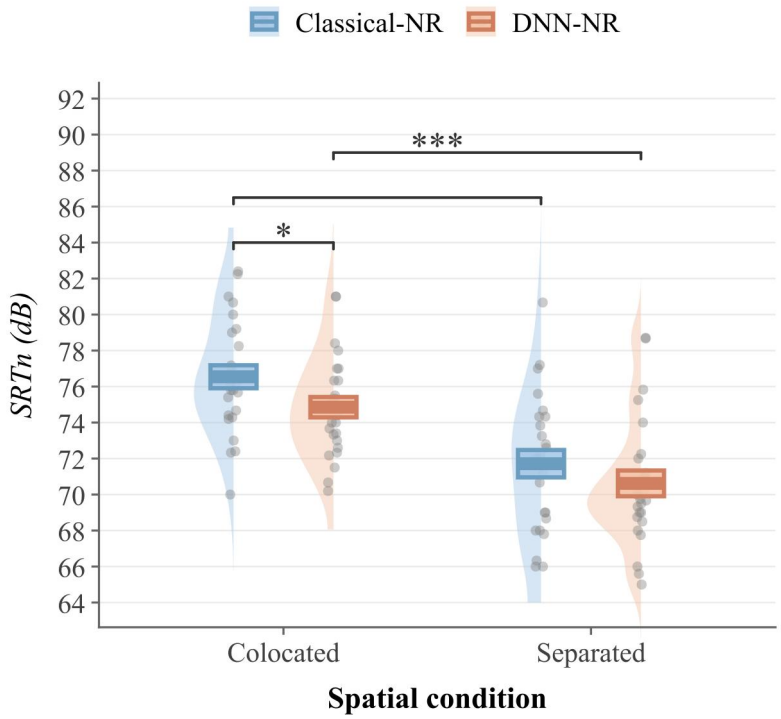
Estimated marginal means (EMMs) and 95% confidence intervals for SRT_n (dB) across all Condition × Program for babble noise

Spatial Condition	Program	EMM	SE	95% CI
Co-located	Traditional-NR	76.5	0.67	[75.2, 77.8]
	DNN-NR	74.9	0.67	[73.6, 76.2]
Separated	Traditional-NR	71.7	0.67	[70.4, 73.0]
	DNN-NR	70.6	0.67	[69.3, 71.9]

Note. EMMs are derived from the GLMM with PatientID as a random intercept. SE = standard error. CI = confidence interval.

Figure 3

Raincloud plot of SRT_n values for babble noise type across spatial and noise reduction conditions



Note. This raincloud plot represents the distribution of speech reception threshold (in dB) values for babble noise type across programs in co-located and separated spatial conditions. Here, SRTn values indicates the lowest value to perceive speech in the presence of background babble noise; Traditional- NR = Traditional Noise Reduction; DNN-NR = Deep Neural Network Noise Reduction. Significance comparisons: * p < .05. *** p <.001.

4.1.2 Traffic Noise

In contrast, under traffic noise, differences between the Traditional-NR and DNN-NR programs were not significant in either spatial condition. In the co-located condition, SRTn did not differ significantly between Traditional-NR and DNN-NR (difference = 1.10 dB, z =

1.86, p = .378). Likewise, in the Separated condition, the difference between the two programs also failed to reach significance (difference = 1.52 dB, z = 2.57, p = .061). These findings indicate that the DNN-NR program did not significantly outperform the Traditional-NR program under the traffic noise condition.

Despite the absence of significant noise reduction program-related differences, spatial separation continued to provide benefit under traffic noise. Within the Traditional-NR program, the co-located condition yielded significantly higher SRTn than the Separated condition (difference = 1.81 dB, z = 3.07, p = .013), indicating poorer performance when speech and noise were co-located. Similarly, within the DNN-NR program, SRTn remained significantly higher in the co-located condition compared to the Separated condition (difference = 2.24 dB, z = 3.78, p = .001). Table 2 depicts the estimated marginal means(EMMs) of SRT_n across the spatial conditions and NR programs with 95% confidence intervals. However, the magnitude of spatial benefit under traffic noise was smaller than that observed under babble noise, suggesting that spatial cues were more effective for overcoming the informational masking associated with babble noise than the predominantly energetic masking characteristics of traffic noise. The figure 3 illustrates the distribution of speech reception threshold(SRT_n) values across the co-located and separated spatial conditions for the two NR programs.

Table 2

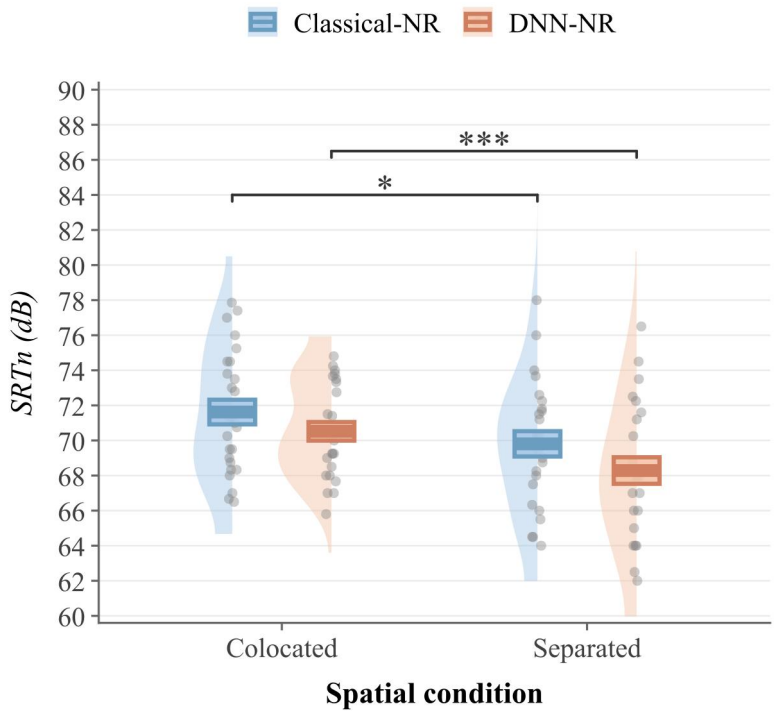
Estimated marginal means (EMMs) and 95% confidence intervals for SRT_n (dB) across all Condition × Program for traffic noise

Spatial Conditions	Program	EMM	SE	95% CI
Co-located	Traditional-NR	71.6	0.67	[70.3, 72.9]
	DNN-NR	70.5	0.67	[69.2, 71.8]
Separated	Traditional-NR	69.8	0.67	[68.5, 71.1]
	DNN-NR	68.3	0.67	[67.0, 69.6]

Note. EMMs are derived from the GLMM with PatientID as a random intercept. SE = standard error. CI = confidence interval.

Figure 4

Raincloud plot of SRT_n values for traffic noise type across spatial and noise reduction conditions

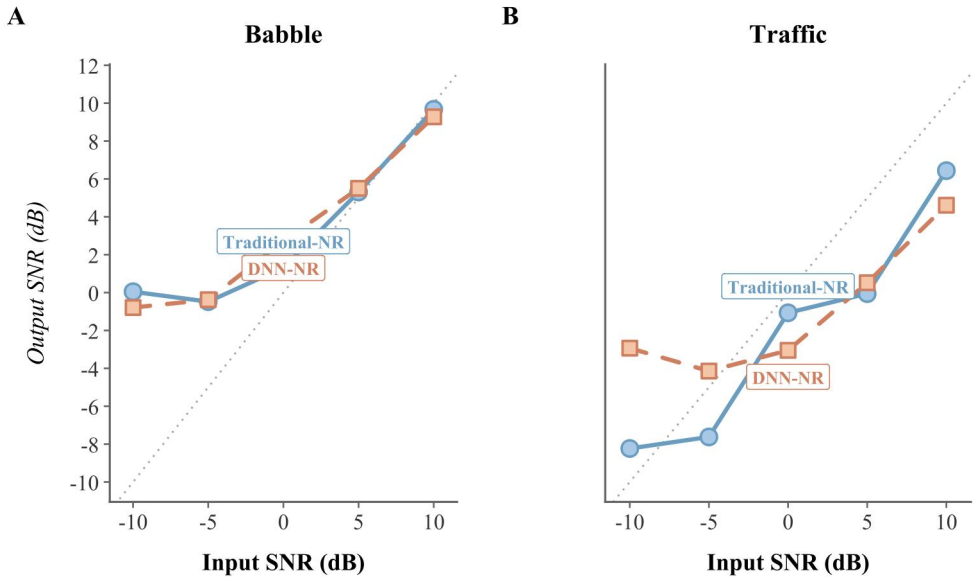


Note. This raincloud plot represents the distribution of speech reception threshold (in dB) values for traffic noise type across programs in co-located and separated spatial conditions. Here, SRTn values indicates the lowest value to perceive speech in the presence of background traffic noise; Traditional- NR = Traditional Noise Reduction; DNN-NR = Deep Neural Network Noise Reduction. Significance comparisons: * $p < .05$. *** $p < .001$.

4.2 Objective Analysis

Figure 5

Input SNR and output SNR across Traditional-NR and DNN-NR conditions in babble and traffic noise in co-located condition



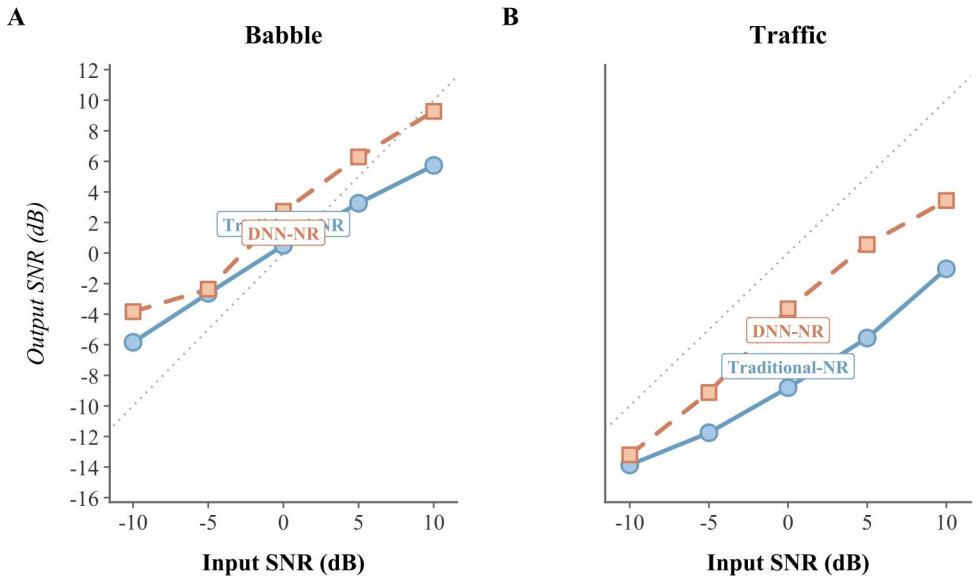
Note. The input and output signal to noise ratio (SNR) of the babble and traffic noise in spatially co-located condition is indicated in panel A and panel B, respectively. The dashed diagonal line (unity line) represents equal input and output SNR Traditional-NR = Traditional Noise Reduction; DNN-NR = Deep Neural Network Noise Reduction.

In the spatially collocated condition, both Traditional-NR and DNN-NR provide substantial SNR improvement under babble noise at negative input levels, elevating a -10 dB input to an output of approximately 0 dB. As the input SNR increases, the output for both algorithms converges to the unity line, indicating neither further improvement nor degradation, with no discernible performance difference between the two algorithms. Conversely, performance in traffic noise is significantly poorer. Both algorithms generally

fail to improve SNR in traffic noise at -5 dB input and above. Traditional-NR strictly tracks or falls below the unity line across all input levels.

Figure 6

Input SNR and output SNR across Traditional-NR and DNN-NR conditions in babble and traffic noise in separated condition



Note. The input and output signal to noise ratio (SNR) of the babble and traffic noise in spatially separated condition is indicated in panel A and panel B, respectively. The dashed diagonal line (unity line) represents equal input and output SNR Traditional-NR = Traditional Noise Reduction; DNN-NR = Deep Neural Network Noise Reduction.The dashed diagonal line (unity line) represents equal input and output SNR.

In the spatially separated condition, the DNN-NR algorithm demonstrates a clear advantage over Traditional-NR across both noise types. Under babble noise, DNN-NR consistently yields higher output SNRs across all tested input levels. While both algorithms successfully improve SNR at negative input levels (-10 dB and -5 dB), Traditional-NR

degrades the signal at positive input SNRs (+5 dB and +10 dB). In contrast, DNN-NR maintains an SNR improvement at +5 dB before falling slightly below unity at +10 dB. Similarly, in traffic noise, DNN-NR consistently provides higher output SNRs than Traditional-NR across the entire input range. Traditional-NR degrades the SNR at all tested input levels, whereas DNN-NR successfully improves SNR at negative input levels (-10 dB and -5 dB), although it still results in SNR degradation at 0 dB input and above.

CHAPTER 5

DISCUSSION

The primary aim of this study was to investigate the efficacy of a deep learning-based noise reduction algorithm (DNN-NR) in a commercially available hearing aid for improving the SNR in listeners with moderate to moderately severe sensorineural hearing loss. The study evaluated the perceptual advantage of this algorithm on SRTn across spatially separated and co-located conditions, utilizing both traffic noise and four-talker babble.

Overall, the findings demonstrated that the DNN-NR algorithm significantly improved SRTn in the spatially co-located condition when exposed to babble noise. However, this perceptual advantage over traditional-NR was not observed in the spatially separated condition. Spatial separation itself consistently improved SRTn across all noise types. Furthermore, neither the traditional NR nor the DNN-NR obscured the natural benefits of spatial hearing. Finally, listeners consistently demonstrated poorer SRTn in babble noise compared to traffic noise, irrespective of the algorithm used.

5.1 Algorithm Efficacy Across Noise Environments

The effectiveness of the DNN-based noise reduction varied significantly depending on the background noise characteristics. In co-located babble noise conditions, the DNN-NR program yielded significantly better SRTn than traditional-NR. Traditional rule-based algorithms, which rely on spectral subtraction and frequency-domain gain reduction, struggle in these environments because competing talkers share core spectral and temporal characteristics with the target speaker (Loizou, 2013). In contrast, DNN-based algorithms are trained on large multi-talker datasets, allowing them to learn complex temporal-spectral representations of acoustic features, such as fundamental frequency, spectral envelope, and

articulation patterns, which enables talker-specific source segregation (Anderson et al., 2021; Kinoshita et al., 2020). Additionally, the DNN architecture can process temporal context to track the target speaker over time by exploiting speech-specific amplitude modulation patterns (4–8 Hz) to differentiate the target from competing talkers (Dau et al., 1999; Drullman et al., 1994).

An important consideration when interpreting the algorithm's performance across these environments is the ecological validity of the noise stimuli and the geographic origin of the device. The hearing aids evaluated in this study are manufactured in Europe, and their DNN algorithms were likely trained on acoustic datasets representative of Western environments. The robust performance of the DNN-NR in local babble suggests that the fundamental acoustic features of speech utilized for source segregation are sufficiently universal to allow for effective cross-linguistic generalization. The algorithm successfully tracked and enhanced the target speech despite potential phonetic differences in the competing babble.

Conversely, in traffic noise, there were no significant differences between the DNN-NR and traditional NR programs. This lack of DNN advantage can be attributed to the region-specific nature of environmental noise. The acoustic characteristics of traffic noise in the Indian subcontinent are likely heterogeneous, high volume fluctuations, more dynamic, erratic, highly non stationary with more spectro-temporal variations and rich in high-intensity, broadband impulsive sounds (e.g., diverse vehicle horns, varied engine types) compared to the European settings. Because the DNN was likely not trained on the specific temporal and spectral complexities of this highly dynamic traffic environment, it failed to recognize these distinct acoustic signatures. Consequently, its advanced feature-extraction capabilities became ineffective, and the system performed no better than traditional spectral subtraction,

which simply treats the interference as steady-state energetic masking (Goehring et al., 2019). These results indicate that while DNN-based noise reduction generalizes well to universal speech patterns, its efficacy in environmental noise is highly dependent on how closely the real-world acoustic environment matches its training data.

5.2 Speech Reception Limitations in Multi-Talker Babble

Participants performed better in traffic noise than in four-talker babble, independent of the algorithm. This difference is caused by the large overlap in acoustics and time between the target speech and the masking speech in babble situations. Recognizing similar-sounding signals is a core constraint of human speech processing in competing-speech conditions, reflecting a perceptual challenge that exceeds what can be predicted by the signal-to-noise ratio alone (Brungart, 2001). Even with optimal noise removal, this perceptual limitation cannot be completely overcome (Bronkhorst, 2000).

In contrast to traffic noise, which concentrates acoustic energy in frequency regions distinct from target speech consonants and fricatives, babble contains competing talkers with spectral distributions nearly identical to those of the target. Thus any suppression of competing speech energy is likely to cause some level of attenuation of target speech energy in the same frequency range (Loizou, 2007). In addition, babble imposes substantial temporal difficulties. Modulation of speech amplitude is usually within the frequency range from 4–8 Hz and has been shown to play a key role for intelligibility (Drullman et al., 1994; Dau et al., 1999). In multi-talker babble, overlapping modulation patterns occur from competing talkers within this critical band and therefore segregation is much more difficult than separating speech from traffic noise.

Last but not least, babble noise is heavy on cognitive load. Specifically, it is the intelligible content of speech from competing talkers that captures attentional resources, increases listening effort; and draws on engaged linguistic systems in ways that non-speech traffic noise seems to do not (Brungart & Simpson, 2003; Hornsby, 2013). These combined acoustic, perceptual, and cognitive factors explain the consistently elevated thresholds observed in babble noise.

5.3 Impact of Spatial Separation on Speech Recognition

Spatial separation of the target speech and masker provided substantial benefits across both noise types and algorithm conditions. In babble noise, spatial separation significantly reduced SRTn for both traditional and DNN-based NR. Significant, though more modest, spatial benefits were also observed in traffic noise.

Notably, while the DNN-based algorithm demonstrated superiority in co-located babble, this advantage disappeared when the target and masker were spatially separated. In co-located conditions (0° azimuth), listeners cannot rely on interaural timing differences (ITD), interaural level differences (ILD), or head shadow cues (Bronkhorst, 2000; Middlebrooks & Green, 1991). Consequently, the DNN's feature-based segregation becomes the primary mechanism for improving speech recognition. However, when sounds are spatially separated, spatial cues alone provide substantial masking release, allowing effective segregation even with traditional noise reduction (Kidd et al., 1994; Culling et al., 2004). This suggests that while DNN processing is highly effective for feature-based segregation, it becomes less necessary when robust spatial information is naturally available to the listener.

5.4 Preservation of Binaural Cues and Spatial Hearing

A critical finding of clinical importance is that neither NR program obscured or distorted spatial hearing benefits. Aggressive signal processing can paradoxically impair spatial hearing by introducing artifacts that distort interaural relationships, even if it improves the monaural SNR (Peissig & Kollmeier, 1997). In this study, both programs maintained the binaural coherence essential for spatial hearing. The algorithms appear to utilize bilateral processing that synchronizes or minimizes latency mismatches between ears, preserving the ITD and ILD cues critical for localization and spatial release from masking (Akeroyd et al., 2007).

Furthermore, the robust spatial benefits observed indicate that the algorithms preserved the spectral relationships encoded in head-related transfer functions (HRTFs), which are necessary for identifying source location (Middlebrooks & Green, 1991; Blauert, 1997). Consistent with recent literature, these findings confirm that neural networks can enhance speech perception in noisy environments without compromising binaural cues (Tokala et al., 2023), underscoring the preservation of spatial hearing as a vital component of real-world hearing aid efficacy (Bentler et al., 2008).

CHAPTER 6

SUMMARY AND CONCLUSIONS

This study investigated the efficacy of an AI-based deep learning noise reduction algorithm (DNN-NR) implemented in a commercially available hearing aid (Phonak Audéo Sphere Infinio i90) in enhancing speech perception in noise among individuals with sensorineural hearing loss. Twenty-five adults aged 50–70 years with bilateral moderate-to-moderately severe sensorineural hearing loss, all native Kannada speakers, participated in the study. Both objective and perceptual evaluations were carried out across co-located and spatially separated conditions using two ecologically valid noise types - four talker Kannada babble and locally recorded traffic noise. The performance of the DNN-NR program was compared against a traditional noise reduction (Traditional-NR) program.

Objectively, SNR improvement was quantified using the modified phase-inversion technique, while perceptually, speech reception thresholds in noise (SRTn) were measured using an adaptive staircase procedure with standardized Kannada sentences. The hearing aids were fitted bilaterally and programmed with two programs with all other processing features held constant across conditions.

The DNN-NR algorithm enhanced the signal-to-noise ratio(SNR) in the negative SNR till when the speech and noise levels were equal and reduced further with increasing SNR. This reveals that in case of noise dominated signal, there was better performance of the algorithm. The results revealed that DNN-NR algorithm provided a statistically significant perceptual benefit over Traditional-NR only under babble noise in the co-located condition, with a mean SRT_n difference of approximately 1.69 dB. In the spatially separated babble condition, this advantage was no longer significant, suggesting that spatial cues independently facilitated adequate source segregation. Under traffic noise, no significant difference was observed between the two programs in either spatial condition. In addition,

76% of participants subjectively preferred the DNN-NR program, when they provided with dichotomous preference scale.

6.1 Clinical Implications of the Study

• **Hearing aid fitting and device selection:** The DNN-NR algorithms benefit patients mainly in co-located multi-talker situations. For environments with spatial separation or stationary noise, traditional algorithms may be equally effective.

• **Programming strategies and feature prioritization:** DNN processing should be isolated from other noise reduction features to avoid interference. Real-ear verification is essential to confirm gain targets are maintained when DNN is active.

• **Spatial Listening Strategies, Counseling and Expectation Management:** Patients should be given realistic counseling about when DNN-NR algorithms provide meaningful benefit and when they may not. Spatial separation of speech and noise provided 4–5 dB benefit in babble and ~2 dB in traffic, so patients should be encouraged to optimize seating and room layout alongside technology use.

• **Population-specific considerations and linguistic diversity:** The present study is one of the first to evaluate DNN hearing aids using authentic Indian speech materials. Algorithm validation across diverse languages and populations is essential before generalizing manufacturer claims.

6.2 Limitations and Future Directions

• **Sample size and participant characteristics:** In the current study, 25 participants with moderate-to-moderately severe hearing loss were included. Larger and more diverse samples are needed to explore effects across age, gender, and hearing loss type.

• **Device-specific findings and algorithm generalization:** The results obtained in this study, apply only to the Phonak Audéo Sphere Infinio i90 and its proprietary DNN algorithm. For

determining broader applicability, comparative studies across multiple brands and DNN configurations are needed.

• **Ecological validity and real-world listening environments:** Since laboratory conditions do not fully replicate the dynamic acoustics of real-world environments, future field studies and longitudinal designs using ecological momentary assessment would improve real-world relevance.

• **Noise types and masking mechanisms:** Only two maskers (babble and traffic noise) were tested, limiting conclusions about other masking conditions. Future research should include a wider range of noise types and talker counts.

• **Listener factors :** Working memory and attention were not assessed, though these modulate speech perception. Future research should incorporate these factors to understand individual differences.

REFERENCES

Aghasoleimani, M., Jalilvand, H., Mahdavi, M., & Ahmadi, R. (2020). Auditory Recognition of Digit-in-Noise under Unaided and Aided Conditions in Moderate and Severe Sensorineural Hearing Loss. *Journal of Audiology & Otology, 25*, 72 - 79. https://doi.org/10.7874/jao.2020.00094

Akeroyd, M. A., Gatehouse, S., & Blaschke, J. (2007). The detection of differences in the cues to distance by elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *121*(2), 1077–1089. https://doi.org/10.1121/1.2404927

Alexander, J. (2021). Hearing Aid Technology to Improve Speech Intelligibility in Noise. *Seminars in Hearing, 42*, 175 - 185. https://doi.org/10.1055/s-0041-1735174

AlSamhori, J. F., AlSamhori, A. R. F., Amourah, R. M., AlQadi, Y., Koro, Z. W., Haddad, T. R. A., AlSamhori, A. F., Kakish, D., Kawwa, M. J., Zuriekat, M., & Nashwan, A. J. (2024). Artificial intelligence for hearing loss prevention, diagnosis, and management. *Journal of Medicine Surgery and Public Health*, *3*, 100133. https://doi.org/10.1016/j.glmedi.2024.100133

AMERICAN NATIONAL STANDARD, Accredited Standards Committee S3, Bioacoustics, & Acoustical Society of America. (1999). *Maximum permissible ambient noise levels for audiometric test rooms*. Acoustical Society of America. (Original work published 1991)

Andersen AH, Santurette S, Pedersen MS, Alickovic E, Fiedler L, Jensen J, Behrens T. Creating Clarity in Noisy Environments by Using Deep Learning in Hearing Aids. Semin Hear. 2021 Aug;42(3):260-281. doi: 10.1055/s-0041-1735134. Epub 2021 Sep 24. PMID: 34594089; PMCID: PMC8463126.

Ashkanichenarlogh, V., Folkeard, P., Scollie, S., Kühnel, V., & Parsa, V. (2025). Objective Evaluation of a Deep Learning-Based Noise Reduction Algorithm for Hearing Aids Under

Diverse Fitting and Listening Conditions. *Trends in Hearing, 29*. https://doi.org/10.1177/23312165251396644

Au, A., Blakeley, J., Dowell, R., & Rance, G. (2018). Wireless binaural hearing aid technology for telephone use and listening in wind noise. *International Journal of Audiology, 58*, 193 - 199. https://doi.org/10.1080/14992027.2018.1538573

Baboukani, P., Graversen, C., Alickovic, E., & Østergaard, J. (2022). Speech to noise ratio improvement induces nonlinear parietal phase synchrony in hearing aid users. *Frontiers in Neuroscience, 16*. https://doi.org/10.3389/fnins.2022.932959

Beck, D. L. (2021). Hearing, listening and deep neural networks in hearing aids. In *Journal of Otolaryngology-ENT Research* (Vols. 1–2021, pp. 5–8) [Research Article]. https://doi.org/10.15406/joentr.2021.13.00481

Bentler, R. A. (2005). Effectiveness of directional microphones and noise reduction schemes in hearing aids: A Systematic Review of the evidence. *Journal of the American Academy of Audiology*, *16*(7), 473–484. https://doi.org/10.3766/jaaa.16.7.7

Bentler, R., Wu, Y., Kettel, J., & Hurtig, R. (2008). Digital noise reduction: Outcomes from laboratory and field studies. *International Journal of Audiology*, *47*(8), 447–460. https://doi.org/10.1080/14992020802033091

Betlehem, T., Mishra, P., Xu, J., Marquardt, D., & McKinney, M. (2024). Leveraging DNN in Edge AI 1. In Starkey. https://cdn.mediavalet.com/usil/starkeyhearingtech/DPIrh4nT5UWv37iI1gwcqA/5FtECxMyD0qjVCcNhAH4Sw/Original/DNN%20Assisted%20MNR%20Whitepaper.pdf

Bhat, G.S., Shankar, N., Reddy, C.K.A. & Panahi, I.M.S. (2019). A Real-Time Convolutional Neural Network-Based Speech Enhancement for Hearing Impaired Listeners Using a Smartphone. *IEEE Access*, **7**, 78421–78433.

Bishop, C. M. (2006). *Pattern recognition and machine learning: Solutions to the Exercises, Web edition*.

Blauert, J. (1997). *Spatial hearing* (2nd, enlarged edition ed.). The MIT Press.

Bronkhorst, A. W. (2000). The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acustica*, *86*(1), 117–128. https://jglobal.jst.go.jp/en/detail?JGLOBAL_ID=200902119753624260

Brungart, D. S., & Simpson, B. D. (2003). Within-ear and across-ear interference in a dichotic cocktail party listening task: Effects of masker uncertainty. *The Journal of the Acoustical Society of America*, *115*(1), 301–310. https://doi.org/10.1121/1.1628683

Chong, F., & Jenstad, L. (2018). A critical review of hearing-aid single-microphone noise-reduction studies in adults and children. *Disability and Rehabilitation: Assistive Technology, 13*, 600 - 608. https://doi.org/10.1080/17483107.2017.1392619

Chung, K. (2004). Challenges and Recent Developments in Hearing Aids: Part I. Speech Understanding in Noise, Microphone Technologies and Noise Reduction Algorithms. *Trends in Amplification, 8*, 124 - 83.

Ciorba, A., Zattara, S., Loroni, G., Prosser, S., & Lo, R. (2014). Quantitative enhancement of speech in noise through a wireless equipped hearing aid. *Acta Otorhinolaryngologica Italica, 34*, 50 - 53.

Copeland, J. (2015). *Artificial intelligence: A Philosophical Introduction*. John Wiley & Sons.

Culling, J. F., Hawley, M. L., & Litovsky, R. Y. (2004). The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. *The Journal of the Acoustical Society of America*, *116*(2), 1057–1065. https://doi.org/10.1121/1.1772396

Dau, T., Verhey, J., Carl von Ossietzky Universita ¨t Oldenburg, AG Medizinische Physik, Graduiertenkolleg Psychoakustik, Kohlrausch, A., IPO Center for Research on User-System

Interaction, & Philips Research Laboratories Eindhoven. (1999). Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers [Journal-article]. *J. Acoust. Soc. Am.*, *106*(5), 2752–2752.

Dell'Antônia, S. F., Ikino, C. M., & Carreirão Filho, W. (2013). Degree of satisfaction of patients fitted with hearing aids at a high complexity service. *Brazilian journal of otorhinolaryngology*, *79*(5), 555–563. https://doi.org/10.5935/1808-8694.20130100

Deniz, B., Gülmez, Z. D., Kara, H., & Kara, E. (2024). Effect of digital noise reduction in hearing aids on speech intelligibility in both quiet and noisy environments. *Noise and Health*, *26*(121), 220–225. https://doi.org/10.4103/nah.nah_67_23

Diehl, P., Zilly, H., Sattler, F., Singer, Y., Kepp, K., Berry, M., Hasemann, H., Zippel, M., Kaya, M., Meyer-Rachner, P., Pudszuhn, A., Hofmann, V., Vormann, M., & Sprengel, E. (2023). Deep learning-based denoising streamed from mobile phones improves speech-in-noise understanding for hearing aid users. *ArXiv, abs/2308.11456*. https://doi.org/10.48550/arxiv.2308.11456

Diehl, P.U, Singer, Y., Zilly, H., Schonfeld, U., Meyer-Rachner, P., Berry, M., Sprekeler, H., Sprengel, E., Pudszuhn, A., & Hofmann, V. (2022). Restoring speech intelligibility for hearing aid users with deep learning. *Scientific Reports, 13*. https://doi.org/10.1038/s41598-023-29871-8

Dong, R., Liu, P., Tian, X., Wang, Y., Chen, Y., Zhang, J., Yang, L., Zhao, S., Guan, J., & Wang, S. (2024). Influences of noise reduction on speech intelligibility, listening effort, and sound quality among adults with severe to profound hearing loss. *Frontiers in Neuroscience, 18*. https://doi.org/10.3389/fnins.2024.1407775

Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of reducing slow temporal modulations on speech reception. *The Journal of the Acoustical Society of America*, *95*(5), 2670–2680. https://doi.org/10.1121/1.409836

Drullman, R., Festen, J. M., & Plomp, R. (1994a). Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America*, *95*(2), 1053–1064. https://doi.org/10.1121/1.408467

Essaid, B., Kheddar, H., Batel, N., & Chowdhury, M. (2025). Deep Learning-Based Coding Strategy for Improved Cochlear Implant Speech Perception in Noisy Environments. *IEEE Access, 13*, 35707-35732. https://doi.org/10.1109/access.2025.3542953

Fabry, D., & Bhowmik, A. (2021). Improving Speech Understanding and Monitoring Health with Hearing Aids Using Artificial Intelligence and Embedded Sensors. *Seminars in Hearing, 42*, 295 - 308. https://doi.org/10.1055/s-0041-1735136

Fitzgerald, M., Athreya, V., Srour, M., Rejimon, J., Venkitakrishnan, S., Bhowmik, A., Jackler, R., Steenerson, K., & Fabry, D. (2025). Effectiveness of deep neural networks in hearing aids for improving signal-to-noise ratio, speech recognition, and listener preference in background noise. *Frontiers in Audiology and Otology*. https://doi.org/10.3389/fauot.2025.1677482

Flynn, M. C., & Lunner, T. (2005). Clinical verification of a hearing aid with Artificial Intelligence. *The Hearing Journal*, *58*(2), 34–38. https://doi.org/10.1097/01.hj.0000286116.92711.77

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state.” *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Geetha, C., Kumar, K.S., Manjula, P., & Pavan, M. (2014). Development and standardisation of the sentence identification test in the Kannada language. *Journal of hearing science, 4*, 18-26. https://doi.org/10.17430/890267

Gelderblom, F. B., Tronstad, T. V., Svendsen, T., & Myrvoll, T. A. (2023). On the Predictive Power of Objective Intelligibility Metrics for the Subjective Performance of Deep Complex Convolutional Recurrent Speech Enhancement Networks. *IEEE/ACM Transactions on Audio Speech and Language Processing*, *32*, 215–226. https://doi.org/10.1109/taslp.2023.3329378

Goehring, T., Keshavarzi, M., Carlyon, R. P., & Moore, B. C. J. (2019). Using recurrent neural networks to improve the perception of speech in non-stationary noise by people with cochlear implants. *The Journal of the Acoustical Society of America*, *146*(1), 705–718. https://doi.org/10.1121/1.5119226

Healy, E., Johnson, E., Pandey, A., & Wang, D. (2023). Progress made in the efficacy and viability of deep-learning-based noise reduction.. *The Journal of the Acoustical Society of America, 153 5*, 2751. https://doi.org/10.1121/10.0019341

Hornsby, B. W. Y. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear And Hearing*, *34*(5), 523–534. https://doi.org/10.1097/aud.0b013e31828003d8

Ibrahim, I., Parsa, V., Macpherson, E., & Cheesman, M. (2012). Evaluation of Speech Intelligibility and Sound Localization Abilities with Hearing Aids Using Binaural Wireless Technology. *Audiology Research, 3*. https://doi.org/10.4081/audiores.2013.e1

Ismail, Q. M. (2025). Artificial Intelligence in Hearing Aids Technology: Meta-Analysis & How to Share the Patient His Needs Depending upon His Life Style. *Open Journal of Applied Sciences*, *15*(07), 2029–2050. https://doi.org/10.4236/ojapps.2025.157134

Jovanovic, M., Mitrov, G., Zdravevski, E., Lameski, P., Colantonio, S., Kampel, M., Tellioglu, H., & Florez-Revuelta, F. (2022). Ambient Assisted Living: Scoping review of artificial intelligence models, domains, technology, and concerns. *Journal of Medical Internet Research*, *24*(11), e36553. https://doi.org/10.2196/36553

Jürgens, T., Ihly, P., Tchorz, J., Nishiyama, T., Tanaka, C., Suzuki, D., Shinden, S., Kitama, T., Ogawa, K., Zaar, J., Laugesen, S., Jones, G., Vatti, M., & Santurette, S. (2025). Closedness of Acoustic Coupling and Audiological Measures Are Associated with Individual Speech-in-Noise Benefit From Noise Reduction in Hearing Aids. *Trends in Hearing, 29*. https://doi.org/10.1177/23312165251325983

Kang, K., & Gong, Q. (2024). Noise Reduction Algorithm for Hearing Aids Based on Smoothed Wiener Filtering. *2024 6th International Academic Exchange Conference on Science and Technology Innovation (IAECST)*, 1198-1203. https://doi.org/10.1109/iaecst64597.2024.11118250

Khan, H., Asif, S., & Nasir, H. (2025). Advances in Intelligent Hearing Aids: Deep Learning Approaches to Selective Noise Cancellation. *ArXiv, abs/2507.07043*. https://doi.org/10.48550/arxiv.2507.07043

Kidd, G., Mason, C. R., Deliwala, P. S., Woods, W. S., & Colburn, H. S. (1994). Reducing informational masking by sound segregation. *The Journal of the Acoustical Society of America*, *95*(6), 3475–3480. https://doi.org/10.1121/1.410023

Kim, S. (2020). Wearable Hearing Device Spectral Enhancement Driven by Non-Negative Sparse Coding-Based Residual Noise Reduction. *Sensors (Basel, Switzerland), 20*. https://doi.org/10.3390/s20205751

Kinoshita, K., Ochiai, T., Delcroix, M., & Nakatani, T. (2020). *Improving Noise Robust Automatic Speech Recognition with Single-Channel Time-Domain Enhancement Network* (pp. 7009–7013). https://doi.org/10.1109/icassp40776.2020.9053266

Klein, S., Rossol, L., Venema, F., Schönewald, S., Karrenbauer, J., & Blume, H. (2025). Noise Reduction in Hearing-Aid Processors: Traditional Methods vs. Neural Networks. *2025 IEEE 36th International Conference on Application-specific Systems, Architectures and Processors (ASAP)*, 172-173. https://doi.org/10.1109/asap65064.2025.00037

Kolberg, C., Holbert, S. O., Bogle, J. M., & Saoji, A. A. (2025). DNN-Based noise reduction significantly improves bimodal benefit in background noise for cochlear implant users. *Journal of Clinical Medicine*, *14*(15), 5302. https://doi.org/10.3390/jcm14155302

Kumar, S., Guruvayurappan, A., Pitchaimuthu, A., & Nayak, S. (2023). Efficacy and Effectiveness of Wireless Binaural Beamforming Technology of Hearing Aids in Improving Speech

Perception in Noise: A Systematic Review. *Ear and Hearing, 44*, 1289 - 1300. https://doi.org/10.1097/aud.0000000000001374

Lazzerini, F., Baldassari, L., Angileri, A., Bruschini, L., Berrettini, S., & Forli, F. (2023). Does the Remote Microphone Still Outperform the Pre-Processing Algorithms? A Group Study in Adult Nucleus Recipients. *Journal of Otorhinolaryngology, Hearing and Balance Medicine*. https://doi.org/10.3390/ohbm4020009

Levitt, H. (2007). A Historical perspective on digital hearing aids: How digital technology has changed modern hearing aids. *Trends in Amplification*, *11*(1), 7–24. https://doi.org/10.1177/1084713806298000

Loizou, P. C. (2013). *SPEECH ENHANCEMENT theory and practice* (Second Edition). CRC Press.

Lovato, A., Monzani, D., Kambo, Y., Franz, L., Frosolini, A., & De Filippis, C. (2024). The Efficacy of Wireless Auditory Training in Unilateral Hearing Loss Rehabilitation. *Audiology Research, 14*, 554 - 561. https://doi.org/10.3390/audiolres14040046

M, N., Parmeshwara, K. S., & Geetha, C. (2018). Effect of number of talkers and language of babble on acceptable noise level in Kannada listeners. *Hearing Balance and Communication*, *16*(4), 241–247. https://doi.org/10.1080/21695717.2018.1542858

Magnusson, L., Claesson, A., Persson, M., & Tengstrand, T. (2013). Speech recognition in noise using bilateral open-fit hearing aids: The limited benefit of directional microphones and noise reduction. *International Journal of Audiology, 52*, 29 - 36. https://doi.org/10.3109/14992027.2012.707335

McCreery, R., Venediktov, R., Coleman, J., & Leech, H. (2012). An evidence-based systematic review of directional microphones and digital noise reduction hearing aids in school-age children with hearing loss.. *American journal of audiology, 21 2*, 295-312. https://doi.org/10.1044/1059-0889(2012/12-0014)

Mehrkian, S., Bayat, Z., Javanbakht, M., Emamdjomeh, H., & Bakhshi, E. (2019). Effect of wireless remote microphone application on speech discrimination in noise in children with cochlear implants. *International journal of pediatric otorhinolaryngology, 125*, 192-195. https://doi.org/10.1016/j.ijporl.2019.07.007

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, *42*(1), 135–159. https://doi.org/10.1146/annurev.ps.42.020191.001031

Miller, C., Bentler, R., Wu, Y., Lewis, J., & Tremblay, K. (2017). Output signal-to-noise ratio and speech perception in noise: effects of algorithm. *International Journal of Audiology, 56*, 568 - 579. https://doi.org/10.1080/14992027.2017.1305128

Moore, B. C. J. (2007). *Cochlear hearing loss: Physiological, Psychological and Technical Issues*. John Wiley & Sons.

Mustafa, M., & Krishnamurthy, A. (2025). A comprehensive review on real-world challenges faced by hearing aid users and innovative solutions for background noise—from lab to life. *The Egyptian Journal of Otolaryngology*, *41*(1). https://doi.org/10.1186/s43163-025-00814-6

Park, G., Cho, W., Kim, K.-S., & Lee, S. (2020). Speech Enhancement for Hearing Aids with Deep Learning on Environmental Noises. *Applied Sciences*, *10*(17), 6077. https://doi.org/10.3390/app10176077dw

Peissig, J., & Kollmeier, B. (1997). Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners. *The Journal of the Acoustical Society of America*, *101*(3), 1660–1670. https://doi.org/10.1121/1.418150

Pujar, R. (2019). Wiener Filter Based Noise Reduction Algorithm with Perceptual Post Filtering for Hearing Aids. *International Journal of Image, Graphics and Signal Processing*. https://doi.org/10.5815/ijigsp.2019.07.06

Rodemerk, K., & Galster, J. (2015). The Benefit of Remote Microphones Using Four Wireless Protocols. *Journal of the American Academy of Audiology, 26 8*, 724-31. https://doi.org/10.3766/jaaa.15008

Rohith, K., & Bhandarkar, R. (2021). A Comparative Analysis of Statistical Model and Spectral Subtractive Speech Enhancement Algorithms. *Advances in VLSI, Signal Processing, Power Electronics, IoT, Communication and Embedded Systems*. https://doi.org/10.1007/978-981-16-0443-0_32

Saoji, A., Sheikh, B., Bertsch, N., Goulson, K., Graham, M., McDonald, E., Bross, A., Vaisberg, J., Kühnel, V., Voss, S., Qian, J., Hogan, C., & Dejong, M. (2024). How Does Deep Neural Network-Based Noise Reduction in Hearing Aids Impact Cochlear Implant Candidacy?. *Audiology Research, 14*, 1114 - 1125. https://doi.org/10.3390/audiolres14060092

Schröter, H., Rosenkranz, T., Escalante, A., Aubreville, M., & Maier, A. (2020). CLCNET: Deep Learning-Based Noise Reduction for Hearing aids using Complex Linear Coding. *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6949-6953. https://doi.org/10.1109/icassp40776.2020.9053563

Schulz, K. M. B., Wittmann, H., Hackenberg, L., Jürgens, T., & Tchorz, J. (2026). Effects of low-latency deep-learning-based noise reduction on speech intelligibility for normal hearing, hearing-impaired listeners and cochlear implant users. *Frontiers in Audiology and Otology*, *3*. https://doi.org/10.3389/fauot.2025.1746635

Schum, D. (2000). COMBINING ADVANCED TECHNOLOGY NOISE CONTROL SOLUTIONS. *SEMINARS IN HEARING, Volume 21*, 0169 - 0178. https://doi.org/10.1055/s-2000-7308

Sheikhzadeh, H., Sameti, H., Deng, L., & Brennan, R. (1994). Comparative performance of spectral subtraction and HMM-based speech enhancement strategies with application to hearing and design. *Proceedings of ICASSP '94. IEEE International Conference on Acoustics, Speech and Signal Processing, i*, I/13-I/16 vol.1. https://doi.org/10.1109/icassp.1994.389367

Snapp, H., Holt, F., Liu, X., & Rajguru, S. (2016). Comparison of Speech-in-Noise and Localization Benefits in Unilateral Hearing Loss Subjects Using Contralateral Routing of Signal Hearing Aids or Bone-Anchored Implants. *Otology & Neurotology*. https://doi.org/10.1097/mao.0000000000001269

Souza, P. E., Jenstad, L. M., & Boike, K. T. (2006). Measuring the acoustic effects of compression amplification on speech in noise. *The Journal of the Acoustical Society of America*, *119*(1), 41–44. https://doi.org/10.1121/1.2108861

Tan, C., Tu, L., Wang, Y., Jin, D., Wang, Y., & Shi, W. (2024). Improving Speech Recognition during Phone Calls in Noisy Environment through the Use of Wireless Audio Streaming in Hearing Aids. *OALib*. https://doi.org/10.4236/oalib.1111343

Thibodeau, L. (2014). Comparison of speech recognition with adaptive digital and FM remote microphone hearing assistance technology by listeners who use hearing aids. *American journal of audiology, 23 2*, 201-10. https://doi.org/10.1044/2014_aja-13-0065

Thibodeau, L. (2019). Benefits in Speech Recognition in Noise with Remote Wireless Microphones in Group Settings. *Journal of the American Academy of Audiology, 31*, 404 - 411. https://doi.org/10.3766/jaaa.19060

Thibodeau, L., & Johnson, C. (2014). Wireless Technology to Improve Communication in Noise. *Seminars in Hearing, 35*, 157 - 157. https://doi.org/10.1055/s-0034-1383500

Tokala, V., Grinstein, E., Brookes, M., Doclo, S., Jensen, J., & Naylor, P. A. (2023). *Binaural Speech Enhancement Using Complex Convolutional Recurrent Networks* (pp. 1130–1134). https://doi.org/10.1109/ieeeconf59524.2023.10476738

Townend O, Nielsen JB, Ramsgaard J.(2018). Real-life applications of machine learning in hearing aids. Hearing Review, 25(4):34-37.

Trifosa, A., & Wulandari, M. (2025). Comparative Study of Noise Reduction Techniques in Audio Data Using PESQ and STOI. *2025 International Conference on Applied Artificial*

Intelligence, Data Engineering and Sciences (ICAIDES), 1-6. https://doi.org/10.1109/icaides67265.2025.11404029

Umashankar, A., Mn, A., & C, P. (2021). APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN HEARING AIDS AND AUDITORY IMPLANTS: a SHORT REVIEW. *Journal of Hearing Science*, *11*(3), 20–23. https://doi.org/10.17430/jhs.2021.11.3.2

Voll, L. (2000). APPLICATION OF TECHNOLOGY TO IMPROVE SIGNAL-TO-NOISE RATIO. *Seminars in Hearing, 21*, 0157-0168. https://doi.org/10.1055/s-2000-7309

Vroegop, J., Homans, N., Goedegebure, A., & Van Der Schroeff, M. (2018). A directional remote-microphone for bimodal cochlear implant recipients. *International Journal of Audiology, 57*, 858 - 863. https://doi.org/10.1080/14992027.2018.1508896

Wang, Z., Dong, R., Fu, X., Wang, Y., & Wang, S. (2021). Effects of Wireless Remote Microphone on Speech Recognition in Noise for Hearing Aid Users in China. *Frontiers in Neuroscience, 15*. https://doi.org/10.3389/fnins.2021.643205

Wolfe, J., Duke, M., Miller, S., Schafer, E., Jones, C., Rakita, L., Dunn, A., Browning, S., & Neumann, S. (2022). Evaluation of Potential Benefits and Limitations of Noise Management Technologies for Children with Hearing Aids.. *Journal of the American Academy of Audiology*. https://doi.org/10.1055/s-0041-1735802

Wu, Y., Stangl, E., Chipara, O., Hasan, S., DeVries, S., & Oleson, J. (2018). Efficacy and Effectiveness of Advanced Hearing Aid Directional and Noise Reduction Technologies for Older Adults With Mild to Moderate Hearing Loss.. *Ear & Hearing*. https://doi.org/10.1097/aud.0000000000000672

Zanin, J., Vaisberg, J., Swann, S., & Rance, G. (2024). Evaluating benefits of remote microphone technology for adults with hearing loss using behavioural and predictive metrics. *International Journal of Audiology, 64*, 327 - 335. https://doi.org/10.1080/14992027.2024.2354500

Zhang, H., Moore, B., Jiang, F., Diao, M., Ji, F., Li, X., & Zheng, C. (2025). Neural-WDRC: A Deep Learning Wide Dynamic Range Compression Method Combined With Controllable Noise Reduction for Hearing Aids. *Trends in Hearing, 29*. https://doi.org/10.1177/23312165241309301