Scholarly Journal of Otolaryngology

DOI: 10.32474/SJO.2025.10.000352

ISSN: 2641-1709

Review Article

# Neural and Cognitive Mechanisms of Speech-In-Noise Perception: Implications for Auditory Training and Rehabilitation

**Leila Hess***

*SAERA, School of Advanced Education Research and Accreditation, Spain*

***Corresponding Author:** Leila Hess, SAERA, School of Advanced Education Research and Accreditation, Spain.

**Received:** September 11, 2025 **Published:** October 07, 2025

## Abstract

Speech-in-Noise (SiN) processing refers to the brains ability to hear speech in the presence of background noise; a common complaint amongst those with hearing impairments. There are various cortical regions and pathways involved in processing speech in noise and this review will explore these regions and whether training can improve SiN and the impacts of auditory training programs on these cortical SiN regions. We will first examine the multiple brain systems involved in SiN processing, from subcortical encoding in the brainstem to higher-order cortical processing involving attention, memory, and prediction. Auditory training has been shown to induce neural plasticity, improving bottom-up encoding (e.g., enhanced frequency-following responses and phoneme discrimination) and top-down modulation (e.g., selective attention and executive control. Multisensory integration—including visual, tactile and gesture cues—further supports training can improve speech comprehension in noise. Understanding how training affects these interconnected systems, and the implications of how auditory training can be added to auditory rehabilitation treatment plans, will enable those with hearing loss find ways to improve on their ability to hear better in noise other than relying on hearing aids alone to improve real-world listening outcomes and overall quality of life.

**Keywords:** Speech-In-Noise, Neural Encoding, Auditory Cortex, Auditory Training, Aural Rehabilitation

## Introduction

Speech-in-Noise (SiN) processing is a critical aspect of auditory perception, particularly for individuals with hearing impairments. Understanding the neural mechanisms underlying SiN processing and the potential for neuroplasticity through auditory training programs is essential for developing effective interventions. This literature review synthesizes current research on the brain regions involved in SiN processing, the impact of auditory training on neuroplasticity and cross-modal plasticity, and the components of auditory training programs that contribute to improved SiN performance.

The ability to effectively process auditory information, particularly speech within noisy environments, is paramount for successful communication and social interaction, and its impairment can lead to significant challenges in daily life, especially for individuals with hearing loss [4]. Hearing loss, a prevalent sensory deficit affecting millions worldwide, not only diminishes the clarity of sound but also disrupts the intricate neural processes that underpin auditory perception [29,43].

As the prevalence of hearing loss continues to rise globally, largely due to an aging population and increased exposure to noise, understanding whether adding auditory training to hearing aid fittings as a general practise will enhance the ability for the patient to hearing better in background noise [13].

Traditionally, hearing aids and cochlear implants have been implemented as an assistive technology, primarily aimed at amplifying sound to compensate for diminished auditory sensitivity. However, simply providing auditory input is not sufficient for optimal hearing outcomes [43]. The brain has a remarkable capacity for plasticity to offer alternative and complementary avenues for enhancing auditory function [40]. The integration of auditory training can enhance the efficacy of hearing aids by fostering neuroplastic changes that can improve auditory processing capabilities in individuals who rely on these devices. This interaction between auditory training, neuroplasticity, and hearing aid fittings represents a complex landscape that merits thorough examination.

Neuroplasticity of the human brain refers to the brain's remarkable capacity to reorganize, adapt, modify, and respond to its everchanging environment. This also includes forming new neural connections throughout life based on experiences [82]. This ability to reorganize its structure and function in response to experience, learning, or injury, provides a compelling framework for understanding how auditory training can improve speech perception in noise for hearing-impaired individuals [29,82].

Auditory training, a structured program of exercises designed to improve auditory skills, has emerged as a promising approach to remediate speech-in-noise deficits by harnessing the brain's inherent plasticity [4,71]. The principles of auditory training are rooted in the concept that focused listening and repeated exposure to specific auditory stimuli can induce neuroplastic changes within the auditory system, leading to enhanced perceptual abilities and cortical representation of sounds [6,79].

Auditory training programs can be computer/app based, in-person/clinic based, home based, or remote and utilize a variety of stimuli (syllables, words, sentences, music) which target different auditory skills (detection, discrimination, identification, comprehension), and may incorporate both analytic (focusing on individual sounds) and synthetic (understanding meaning in context) exercises. [79]. As such, understanding the neural mechanisms that drive auditory training-induced plasticity is crucial for optimizing training protocols and maximizing their benefits for hearing-impaired individuals.

## Need of the Study

The field of audiology has witnessed significant advancements in hearing aid technology and audiological rehabilitation strategies, yet challenges persist in optimizing outcomes for individuals with hearing loss [4]. Auditory training has emerged as a pivotal intervention for individuals with hearing impairments, particularly in the context of aural rehabilitation. The concept of auditory training encompasses a variety of techniques aimed at improving auditory skills, including sound discrimination, auditory memory, and listening in challenging environments [75]. Recent advancements in neuroscience have deepened our understanding of auditory training's potential to induce neural plasticity—changes in the brain's structure and function in response to auditory experiences. This master's thesis will explore the potential benefits of auditory training programs in improving cognitive and behavioral outcomes associated with them, including improvements in memory, attention, and language processing [51,68,71,73,76] and investigate how these neuroplastic changes can harness improved hearing aid fitting outcomes [4]. It will also explore the potential applications of this knowledge to address how targeted auditory training interventions may be utilized in the field of aural rehabilitation to improve success in hearing aid fittings to enhance functional outcomes for individuals with hearing impairments.

## Method

A review of the literature was conducted to examine the neural mechanisms underlying speech-in-noise (SiN) processing and the impact of auditory training on these mechanisms. Searches were conducted across PubMed and Google Scholar for relevant content using the specified inclusion and exclusion criteria.

Search terms were grouped into three key domains and combined using Boolean operators:

I. *Speech-in-noise*: "speech-in-noise", "SiN perception", "auditory scene analysis", "cocktail party effect", "speech intelligibility".

II. *Neural mechanisms*: "neural encoding", "auditory cortex", "brainstem responses", "frequency following response (FFR)", "auditory evoked potentials (AEPs)", "EEG", "fMRI".

III. *Auditory training*: "auditory training", "auditory learning", "phoneme discrimination training", "neuroplasticity", "cross-modal plasticity", "LACE", "cognitive training", "musical training", "selective attention", "working memory training", "aural rehabilitation", "auditory perception"

### Inclusion Criteria

Inclusion criteria focusing on speech-in-Noise (SiN) processing included human and animal studies that included auditory learning tasks. Specifically, studies were included if they met the following criteria: 1. Focus on SiN Processing: Research that investigated brain mechanisms involved in perceiving speech in noisy environments. 2. Neural Mechanisms: Articles that reported findings on the neural correlates associated with speech processing, including structural and functional imaging studies. 3. Neuroplasticity and Auditory Training: Studies that explored the effects of auditory training on neuroplastic changes in individuals, particularly those with hearing impairments. 3. Cross-Modal Plasticity: Research that examined the multimodal interactions between auditory processing and other brain regions in relation to SiN performance.

Inclusion criteria for reviewing auditory training programs and how they enhance speech-in-noise processing was specified to individuals, both adults and children who were tasked to an Auditory Training Therapy (ATT) program. These were either computer-based, application "app" based, and either clinician or non-clinician-based delivery. These programs could include but are not limited to Listening and Communication Enhancement (LACE), Read my Quips, and other auditory brain training games, music therapy, or passive auditory stimulation.

The studies included in this review consisted of randomized controlled trials, non-randomized controlled trials, cohort studies with repeated measures in obtaining the results of pre- and post-training as well as systematic reviews and meta-analyses of the existing data.

The outcome measure(s) related to speech perception/intelligibility in the presence of background noise, using validated speech tests and self-reporting questionnaires.

### Exclusion Criteria

In this review we excluded studies that did not meet the inclusion criteria or did not relate to the purpose of this review. Studies evaluating the effects of medications or drugs on the auditory cortex were excluded. Pilot studies, non-peer reviewed studies, and non-English studies, where translation was not available, were also excluded.

## Procedure

A search for peer-reviewed articles published between January 2000 and May 2025 was conducted, emphasizing English-language articles. This broader date range allows the review to capture all relevant studies on the various speech in noise processes to ensure we capture all existing evidence on this topic. It allows for more studies to be included for a more comprehensive and complete review of the existing literature. This also prevents missing important foundational or pivotal studies that might have been published earlier, which could significantly influence the understanding of how the brain processes speech in noise.

In addition to the database-based literature search, qualified additional relevant studies uncovered through citation tracking, and articles in references from central key papers were also reviewed. This approach helped to uncover seminal works and recent advancements that may not have been captured in the initial database search. Furthermore, conference proceedings and clinical trial registries were examined to identify ongoing research and unpublished findings relevant to the topic. Initial screenings of the articles were done on titles and abstracts according to the selection criteria and identified. Related cited works were also included when relevant.

Selection bias does exist by only relying on the availability of full text versions available on PubMed and Google scholar. Language bias exists in including only English-language studies which could exclude important findings although most studies were either available or translated into English Figure 1.

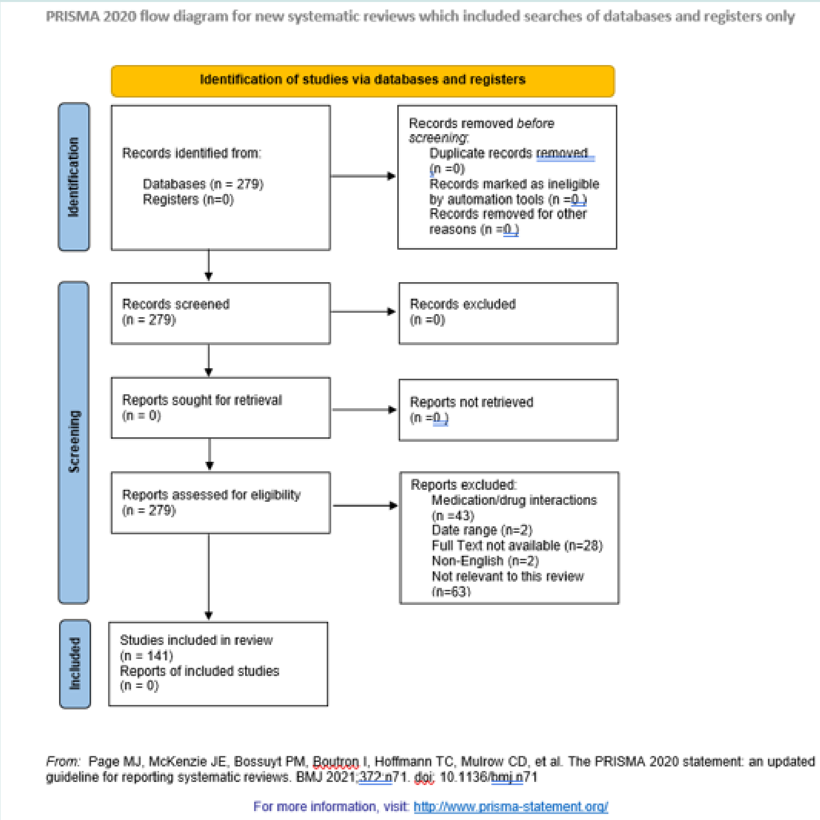


**Figure 1**

## Results

Speech-in-Noise (SiN) processing is the ability to understand speech in the presence of background noise and is the primary complaint of the hearing impaired. This skill is essential for communication in everyday environments, such as noisy restaurants or busy streets, and by improving the individual's ability to process this information, hearing aid satisfaction and quality of life may increase.

Several areas and mechanisms within the brain work to extract speech from noise [70]. Understanding how the brain functions when processing SiN is vital for designing useful auditory training programs. This investigation can identify ways to optimize training protocols and, ultimately, the greatest benefit for those who are hearing impaired. Thus, this review highlights the specific areas of the brain that are used when listening to speech in noise and assesses the influence of auditory training and plasticity on these areas and pathways.

## Speech-in-Noise Processes and Impact of Auditory Training

### Neural Speech Encoding

Initially, the brain encodes speech through a bottom-up auditory analysis that involves the initial processing of the signal within the brainstem. In these initial stages of auditory processing, the encoding of the basic acoustic features of sound, including frequency discrimination, intensity, and timing, is necessary for binaural processing and sound localization (Dubinsky, et al., 2024).

Phonemes are small units of sound that give words meaning and are part of speech encoding. Recognizing phoneme discrimination acts as a pillar for comprehension and meaning in any language. Nuances in phonemes change the meaning of the word, and being able to discriminate between phonemes improves the individual's ability to comprehend speech. Adaptive syllable training and syllable speech variant training have been shown to improve phoneme discrimination and related auditory processing abilities [64,80]. Anderson, et al., 2013 [5] demonstrated that participants who engaged in phoneme discrimination training showed improved performance on tests of speech perception in noise, underscoring the efficacy of targeted auditory training in enhancing real-world listening skills. Phoneme discrimination and temporal processing also improved following auditory training with Listening and Communication Enhancement Program "LACE" [5,71,76] and in another study involving older listeners enrolled in a 12-week program using a word-based auditory training protocol [11].

Frequency following response (FFR) is an electrical brainwave that measures the neural activity following the fundamental frequency (F0) of a repeating sound such as speech or music. FFR is used to study how the auditory system processes more complex sounds and how it changes with learning and experience. In SiN tasks, researchers observed the strength of EEGs and source reconstruction to investigate which brain regions contribute to FFR and their relationship to SiN processing [8].

The researchers found that FFR strength decreased from subcortical to cortical areas. Importantly, the strength of FFRs in the auditory nerve and brainstem/midbrain, but not the primary auditory cortex, was related to how well participants performed on a SIN task, suggesting that subcortical activity is key to the link between electrical FFRs and SIN skills measured by EEG [8]. Similarly, findings show that accurate encoding of F0 within the brainstem is crucial for SiN processing [70].

EEGs related to slow syllable rate changes, the rate of the fundamental frequency (pitch), and the fine details of the sound showed that better SPIN performance was linked to stronger brain responses to the slow syllable rate and the fine details of resolved harmonics. Responses to the fundamental frequency rate only showed a relationship to SPIN when hearing loss was taken into account [47]. The study suggests that how well the brain locks onto these temporal cues is important for hearing speech in noise and has implications for fitting hearing devices.

Using magnetoencephalography (MEG), researchers determined that stronger neural encoding of the F0 in the brainstem, thalamus, and cortex, along with a larger cortical P2 response, correlated with better SiN accuracy. These enhanced neural responses were also linked to musical training [17]. The findings suggest that musical training improves neural networks with robust bottom-up sound encoding which lays a good foundation of better auditory information for later processing.

Short-duration auditory training programs have been linked to changes in the frequency following response (FFR) to F0; therefore, the underlying physiological response appears to index improved neural processing of this critical speech cue. One study showed that short-term auditory training aimed at speech-in-noise perception in young adults resulted in enhanced FFRs to speech sounds following training. This study reported enhancements in the encoding of pitch-related cues, including the fundamental frequency [71].

Another recent study examined the effects of LACE on the neural encoding of speech in multi-talker noise, as measured by the FFR response. They found that the latency of the FFR response was earlier, suggesting faster neural timing in the reference speech presented in noise. Although the focus of this study was on latency, faster and more robust FFRs are often indicative of enhanced neural synchrony to F0 and other speech features. Furthermore, this study suggests that auditory training, such as LACE, can reduce the neural delays encountered when processing background noise and may improve speech perception in noise through specialized auditory training [67].

Musical training has been shown to significantly improve the ability to understand speech in noisy environments through various mechanisms [17]. This phenomenon, often referred to as the "musician advantage," arises from the shared neural pathways utilized by music and speech perception, which involve similar processes for pitch, timing, and timbre [3].

Research indicates that musicians demonstrate superior neural responses to intricate sounds, such as speech and music, characterized by more pronounced N1 and mismatch negativity (MMN) components when responding to auditory inputs [59,78].

Passive auditory training such as singing in a choir also shows enhanced performance for SiN processing. Over 10 weeks, a choir group showed improvements in SIN perception, pitch discrimination, and the neural representation of speech pitch compared to a control group. The choir group had better pitch discrimination, which itself was associated with stronger neural processing of speech sounds (Dubinsky, et al.,2024). The findings suggest that even short-term choir participation can be an effective way to counteract age-related hearing difficulties.

As a result of this enhanced processing capability, musicians are better equipped to distinguish target sounds amid background noise, an essential skill in “cocktail party” scenarios where several conversations occur simultaneously [58].

## Selective Attention and Filtering

The “cocktail party effect” was originally coined by Colin Cherry in 1953 and is often referred to when describing the ability to focus on speech while ignoring conflicting noise. Auditory selective spatial attention (ASSA), which was explored by Liu, et al. alludes to the fact that there are many processes involved in this type of selective attention [45].

Brainstem activity is modulated by selective attention, which enhances the processing of target speech and suppresses background noise. This modulation is evident in the amplitude and latency of brainstem responses, such as wave V of the ABR, which is larger and faster when attention is directed toward target speech [49,63].

Selective attention and filtering mechanisms operate in both bottom-up and top-down auditory processing. Bottom-up processing regards the physical characteristics of sounds, such as spectral and temporal cues, to replace speech with noise, while top-down processing brings in cognitive factors such as attentional states or prior knowledge to weigh the relevance of auditory signals [23,81].

Selective attention is a crucial mechanism that regulates neural activity to focus on the most relevant stimuli while minimizing distractions from irrelevant information [73]. This selection process plays a vital role in decoding meaningful signals amid competing sounds, facilitating effective communication and enhancing awareness of one’s surroundings [69].

For those with hearing impairments, the ability to filter out unnecessary auditory information becomes essential when processing speech against background noise [11]. In individuals with normal hearing sensitivity, these filtering mechanisms operate efficiently, aiding communication even in challenging acoustic environments. Conversely, hearing loss complicates this filtering process and can significantly impair speech understanding in situations where multiple voices are present (Carta, et al., 2023, 2024).

The prefrontal cortex is crucial for managing attention by adjusting sensory processing based on behavioral objectives, particularly in challenging situations [73]. Subcortical structures, such as the brainstem, contribute to the recognition of F0 cues that are essential for understanding speech amid background noise. When individuals concentrate on a specific sound, the variability in cortical auditory-evoked responses decreases, demonstrating improved neural consistency. This effect is especially strong in individuals with musical training, highlighting how experience can influence attentional processes [73].

Musical training fosters better inhibitory control, enabling individuals to filter out irrelevant sounds more efficiently. For instance, musicians often show diminished P3a responses to distracting new noises, reflecting less distraction and a greater ability to focus on relevant stimuli [59].

The secondary auditory cortex is involved in more complex tasks, such as speech recognition, auditory scene analysis, and segregation of target speech from background noise. One study measured the effects of Automatic Auditory Scene Classification (SCAN) in children wearing cochlear implants. Post-training there was a 4.6dB improvement for speech in noise perception compared to the control group [16].

Speech-in-noise perception is enhanced when the auditory cortex and brainstem communicate with each other to improve the neural encoding of speech [9]. This is shown more specifically in the right superior temporal gyrus (STG) [71]. However, significant activity was also found in the left STG for speech-in-noise [10]. One study showed that the left middle temporal gyrus and bilateral cerebellum also showed increased activity along with the left STG in normal hearing individuals when processing SiN [48].

For auditory training to enhance selective attention and filtering, the brain utilizes a top-down learning method to understand speech in noise. Top-down learning generally refers to cognitive elements, specifically functional memory and attention control. These types of programs usually include assignments in which participants have to focus on some auditory streams while ignoring others, thus focusing more on selective attention [30,60]. Neurofeedback training, for example, is an attempt to improve attentional regulation of auditory responses, which has been shown to improve cortical responses to target speech items and subsequent performance in background noise [38].

Research indicates that neurofeedback training can increase beta oscillations associated with top-down processing and the prediction of specific sounds [38,66]. In addition, alpha oscillations have been associated with improved spatial inhibition processing, which enables individuals to focus their attention on desired sounds amidst a noisy environment [38,66].

Cortical late event-related potentials and the Hearing in Noise Test (HINT) were used to determine the efficacy of hearing aid use and the ReadMyQuips (RMQ) auditory training program. They used the P3a and P3b components to provide insights into the neural mechanisms of selective attention. The P3a component is related to involuntary attentional switches to distractors, whereas the P3b component is related to target detection and attentional allocation [60]. In this study, the RMQ reduced the P3a amplitudes, indicating less distraction, and increased the P3b amplitudes, representing a higher success rate of target detection and better speech perception in noise [60].

## Cognitive Integration and Prediction

Cognitive integration is the interaction between the auditory and cognitive systems to improve speech perception. Speech processing and perception in noise are part of top-down mechanisms, including lexical, syntactic, semantic, contextual, and visual cues, as well as executive functions (switching, working memory, and cognitive flexibility).

**Lexical knowledge:** which pertains to the understanding of words and their meanings, plays a crucial role in anticipating words during speech processing, especially in noisy conditions. For instance, EEG data indicate that context-based word predictability is closely correlated with speech comprehension [77]. Additionally, research on perceptual learning emphasizes that lexical predictability is beneficial in learning to perceive degraded speech, and greater learning gains occur with low predictability [44]. Evidence has indicated that people with stronger lexical knowledge benefit more in terms of speech recognition in noise because when the auditory signal is degraded, lexical knowledge can compensate for the missed portions of the auditory signal [37,42].

Syntactic knowledge involves the rules governing sentence structure and aids in predicting upcoming words and phrases, thereby enhancing comprehension under challenging listening conditions [22,25]. For example, slower speech rates have been shown to reduce contextual facilitation, particularly for low-predictability sentences, indicating that syntactic processing is sensitive to temporal factors [7]. Early speed clues remain important even when later words provide a different meaning. The influence of early speech rate cues persist despite the presence of later, lexically disambiguating information [36].

**Semantic knowledge:** involves the meaning of language and works into meaningful contexts, aiding comprehension by providing a framework for interpreting sentences. Research has demonstrated that semantic predictability enhances speech processing, with highly predictable words eliciting reduced N400 amplitudes in clear speech and delayed effects in noisy conditions [34]. The brain's semantic prediction processes are correlated with frontoparietal connectivity, particularly in the left temporoparietal junction and bilateral inferior frontal gyrus, which are coupled more with the control and memory networks when anticipating highly predictive texts [32]. In large-scale network interactions, semantic integration demands modulate with increased activity in the left frontal and temporal regions, along with stronger interactions between task-positive and default mode networks. The brain uses predictive processing to anticipate context in noisy environments to "hear better" [52]. Older adults with hearing loss exhibit increased use of semantic context, particularly in high-context tasks, suggesting a compensatory mechanism [65].

**Semantic knowledge:** is a critical component of auditory training. By leveraging semantic context, listeners can better interpret ambiguous or degraded speech signals, especially in noisy environments [42,55].

**Contextual Cues:** can be linguistic or non-linguistic in nature. Essentially, if listeners can use contextual cues indicating what they should expect, they typically have an advantage in processing speech in noise. As previously mentioned, linguistic context (i.e., constraints at the sentence level) supports listeners in recognizing words (and sentences) by alluding to some predictive information [24]. Semantic context aided the identification of a word-in-noise, even though the linguistic cue was presented after the target word, indicating that listeners possess a function for engaging memory to support contextual processing [14].

Meaningful knowledge of listeners is not only an important aspect of auditory training, but when the semantic context is engaged, listeners can make informed inferences when any speech signal is ambiguous or degraded, especially in noise [30,31].

In a recent study participants listen to AI-generated dialogues or monologues, either intact or scrambled, and were able to identify changes in a target sentence presented in noise demonstrating that both social context (dialogue vs. monologue) and semantic context (intact vs. scrambled) improved speech-in-noise processing [1].

**Visual Cues:** Auditory signals combined with visual elements, such as lip movements and facial expressions, provide contextual awareness for auditory processing, particularly in the presence of noise [31,50]. For example, visual cues associated with lip-syncing can assist in speech comprehension and have been shown to improve intelligibility by providing supporting information to the auditory signal [31,35].

The integration of auditory and visual information in the Posterior Superior Temporal Sulcus (pSTS) and Superior Temporal Gyrus (STG) enhances speech perception in noise. This integration is thought to occur at multiple stages of processing, from basic sensory processing to higher-level linguistic processing [18,62].

Visual cues are used to predict what the next auditory signal will be. This cross modal predictive process is believed to happen at the level of the STG/pSTS which relies on visual information to construct an expectation for what the individual will receive as auditory input. This approach seems to operate most effectively in an environment with noise vs one of clarity, when the auditory signal is less reliable [28].

The visual cortex within the occipital cortex is responsible for visual-speech processing. Visual speech cues, such as lip movements and facial cues, are used to extrapolate and supplement auditory-based speech information. Research has shown that the visual cortex can phase-lock to auditory speech signals, particularly in noisy conditions, which shows that these structures are important for audiovisual integration [2,83]. The amplitude envelope of speech plays an important role in representing rhythmic and syllable structure. In noisy conditions, visual speech information improves neural tracking of the speech envelope, particularly within the right auditory cluster [26].

## Multisensory Integration

Multisensory integration of cues from amplitude envelopes, can also improve speech perception in noise. Researches showed that adding synchronized visual, tactile, or both cues to auditory input significantly improved word recognition, particularly in higher noise levels, with the combination of visual and tactile cues yielding the greatest benefit [53,54]. Body language and gestures are also integrated in this multisensory analysis, for more precise word recognition. This neural processing suggests multisensory integration beyond the mere sum of its parts, emphasizing the brain's capability to access context rich cues to enhance the capacity to comprehend speech amid competing sounds [50,56].

One study involving cochlear implant users found that individually tailored auditory-cognitive training and therapeutic intervention can improve speech understanding in noise, showing some potential for integrating cognitive abilities as part of the auditory rehabilitation process [46].

## Working Memory

An individual with stronger working memory can process and hold onto auditory material in an active way, which contributes to their ability to perform better in noisy listening situations. The ability to improve working memory through auditory training programs has been linked to enhanced speech recognition abilities in noise, especially among older individuals or those with hearing loss [25,84].

Auditory training programs that included cognitive demands like working memory, attention, and executive function enhancement provided a significant benefit for speech-in-noise perception. For example, participants of an auditory-cognitive training (ACT) program were found to have improved speech-in-noise perception, especially when the auditory-cognitive training involved multi-modal options [21,25].

## Prediction

In noisy environments the brain uses a "predictive coding model" to generating expectations or predictions about incoming sensory information. One study looked at electrophysiological markers of anticipation by measuring Stimulus-Preceding Negativity (SPN) to investigate the role of prediction in speech comprehension in noise in young adults with normal hearing. When testing speech-in-noise performance, the researchers found that the amplitude of the SPN was larger, suggesting that the SPN could be a potential neural marker to use while investigating how individuals use top-down processing to predict and comprehend speech in challenging listening contexts. In this type of situation individuals take sensory information and prior knowledge and expectations into consideration [55]. When then brain engages in top-down predictions it relies on the context or "world knowledge" of language, and situational cues, in order to predict what sounds and/or words may occur and to then match the prediction within the noisy sensory input [74]. Studies examining event-related potentials (ERPs) assessing predictive processing have identified that the specific ERP components used in predictive processing, such as the N400, reflect the integration of both semantic and syntactic information [34]. The effect of the N400 is reduced in noisy environments, indicating the interference of predictive processes.

Behavioral studies with hearing aid users with mild to moderate hearing losses have demonstrated that training programs that emphasize predictive processing, including functional real-world listening tasks like competing speech and listening-memory dual-tasks, can lead to significant improvements in speech recognition in challenging environments, not only on the trained task but also in the functional listening tasks [33]. This type of auditory training takes the listener's prior knowledge and expectations to focus on cognitive integration and prediction to hear better in noise [22,55].

## Auditory Motor Interactions

According to the motor theory of speech perception, we comprehend speech by subconsciously considering our own vocal tract movements, or motor gestures, used to generate those sounds (Galan Tucci, et al., 2006). Within the motor cortex, researchers found distinct neural patterns with listening which evoked activity in specific ventral motor cortex regions. They were organized by acoustic features, rather than the widespread, articulator-specific activity seen during speaking. The findings suggest the motor cortex represents auditory vocal information during listening, not the articulatory gestures themselves [15,72].

High-resolution fMRIs allow for precise localization to identify the overlap between brain regions involved in speech production and perception [41]. Activity in the left motor, premotor, and frontal cortex were found to be correlated with phoneme category differences during a /ba/-/da/ discrimination task, and multivariate analysis which indicated that areas associated with speech production also contribute to perceptual categorization.

To study the link between speech production and perception, participants were asked to either repeat syllable pairs aloud or perform a same-different judgment task. They were then tested on their ability to discriminate syllables in noise. The consistent finding across four experiments was that prior syllable repetition improved subsequent syllable discrimination in noise, both for repeated and new syllables, suggesting enhanced sensitivity to phonological details. This demonstrates significant improvements in speech perception in noise, particularly for both primed and new syllable pairs. The benefits were observed even after a delay, suggesting long-term plastic changes in the brain [57].

The dorsal stream in our brains is vital for understanding sounds in space and time, and for turning what we hear into actions [39,61,85]. For example, when it comes to music, this part of the brain helps us coordinate our actions to produce musical sounds, which relate to timing in music. It also helps us predict when certain sounds will happen and connects that timing with our movements. Interestingly, musical training seems to boost the links between the parts of the brain that handle hearing and movement, primarily linking the auditory cortex with the parietal lobe, dorsal premotor cortex, and the dorsolateral prefrontal cortex, among others [85]. Musical training has also been shown to strengthen the functional connectivity between auditory and motor cortices, particularly in the left hemisphere [19,86]. Musicians and individuals with musical training consistently outperform non-musicians in speech perception tasks in noise. This advantage is associated with stronger activation of the left inferior frontal gyrus and premotor cortex, as well as greater specificity of phoneme representations in these regions [19,86].

## Discussion

Hearing in the presence of background noise is the most challenging obstacle for individuals with hearing loss. This poses a great challenge on the auditory and related systems. The brain is adaptable and constantly adjusts, using all necessary systems to reach the goal of extracting speech from background noise.

This literature review explored a multifaceted array of mechanisms to understand the neural and cognitive mechanisms involved in speech-in-noise processing. Understanding the intricate interplay of these mechanisms has profound implications for the development and refinement of auditory training programs and hearing aid fitting rationales.

By synthesizing a wide range of studies, the review delineates the complex interplay between bottom-up neural encoding, top-down cognitive modulation, multisensory integration, and predictive processing, and underscores the clinical relevance of these findings in auditory rehabilitation.

Neural Encoding and Auditory Training: Enhancing Brainstem and Cortical Responses

One of the broadest findings derived from the literature review is that SiN processing starts with the encoding of different acoustic features at the brainstem level. The Frequency-Following Response (FFR) and the Auditory Brainstem Response (ABR), provide the clearest evidence, of the qualitative extent of early auditory processing at the neural encoding level. For example, F0 Frequency-Following Responses (FFR) related to brainstem stability show that individuals with strong FFR demonstrate superior SiN perception. As such, we have proposed that subcortical precision is a fundamental prerequisite for managing effective speech comprehension in noise [8,17].

These early neural processes are certainly not fixed. The summarized findings demonstrate that short-duration auditory training can result in brainstem-level plasticity with evidence of improved neural timing, improved phase-locking, and increased FFR amplitude [67,71]. These neuroplastic changes support the fact that even transient interventions can make a true and meaningful difference in neural speech encoding. This is especially encouraging for people with hearing loss who may show brainstem responses that are delayed or distorted in noise [71].

Musical training also increased the fidelity of brainstem responses and SiN perception, which adds stronger evidence that various long term, goal-directed auditory experiences shape early auditory processing. These considerations carry significant implications since auditory training programs that mirror the demands or potential benefits of musical training—with a focus on pitch, timing, and temporal structure—might translate to improved SiN perception.

Selective Attention and Cognitive Filtering: Neural Gatekeeping in Noisy Environments

Selective attention to relevant speech while ignoring irrelevant noise, is important for performance in SiN. This selectivity uses both bottom-up and top-down mechanisms which is important for enhancing representation of targeted stimuli while suppressing a distractor [23]. The review identified the effects of attentional modulation to brainstem responses, such as wave V of the ABR and cortical event-related potentials, P3a, and P3b [49,73].

Auditory training seems to be able to engage the mechanisms of attention. For example, neurofeedback training has been found to result in greater beta oscillations and attention in SiN tasks [38]. Similarly, using top-down training protocols, where participants have to focus on a target while ignoring distractors, has yielded greater P3b amplitude and lower P3a, suggesting greater allocation of attention and decrease susceptibility to distraction [60].

The role of the prefrontal cortex in modulating attention reinforces a top-down model of auditory processing. The prefrontal cortex serves as a neural gatekeeper, by guiding attentional resources, influencing sensory areas to facilitate efficient SiN processing. Again, it is interesting to find that musical training had greater effects, and that musicians have inhibitory control and reduced neural variability that could transfer to non-musical listening tasks.

Cognitive Integration and Predictive Processing: Leveraging Prior Knowledge

Cognitive integration—the coordination of attention, working memory, executive function, and linguistic knowledge—all work together to enable successful speech understanding, especially in noisy situations. The literature showing evidence of top-down cognitive mechanisms allow listeners to draw inferences, fill in perceptual holes, and use context to aid recognition and understanding [21,46].

One notable finding is the reflexive role of prediction in perception of SiN. Predictive coding models suggest that the brain is continuously generating predictions based on prior knowledge, which will help direct perception when sensory input is ambiguous or degraded [55]. The N400 amplitude in electrophysiological studies showed reduced N400 amplitude for semantically predictable words, even in noisy conditions, demonstrating that semantic prediction is an important cognitive mechanism for speech understanding [34].

Training based on predictive processing could lead to improved speech recognition in noise by utilizing lexical, syntactic, and semantic training [12]. An example of this is training using events or passages with rich data or contextual information. An even better training effect may occur if the tasks prompted the listener to predict what the next word or that the listener had to anticipate a certain word order; this is like the language process of contextual access. This type of cognitive-linguistic training may be highly relevant for older adults with hearing loss who depend more on contextual knowledge in uncertain listening conditions with reduced audibility [65]. Working memory is paramount to cognitive integration. Those with superior working memory perform better with SiN scenarios, presumably because they can retain the auditory information over time, while manipulating the information. Training members of the group on the characteristics of working memory—through dual-task training or through tasks that develop memory span capacity—could improve SiN [25].

Multisensory and Audiovisual Integration: Enhancing Perception with Additional Modalities

Significant improvements in speech perception in noise occur with multisensory integration, such as the presence of visual and tactile inputs. The posterior superior temporal sulcus (pSTS) and superior temporal gyrus (STG) are two important hubs for audiovisual integration, as it permits listeners the ability to rely on lip-reading, facial expressions, and potentially tactile feedback when hearing is degraded [18,31]. Visual information not only enhances the neural tracking of the speech envelope, but also allows for auditory cortical responses to be phase-locked, allowing auditory cortical responses to compensate from reduced auditory clarity [2]. This realization is crucial, especially in real-world listening contexts when visual aids are at the listeners' disposal, but typically not maximized by formal auditory training programs. Auditory training programs can be modified to include audiovisual elements such as videos, or face to face modelling, to ensure they are maximally effective.

Visual cues predict subsequent auditory information through cross modal anticipatory proceeses, which can shorten processing delays and further improve speech intelligibility [28]. This provides compelling rationale for embedding cross modal tasks into auditory training programs, especially for individuals who have difficulties with temporal processing, or who experience auditory-only limitations.

**Auditory-Motor Integration: Sensorimotor Facilitation of Perception**

According to the motor theory of speech perception, the motor system is engaged in interpreting speech sounds, particularly in degraded listening conditions [27]. In particular, this review draws attention to the motor regions (i.e., premotor) involved in speech perception (e.g., Broca's area) which may provide a "mirror" template for interpreting ambiguous auditory stimuli [15,72].

These kinds of observations have important implications for auditory training. Programs that involve vocal repetition, speech production, or even imagined articulation may activate the motor system to strengthen perception. For example, shadowing tasks that involve participants repeating speech in real-time may foster favour auditory-motor coupling and enhance their integration during sensorimotor tasks.

Additionally, the involvement of motor areas in spoken language processing may indicate some overlap with other remedial approaches such as speech-language therapy or neurostimulation that targets motor circuits. Investigating how to coordinate additional therapy or treatment may present better opportunity in fewer sessions from a more integrative approach to rehabilitating those with hearing impairments.

The following table summarizes SiN processes, the region of the brain associated to the process, and the respective auditory training programs shown to enhance SiN processing in these areas Table 1.

**Table 1:** SiN Process, Brain Regions, and Associated Auditory Training Programs (ATP).

| SiN Process | Brain Region | Auditory Training |
|---|---|---|
| Neural Speech Encoding | Brainstem (bottom-up processes) | Adaptive Syllable training, syllable speech variant training, LACE, word based auditory training, musical training |
| Selective Attention and Filtering | Prefrontal Cortex, STG, cerebellum, and Modulating brainstem activity | SCAN, musical training, selective attention and neurofeedback training, ReadMyQuips |
| Cognitive Integration and Prediction | Top-down processes | Individually tailored ATP, working memory training, Auditory-Cognitive Training, Predictive Processing Training (competing speech and listening-memory dual tasks) |
| Multisensory and Audiovisual | pSTS, STG | Musical Training, Face-to-face modelling, audiovisual training |
| Auditory Motor Integration | Motor regions, Broca's area | Vocal repetition, speech production, shadowing tasks, reading out loud |

**Implications for Clinical Practice and Future Research**

The evidence presented here illustrates the complexity of SiN perception and highlights the need to address multiple neural/cognitive systems in auditory training. We will not have one-size-fits-all protocols. To be effective, auditory training needs to be personalized by targeting an individual's hearing status, cognitive limits, and previous experiences (e.g., musical experience).

Secondly, training protocols should certainly go beyond auditory training alone, even though auditory training should be the primary attention, and will also need to be multimodal, we must consider visual, language, cognition, and motor skills. Programs like LACE and ReadMyQuips, however they may benefit, will require further development and thoughtful examination to best utilize neural plasticity and address additional systems. Also, objective metrics like EEG, FFRs, and ERPs could, and should also be, consistently used to monitor progress and personalized intervention strategies.

From the experiential research point-of-view, longitudinal studies are needed to observe that the training effects are sustained and to evaluate the best duration, intensity, and time training within each session. We need to design training programs where the duration is sufficiently long and challenging, to ensure the maximal effectiveness and best training outcomes for a person with hearing loss including those with auditory processing disorders. Future studies could also include the observation of neural mechanisms involved with changes induced by training, and how we can develop better auditory rehabilitation methods, including the use of neurofeedback, brain-computer interfaces, or a real-time picture of the neural activity.

Future study directions could include a more in-depth review of how specifically aging, retrocochlear pathologies, and those with cochlear synaptopathy or "hidden hearing loss" affects speech-in-noise processing. A review on whether auditory training can improve on the processing of these areas can give insight into how these populations can enhance communication and quality of life.

## Conclusion

Orchestrating Clarity from Noise – The Promise of Auditory Training

Understanding speech-in-noise (SiN) is an integral part of everyday conversations and quality of life, however especially difficult for individuals with hearing impairments. In this review, we have established that the complex processes involved in SiN, from the encoding of auditory stimuli in the brainstem to the integrated cortical processes and predictive aspects, are complex and multidimensional.

We have illuminated some aspects of how aspects of auditory training may improve the processes of neural encoding of speech, by promoting crucial early brainstem responses to fundamental speech, which could result in phoneme discrimination. Training also benefits selective attention in noise and cognitive filtering in noise, when using working memory and cognitive processes to predict speech, with lexical, semantic, and visual sources of information. Auditory-motor processes can also support SiN processing by engaging the speech-motor system through various activities, such as musical or articulation training.

In summary, improving SiN processing involves much more than amplifying sound for "better hearing", but rather engaging the brain to intentionally interpret stimuli, adapt and refine cognitive function and information. There is robust evidence to suggest that, with targeted auditory training inducing brain plasticity in specific neural pathways, enhancing well-documented auditory behaviours for communication will be engendered. By refining training protocols to take into consideration the interactive and fluent nature of auditory and visual information, where both are active and intertwined, we could significantly improve hearing aid use, and ultimately, people's quality of life in difficult listening social situations.

## Competing Interest

None.

## Acknowledgements

None.

## References

1. Abassi E, Zatorre RJ (2024) Influence of social and semantic contexts on phonetic encoding in naturalistic conversations. Neuroscience.
2. Aller M, Økland HS, MacGregor LJ, Blank H, Davis MH (2022) Differential Auditory and Visual Phase-Locking Are Observed during Audio-Visual Benefit and Silent Lip-Reading for Speech Perception. The Journal of Neuroscience 42(31): 6108-6120.
3. Anderson S, Kraus N (2011) Neural Encoding of Speech and Music: Implications for Hearing Speech in Noise. Seminars in Hearing 32(02): 129-141.
4. Anderson S, Kraus N (2013) Auditory Training: Evidence for Neural Plasticity in Older Adults. Perspectives on Hearing and Hearing Disorders Research and Diagnostics 17(1): 37.
5. Anderson S, White Schwoch T, Choi HJ, Kraus N (2013) Training changes processing of speech cues in older adults with hearing loss. Frontiers in Systems Neuroscience 7: 97.
6. Atilgan H, Walker KM, King AJ, Schnupp JW, Bizley JK (2023) Auditory training alters the cortical representation of complex sounds. Neuroscience 45(18): e0989242025.
7. Bhandari P, Demberg V, Kray J (2024) Speaking fast and slow: How speech rate affects contextual facilitation in degraded speech comprehension. PsyArXiv.
8. Bidelman GM, Momtaz S (2020) Subcortical rather than cortical sources of the frequency-following response (FFR) relate to speech-in-noise perception. Neuroscience 746: 135664.
9. Bidelman GM, Price CN, Shen D, Arnott SR, Alain C (2019) Afferent-efferent connectivity between auditory brainstem and cortex accounts for poorer speech-in-noise comprehension in older adults. Neuroscience 382: 107795.
10. Bishop CW, Miller LM (2009) A Multisensory Cortical Network for Understanding Speech in Noise. Journal of Cognitive Neuroscience, 21(9): 1790-1804.

11. Burk MH, Humes LE (2008) Effects of Long-Term Training on Aided Speech-Recognition Performance in Noise in Older Adults. Journal of Speech Language and Hearing Research 51(3): 759-771.
12. Burkhardt P, Müller V, Meister H, Weglage A, Lang Roth R, et al. (2022) Age effects on cognitive functions and speech-in-noise processing: An event-related potential study with cochlear-implant users and normal-hearing listeners. Frontiers in Neuroscience 16: 1005859.
13. Chadha S, Kamenov K, Cieza A (2021) The world report on hearing, 2021. Bulletin of the World Health Organization 99(4): 242-242A.
14. Chan TMV, Alain C (2019) Listening back in time: Does attention to memory facilitate word-in-noise identification? Attention Perception Psychophysics 81(1): 253-269.
15. Cheung C, Hamilton LS, Johnson K, Chang EF (2016) The auditory representation of speech sounds in human motor cortex. eLife 5: e12577.
16. Ching TYC, Zhang V, Nel E, Hou S, Incerti P (2024) Effects of automatic auditory scene classification on speech perception in noise and real-world functional communication in children using cochlear implants. Cochlear Implants International 25(6): 422-433.
17. Coffey EBJ, Chepesiuk AMP, Herholz SC, Baillet S, Zatorre RJ (2016) Neural correlates of early sound encoding and their relationship to speech in noise perception. Neuroscience 11: 479.
18. Crosse MJ, Di Liberto GM, Lalor EC (2016) Eye Can Hear Clearly Now: Inverse Effectiveness in Natural Audiovisual Speech Processing Relies on Long-Term Crossmodal Temporal Integration. The Journal of Neuroscience 36(38): 9888-9895.
19. Du Y, Zatorre RJ (2017) Musical training sharpens and bonds ears and tongue to hear speech better. Proceedings of the National Academy of Sciences 114(51): 13579-13584.
20. Dubinsky E, Wood EA, Nespoli GA, Russo F (2019) Short-Term Choir Singing Supports Speech-in-Noise Perception and Neural Pitch Strength in Older Adults with Age-Related Hearing Loss. Toronto Metropolitan University 13: 1153.
21. Ferguson MA, Henshaw H (2015) Auditory training can improve working memory, attention, and communication in adverse conditions for adults with hearing loss. Frontiers in Psychology 6: 556.
22. Ferguson M, Henshaw H (2015) How Does Auditory Training Work? Joined-Up Thinking and Listening. Seminars in Hearing, 36(04): 237-249.
23. Fiedler L, Johnsrude I, Wendt D (2023) Salience-dependent disruption of sustained auditory attention can be inferred from evoked pupil responses and neural tracking of task-irrelevant sounds. Neuroscience 45(14): e2066232025.
24. Fitzgerald LP, DeDe G, Shen J (2024) Effects of linguistic context and noise type on speech comprehension. Frontiers in Psychology 15: 1345619.
25. Frei V, Giroud N (2024) A short-term immersive auditory-cognitive training improves speech in noise perception in older adults with varying hearing acuity and working memory capacity. In Review.
26. Frei V, Schmitt R, Meyer M, Giroud N (2023) Visual speech cues enhance neural speech tracking in right auditory cluster leading to improvement in speech in noise comprehension in older adults with hearing impairment. In Review.
27. Galantucci B, Fowler CA, Turvey MT (2006) The motor theory of speech perception reviewed. Psychonomic Bulletin & Review 13(3): 361-377.
28. Ganesan K, Plass J, Beltz AM, Liu Z, Grabowecky M (2020) Visual speech differentially modulates beta, theta, and high gamma bands in auditory cortex. Neuroscience 54(9): 7301-7317.
29. Glick H, Sharma A (2017) Cross-modal Plasticity in Developmental and Age-Related Hearing Loss: Clinical Implications. Hearing Research 343: 191-201.
30. Gohari N, Dastgerdi ZH, Rouhbakhsh N, Afshar S, Mobini R (2023) Training Programs for Improving Speech Perception in Noise: A Review. Journal of Audiology and Otology 27(1): 1-9.
31. Guastamacchia A, Galletto A, Riente F, Shtrepi L, Puglisi GE (2024) Impact of contextual and lip-sync-related visual cues on speech intelligibility through immersive audio-visual scene recordings in a reverberant conference room. INTER-NOISE and NOISE-CON Congress and Conference Proceedings 270(2): 9174-9185.
32. He Y, Shao X, Liu C, Fan C, Jefferies E (2025) Diverse Frontoparietal Connectivity Supports Semantic Prediction and Integration in Sentence Comprehension. The Journal of Neuroscience 45(5): e1404242024.
33. Henshaw H, Ferguson MA (2014) Assessing the benefits of auditory training to real-world listening: Identifying appropriate and sensitive outcomes 4: 45-52.
34. Hsin C H, Chao P C, LEE C Y (2023) Speech comprehension in noisy environments: Evidence from the predictability effects on the N400 and LPC. Frontiers in Psychology 14: 1105346.
35. Hussain T, Dashtipour K, Tsao Y, Hussain A (2024) Audio-Visual Speech Enhancement in Noisy Environments via Emotion-Based Contextual Cues (Version 1). arxiv.
36. Kaufeld G, Naumann W, Meyer AS, Bosker HR, Martin AE (2020) Contextual speech rate influences morphosyntactic prediction and integration. Language Cognition and Neuroscience 35(7): 933-948.
37. Kim MW, Ryu JW, Kim EJ (2005) Speech Recognition by Integrating Audio, Visual and Contextual Features Based on Neural Networks. In L Wang K Chen, YS Ong (Eds), Advances in Natural Computation (Vol. 3611, pp. 155-164). Springer Berlin Heidelberg.
38. Kim S, Emory C, Choi I (2021) Neurofeedback Training of Auditory Selective Attention Enhances Speech-In-Noise Perception. Frontiers in Human Neuroscience 15: 676992.
39. Kim S, Schwalje AT, Liu AS, Gander PE, McMurray B (2019) Pre- and post-target cortical processes predict speech-in-noise performance. Neuroscience 228: 117699.
40. Kral A, Sharma A (2023) Crossmodal plasticity in hearing loss. Trends in Neurosciences 46(5): 377-393.
41. Lankinen K, Ahveninen J, Uluç I, Daneshzand M, Mareyam A, et al. (2023) Role of Articulatory Motor Networks in Perceptual Categorization of Speech Signals: A 7 T fMRI Study. Neuroscience 33(24): 11517-11525.
42. Lee H, Shin E-Y (2023) The Necessity for Auditory Training with Idiomatic Expressions. Audiology and Speech Research 19(4): 215-221.
43. Lesica NA (2018) Why Do Hearing Aids Fail to Restore Normal Auditory Perception? Trends in Neurosciences 41(4): 174-185.
44. Li Y, Fan C, Liu C, Li X (2023) The modulating effect of lexical predictability on perceptual learning of degraded speech. Frontiers in Language Sciences 2(1): 1139073.
45. Liu H, Bai Y, Zheng Q, Liu J, Zhu J, et al. (2024) Electrophysiological correlation of auditory selective spatial attention in the "cocktail party" situation. Human Brain Mapping 45(11): e26793.
46. Lloret G, Vincent C, Risoud M, Beck C, Lemesre PE, et al. (2024) Evaluation of a personalized auditory-cognitive training on the improvement of speech understanding in noise in cochlear implanted patients. Cochlear Implants International 25(6): 467-476.
47. Mai G, Tuomainen J, Howell P (2018) Relationship between speech-evoked neural responses and perception of speech in noise in older adults. The Journal of the Acoustical Society of America 143(3): 1333-1345.
48. Manan HA, Yusoff AN, Franz E A, Mukari SZ-MS (2013) The effects of background noise on brain activity using speech stimuli on healthy young adults. Neurology, Psychiatry and Brain Research, 19(4): 207-215.

49. Mankel K, Comstock DC, Bormann BM, Das S, Sagiv D, et al. (2024) Auditory brainstem responses to speech-in-noise reflect selective attention, comprehension, and subjective listening effort. Neuroscience.

50. Matyjek M, Kita S, Torralba Cuello M, Soto Faraco S (2024) Multisensory integration of speech and gestures in a naturalistic paradigm. Human Brain Mapping 45(11): e26797.

51. Miendlarzewska EA, Trost WJ (2014) How musical training affects cognitive development: Rhythm, reward and other modulating variables. Frontiers in Neuroscience 7(1): 279.

52. Nieberlein L, Martin S, Williams KA, Gussew A, Cyriaks SD, et al. (2024) Semantic Integration Demands Modulate Large-Scale Network Interactions in the Brain. Human Brain Mapping 45(18): e70113.

53. Oh Y, Kalpin N, Hunter J, Schwalm M (2023) The impact of temporally coherent visual and vibrotactile cues on speech recognition in noise. JASA Express Letters 3(2): 025203.

54. Oh Y, Schwalm M, Kalpin N (2022) Multisensory benefits for speech recognition in noisy environments. Frontiers in Neuroscience 16(1): 1031424.

55. Okamoto K, Hoyano K, Saiki Y, NomuraT, Irie K, et al. (2024) Predictive brain activity related to auditory information is associated with performance in speech comprehension tasks in noisy environments. Frontiers in Human Neuroscience 18(1): 1479810.

56. O'Sullivan AE, Crosse MJ, Di Liberto GM, De Cheveigné A, Lalor EC, et al. (2020) Neurophysiological indices of audiovisual speech integration are enhanced at the phonetic level for speech in noise. Neuroscience.

57. Perron M, Liu Q, Tremblay P, Alain C (2024) Enhancing speech perception in noise through articulation. Annals of the New York Academy of Sciences 1537(1): 140-154.

58. Puschmann S, Baillet S, Zatorre RJ (2019) Musicians at the Cocktail Party: Neural Substrates of Musical Training During Selective Listening in Multispeaker Situations. Cerebral Cortex 29(8): 3253-3265.

59. Putkinen V, Saarikivi K, Chan TMV, Tervaniemi M (2021) Faster maturation of selective attention in musically trained children and adolescents: Converging behavioral and event-related potential evidence. European Journal of Neuroscience 54(1): 4246-4257.

60. Rao A, Rishiq D, Yu L, Zhang Y, Abrams H et al. (2017) Neural Correlates of Selective Attention with Hearing Aid Use Followed by Read My Quips Auditory Training Program. Ear & Hearing 38(1): 28-41.

61. Rauschecker, JP (2013) Processing Streams in Auditory Cortex. In Y. E. Cohen, A. N. Popper, & R. R. Fay (Eds.), Neural Correlates of Auditory Cognition (Vol. 45, pp. 7-43). Springer New York.

62. Rennig J, Beauchamp MS (2020) Linking Activity in Human Superior Temporal Cortex to Perception of Noisy Audiovisual Speech. Neuroscience

63. Saiz-Alía M, Forte AE, Reichenbach T (2019) Individual differences in the attentional modulation of the human auditory brainstem response to speech inform on speech-in-noise deficits. Scientific Reports 9(1): 14131.

64. Schumann A, Ross B (2022) Adaptive Syllable Training Improves Phoneme Identification in Older Listeners with and without Hearing Loss. Audiology Research 12(6): 653-673.

65. Shen J, Sun J, Zhang Z, Sun B, Li H, et al. (2024) The Effect of Hearing Loss and Working Memory Capacity on Context Use and Reliance on Context in Older Adults. Ear & Hearing 45(3): 787-800.

66. Shim H, Gibbs L, Rush K, Ham J, Kim S, et al. (2023) Neural Mechanisms Related to the Enhanced Auditory Selective Attention Following Neurofeedback Training: Focusing on Cortical Oscillations. Applied Sciences 13(14): 8499.

67. Skoe E, Kraus N (2024) Neural Delays in Processing Speech in Background Noise Minimized after Short-Term Auditory Training. Biology 13(7): 509.

68. Smith GE, Housen P, Yaffe K, Ruff R, Kennison RF, et al. (2009) A Cognitive Training Program Based on Principles of Brain Plasticity: Results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) Study. Journal of the American Geriatrics Society 57(4): 594-603.

69. Söderlund GBW, Åsberg Johnels J, Rothén B, Torstensson-Hultberg E, Magnusson A, et al. (2021) Sensory white noise improves reading skills and memory recall in children with reading disability. Brain and Behavior 11(7): e02114.

70. Song JH, Skoe E, Skoe E, Banai K (2011) Perception of speech in noise: Neural correlates. Journal of Cognitive Neuroscience 23(9): 2268-2279.

71. Song JH, Skoe E, Skoe E, Banai K (2012) Training to Improve Hearing Speech in Noise: Biological Mechanisms. Cerebral Cortex 22(5): 1180-1190.

72. Stokes RC, Venezia JH, Hickok G (2019) The motor system's [modest] contribution to speech perception. Psychonomic Bulletin & Review 26(4): 1354-1366.

73. Strait DL, Kraus N (2011) Can You Hear Me Now? Musical Training Shapes Functional Brain Networks for Selective Auditory Attention and Hearing Speech in Noise. Frontiers in Psychology 2(1): 113.

74. Sussman ES, Bregman AZ, Lee W-W (2014) Effects of task-switching on neural representations of ambiguous sound input. Neuropsychologia 64: 218-229.

75. Sweetow RW, Henderson Sabes J (2010) Auditory Training and Challenges Associated with Participation and Compliance. Journal of the American Academy of Audiology 21(9): 586-593.

76. Sweetow RW, Sabes JH (2006) The Need for and Development of an Adaptive Listening and Communication Enhancement (LACETM) Program. Journal of the American Academy of Audiology 17(8): 538–558.

77. Synigal SR, Anderson AJ, Lalor EC (2023) Electrophysiological indices of hierarchical speech processing differentially reflect the comprehension of speech in noise. Neuroscience.

78. Tervaniemi M, Kruck S, De Baene W, Schröger E, Alter K, et al. (2009) Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. European Journal of Neuroscience 30(8): 1636–1642.

79. Tremblay K (2007) Training-Related Changes in the Brain: Evidence from Human Auditory-Evoked Potentials. Seminars in Hearing 28(2): 120-132.

80. Tremblay K, Kraus N, McGee T, Ponton C, Otis AB, et al. (2001) Central Auditory Plasticity: Changes in the N1-P2 Complex after Speech-Sound Training: Ear and Hearing 22(2): 79-90.

81. Viswanathan V, Heinz MG, Shinn-Cunningham BG (2024) Impact of Reduced Spectral Resolution on Temporal-Coherence-Based Source Segregation. Neuroscience.

82. Voss P, Thomas ME, Cisneros-Franco JM, De Villers-Sidani É (2017) Dynamic Brains and the Changing Rules of Neuroplasticity: Implications for Learning and Recovery. Frontiers in Psychology 8(1): 1657.

83. Yu L, Rao A, Zhang Y, Burton PC, Rishiq D, et al. (2017) Neuromodulatory Effects of Auditory Training and Hearing Aid Use on Audiovisual Speech Perception in Elderly Individuals. Frontiers in Aging Neuroscience.

84. Yusof Y, Mukari SZS, Dzulkifli MA, Chellapan K, Ahmad K, et al. (2019) Efficacy of a newly developed auditory–cognitive training system on speech recognition, central auditory processing and cognitive ability

among older adults with normal cognition and with neurocognitive impairment. Geriatrics & Gerontology International 19(8): 768-773.

85. Zatorre R (2024) Communicating Between Auditory Regions and the Rest of the Brain: The Dorsal Stream. In R Zatorre, From Perception to Pleasure (1st ed., pp. 100–135). Oxford University Press New York.

86. Zendel BR (2022) The importance of the motor system in the development of music-based forms of auditory rehabilitation. Annals of the New York Academy of Sciences 1515(1): 10-19.



**DOI:** 10.32474/SJO.2025.10.000352