Hearing Research 466 (2025) 109367

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Neural tracking of continuous speech in adverse acoustic conditions among healthy adults with normal hearing and hearing loss: A systematic review

David Ratelle [a,b,c], Pascale Tremblay [a,b,c,*]

[a] CERVO Brain Research Center, 2301 Avenue D'Estimauville, Quebec City, QC G1J 2G3, Canada
[b] Université Laval, Faculté de Médecine, École des Sciences de la Réadaptation, Pavillon Ferdinand-Vandry, 1050 Avenue de la Médecine, Quebec City, QC G1V 0A6, Canada
[c] Centre for Research on Brain, Language and Music (CRBLM), McGill University, Faculty of Medicine, Rabinovitch House, 3640 Rue de la Montagne, Montréal, QC H3G 2A8, Canada

ARTICLE INFO

Keywords:
Speech perception
Hearing loss
Aging
Speech-in-noise
Electroencephalography (EEG)
Magnetoencephalography (MEG)
Speech tracking
Neural responses
Auditory processing

ABSTRACT

The study of neural speech tracking (NST) has gained increasing attention in the field of auditory neuroscience in recent years. However, its contribution to speech perception in noise (SPiN), especially regarding aging and hearing loss, has yet to be fully explored. This systematic review examined NST in adults with and without hearing loss, focusing on its modulation by age, hearing impairment, and adverse acoustic conditions, as well as its relationship with behavioral SPiN performance. A systematic literature search identified studies using electroencephalography (EEG) or magnetoencephalography (MEG) to investigate NST in continuous speech processing under adverse acoustic conditions. Studies included participants with and without hearing loss, excluding those with neurological disorders. Various NST methods, including forward and backward modeling, coherence, and cross-correlation, were examined. Fifty-four studies met the inclusion criteria. Most studies focused on young adults, with fewer studies including older adults or individuals with hearing loss. Findings suggest that older adults exhibit increased NST compared to younger adults, potentially reflecting compensatory mechanisms for auditory processing declines. Similarly, hearing impairment was generally associated with enhanced NST, likely due to altered neural encoding and increased reliance on cognitive resources. The impact of adverse acoustic conditions, as reflected by the signal-to-noise ratio (SNR), on NST was predominantly negative, with NST decreasing as noise levels increased. However, some studies suggested a non-linear relationship, with NST peaking at intermediate SNRs. Furthermore, most studies reported a positive correlation between NST and SPiN performance, typically observed across individuals or conditions within homogeneous groups or pooled samples. While stronger tracking was generally associated with better behavioral outcomes, this relationship does not imply that higher NST always corresponds to better performance across different populations. Age and hearing loss appear to modulate NST, likely through both neural compensation and auditory processing adaptations. The complex effects of SNR on NST highlight the need for additional research to better understand its underlying mechanisms. Future studies should delve deeper into the interplay between aging, cognition, and auditory deficits in shaping NST, offering a more comprehensive understanding of speech processing in challenging acoustic environments.

1. Introduction

The study of auditory speech perception has been a topic of interest in electroencephalography (EEG) research for over 50 years. In the early 70 s, research addressed subjects such as the effect of attention on the evoked response to speech sounds (Roth et al., 1970), the neural correlates of the lateralization of speech functions (Matsumiya et al., 1972; Morrell and Salamy, 1971), differences in neural responses between verbal and non-verbal stimuli (Buchsbaum and Fedio, 1970; Cohen, 1971; Greenberg and Graham, 1970), and the effect of verbal meaning on EEG evoked potentials (Brown et al., 1973). A common feature of all these studies is the use of brief speech material such as syllables or monosyllabic words, with a narrow EEG signal analysis window restricted to the milliseconds corresponding to the length of these

* Corresponding author.
E-mail addresses: david.ratelle.2@ulaval.ca (D. Ratelle), pascale.tremblay@fmed.ulaval.ca (P. Tremblay).

https://doi.org/10.1016/j.heares.2025.109367
Received 18 February 2025; Received in revised form 21 June 2025; Accepted 10 July 2025
Available online 11 July 2025
0378-5955/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

stimuli, and neural responses averaged over multiple repetitions of the stimulation. This technique, known as auditory event-related potentials, has been used since then and has allowed researchers to address various questions related to speech perception. However, this technique is not suited to capture the extended neural response to the complex and dynamic fluctuations in speech over time (Crosse, Di Liberto, et al., 2016) which include changes in both frequency and intensity, often referred to as spectro-temporal modulations.

In recent years, research on speech perception has begun to use continuous speech stimuli to characterize the neural impulse response in more ecological auditory scenarios as opposed to short speech stimuli (see a review by Brodbeck and Simon (2020)). Concurrently, new EEG and magnetoencephalography (MEG) response analysis techniques for continuous speech stimuli have emerged and are now well established in the scientific literature (see a review by Gillis, Van Canneyt, et al. (2022)). Recent neurophysiological evidence highlights that aging and hearing loss are not only associated with degraded temporal processing, but also with exaggerated cortical and subcortical envelope encoding (Herrmann and Butler, 2021), potentially due to neural disinhibition and central hyperactivity. These changes, which are particularly evident for low-frequency modulations, may contribute to altered speech perception in adverse conditions and warrant further exploration in continuous speech paradigms (McClaskey, 2024).

There is no consensus in the literature regarding the specific term to use when describing how the brain follows a long stream of speech, whether it is referred to as speech tracking, neural tracking of speech, phase coding, or neural entrainment to speech. Many of these terms are used interchangeably, but Obleser and Kayser (2019) emphasize distinctions between them. For example, neural entrainment refers specifically to the synchronization of neural oscillations with sensory inputs, while phase coding relates to the timing of neural activity relative to the phase of the local field potential. Speech tracking and neural tracking, as Obleser and Kayser (2019) point out, are broader terms that encompass both quantitative and qualitative relationships between brain activity and speech stimuli, and they can be measured using methods like phase or power entrainment. Accordingly, we have chosen to use the term *neural speech tracking* (NST) since it captures the broader process of interest in this review, by which brain activity aligns with continuous speech, especially under adverse acoustic conditions.

Speech perception can be challenging when it occurs in the presence of various acoustic sources, such as environmental noise (i.e., energetic masking) or competing speech, whether from a single or multiple talkers (i.e., informational and energetic masking in the case of multiple talkers) (Brungart, 2001; Brungart et al., 2001; Culling and Stone, 2017; Freyman et al., 2004, 2007). This latter situation is commonly referred to as the cocktail party problem (Cherry, 1953). In this review, we will use the acronym SPiN (Speech Perception in Noise) to refer to speech perception in the presence of any adverse acoustic signals. The location of the noise source relative to the target speech can further exacerbate listening difficulties to varying degrees (Arbogast et al., 2002; Freyman et al., 2001, 1999). Additionally, differences in decibel (dB) levels between the target speech and the masking sound, referred to as the signal-to-noise ratio (SNR) (see Wang and Xu (2021) for a review on masking), are critical for speech intelligibility and comprehension and play a key role in evaluating the brain's ability to track the speech signal in SPiN contexts.

SPiN difficulties present a significant challenge for many adults (Davis, 1989; Gilliver et al., 2013, 2015; Hannula et al., 2011; Tremblay et al., 2015). In addition to their negative psychological impacts, such as stress, anxiety, and depression, SPiN difficulties are also known to reduce social participation (Heine and Browning, 2002), which in turn leads to a higher risk of premature mortality, as shown in a meta-analysis by Holt-Lunstad et al. (2015). It is estimated that SPiN difficulties account for 5 to 15 % of all consultations in hearing-related health clinics (Pang et al., 2019).These difficulties can begin as early as the third decade of life (Jokinen, 1973) and are also observed in older adults without clinical hearing loss (Frisina and Frisina, 1997). As reported by Beck et al. (2018), approximately 12 to 15 % of adults with normal hearing thresholds have difficulty hearing and struggle to understand speech in noise. This clinically distinct condition is known in the audiological and ear-nose and throat domain as the King-Kopetzky syndrome (King, 1954; Kopetzky, 1948; Pryce, 2006), or Obscure Auditory Dysfunction (Saunders and Haggard, 1989, 1992). More recently, the term hidden hearing loss has emerged in the literature, referring to cochlear synaptopathy, where it is thought that a deafferentation occurs between inner hair cells and auditory nerve fibers. This condition is observed despite the presence of preserved outer hair cells, which produce normal hearing thresholds in pure tone audiometry but lead to significant speech-in-noise (SPiN) impairment (see a review by Liu et al. (2024)). SPiN difficulties are also present in the hearing-impaired population, generally characterized by audiometric thresholds elevated at or above 25 dB (CHABA, 1988; Helfer and Freyman, 2008; Marrufo-Perez and Lopez-Poveda, 2022; Quist-Hanssen et al., 1978; Shapiro et al., 1972). For this population, hearing aid amplification provides only limited benefits (Healy and Yoho, 2016; Lavie et al., 2014). Taken together, evidence suggests that age and hearing loss are not the only contributing factors to SPiN difficulties and that other neurological factors might also be involved.

Neurological correlates of SPiN have been investigated through various methods, including MRI studies that highlight the impact of cortical structures on SPiN performance. Tremblay et al. found that decreased integrity in the arcuate fasciculus (Perron et al., 2021; Tremblay et al., 2018) and cortical thinning in regions of the dorsal speech stream including the superior temporal cortex, ventral premotor cortex and inferior frontal gyrus is associated with poorer SPiN (Tremblay et al., 2021), while Wong et al. (2008) demonstrated that higher noise levels were associated with increased left posterior superior temporal gyrus activity, correlating with decreased SPiN performance. In contrast, heightened activity in the right hemisphere was associated with better performance. Wong et al. (2009) further noted that improved SPiN performance in older adults was linked to increased activation in frontal and parietal regions, suggesting compensatory mechanisms for aging. Du et al. (2016) observed improved SPiN with increased activation in the left inferior frontal gyrus and precentral gyrus. In intracranial EEG studies, Berezutskaya et al. (2020) reported that the dorsal precentral cortex was particularly sensitive to speech compared to a mix of speech, music, and noise while Glanz et al. (2018) reported that the left ventral premotor cortex showed consistent activation during speech perception, regardless of whether the speech was presented in noise or in clear conditions. Finally, Mesgarani and Chang (2012) showed enhanced encoding of temporal and spectral aspects of attended speech in the superior and middle temporal gyri during a two-speaker task, and Nourski et al. (2024) linked better performance in degraded speech tasks to increased activity in the posterior superior temporal and precentral gyri. Taken together, these studies suggest a neurological contribution to SPiN difficulties, especially in dorsal speech stream regions. Although these studies provide valuable insights into the neural mechanisms of SPiN, it is important to note that most of them, apart from Glanz et al. (2018), used short speech stimuli. This limits their ability to fully capture the complexity of real-world speech perception, where speech is typically heard in a continuous stream. Therefore, the applicability of these findings to most everyday listening situations remains uncertain.

There is a growing interest in the literature in using various speech tracking techniques to study central auditory processing. However, to our knowledge, no review has yet explored how these methods can improve our understanding of SPiN difficulties in adults and older individuals, particularly in relation to age-related changes, hearing loss, the impact of SNR on speech tracking, and the connection to behavioral data. In this review, we focused on studies that utilized continuous speech stimuli and included at least one comparative adverse acoustic condition. Given the similarities between EEG and MEG data and the

fact that both can be analyzed using similar methods, we have chosen to include both techniques in this review. Young normal-hearing participants, who are typically considered a reference group in comparative studies (Presacco et al., 2019), were included to provide a baseline for comparison with older adults with and without hearing loss, in order to highlight differences in NST across these populations.

The primary objective of this review was to provide an analysis of the literature on continuous speech tracking in adverse acoustic conditions among healthy adults with normal hearing or hearing loss. First, we aimed to descriptively evaluate the literature from a methodological standpoint. Second, we sought to assess the influence of adverse acoustic conditions on speech tracking, considering the various analysis techniques used, the populations studied, and the relationship between speech tracking and SPiN performances. Specifically, age (younger vs. older adults) and hearing status (normal hearing vs. hearing impaired) were considered as key variables of interest. Finally, we aimed to evaluate the contribution of different speech tracking methods by analyzing their convergence in capturing neural processes underlying continuous speech processing in adverse acoustic conditions. Through this approach, we aimed to deepen the understanding of the neurological processes involved at the macroscopic level in continuous speech analysis under adverse acoustic conditions, and to provide insights to guide future research.

2. Method

2.1. Search Strategy

We carried out a search of online databases on January 14, 2024, specifically PubMed and PsycNet, aiming to identify peer-reviewed studies that utilize EEG or MEG in the context of SPiN tasks, with a particular focus on NST to continuous speech methodologies. Intracortical EEG studies were not included because of our focus on neurologically healthy adults. The search strategy was designed to encompass a broad range of relevant studies while maintaining specificity. Keywords were selected based on their relevance to EEG or MEG, NST methodologies, and SPiN conditions:

PubMed

("Electroencephalography"[MeSH Terms] OR "EEG"[All Fields] OR "Magnetoencephalography"[MeSH Terms] OR "MEG"[All Fields]) AND ("Speech Perception"[MeSH Terms] OR "Auditory Perception"[MeSH Terms] OR "Speech"[MeSH Terms] OR "speech tracking"[All Fields] OR "neural tracking"[All Fields] OR "speech envelope tracking"[All Fields] OR "phase coding"[All Fields] OR "neural entrainment"[All Fields] OR "neural speech tracking"[All Fields] OR "FFR"[All Fields]) AND ("Noise"[MeSH Terms] OR "noisy environments"[All Fields] OR "white noise"[All Fields] OR "speech babble"[All Fields])

PsycNET

("Electroencephalography" OR "EEG" OR "Magnetoencephalography" OR "MEG") AND ("Speech Perception" OR "Auditory Perception" OR "Speech" OR "speech tracking" OR "neural tracking" OR "speech envelope tracking" OR "phase coding" OR "neural entrainment" OR "neural speech tracking" OR "FFR") AND ("Noise" OR "noisy environments" OR "white noise" OR "speech babble")

2.2. Selection Process

A preselection on the title and the abstract of all retrieved articles was performed independently by the two authors. The preselected articles were then independently assessed on the full text by the two authors. After full text assessment, the references of the selected articles were screened for additional articles, which were then independently assessed via the same preselection and selection processes. A web-based collaboration software platform that streamlines the production of systematic and other literature reviews was used for this process (Covidence).

2.3. Selection Criteria

To be eligible for this review, studies were required to be peer-reviewed, quantitative in nature, and published in English. We imposed no restrictions on the publication date. The participants in these studies had to be neurologically healthy adults, including both younger and older individuals (i.e., without neurological, psychological, psychiatric, cognitive, or communication disorders), with normal hearing or hearing loss. Participants using hearing aids were included, but cochlear implant users were excluded. We selected EEG and MEG studies that analyzed NST in SPiN tasks using continuous speech stimuli (at least at the sentence level) with either male or female speakers. Eligible studies could include speech in the presence of spatially separated maskers, energetic maskers (e.g., white noise), or informational maskers (e.g., speech babble) employed at different SNR, specifically above 0 dB, at 0 dB, and below 0 dB. However, these studies were required to include natural human speech (i.e., non-synthetic speech) and at least one comparative condition (e.g., speech in quiet versus speech in noise, or speech in noise at one SNR level compared to another). Single case reports, case series, multiple case studies, and studies focused on technical, or engineering aspects were excluded.

2.4. Data Extraction

Data extraction was done independently by the first author and verified by a member of the laboratory or the second author. Discrepancies were solved by consensus. The following information was extracted for each study.

- **Group characteristics:** This included the total number of participants, their sex distribution, age range, and handedness. Additional details such as neurological health, cognitive status, and hearing status were also systematically extracted.
- **Test conditions, stimuli, and SPiN tasks:** Details on the mode of presentation, transducer type, presentation level, and whether soundproof booths were used were recorded. Characteristics of the target stimuli, their duration, the number of presentations, and repetitions were noted for both quiet and adverse acoustic conditions. The type of adverse conditions (e.g., competing talkers, multitalker babble, noise, music), the use of spatialization techniques, and specific SNR values or individualized presentation levels were also documented. SPiN tasks were identified based on whether they were used to maintain participant engagement or to serve as a correlation measure.
- **EEG and MEG NST methods and extracted features of the stimuli:** The methods used for NST were recorded, including their specific applications and the neurophysiological data they analyzed. Features of the speech signal used in NST analysis were documented. The use of model-informed filtering methods, such as those simulating cochlear or subcortical processing, was also noted.
- **Results from neural speech tracking methods:** This included data on how age, hearing status, and SNR influenced NST. Associations between NST and behavioral data were also systematically reviewed.

3. Results

3.1. Study Selection

Online literature searches identified a total of 1102 studies matching the search terms: 450 in PubMed and 652 in PsycNet. After removing duplicates, 908 unique records remained, with an additional 11 articles identified from the bibliographies of the final selection of articles. Titles and abstracts were screened, and records not meeting the inclusion criteria were discarded ($N = 816$). All disagreements were resolved by consensus between the two authors. A full-text evaluation of the remaining 103 articles led to the exclusion of 49 articles that did not

meet the inclusion criteria for the systematic review, leaving 54 articles. In total, the systematic review included 54 studies conducted between 2012 and 2024 (see Fig. 1). See Table 1 for a complete list of all included studies.

3.2. Group characteristics

3.2.1. Age, Number of Participants, Sex, and Laterality

Among all studies, 34 focused on young adults (range: 18–40 years old), 7 included both young and older adults (range: 20–82 years old), 7 compared young and older adults in separate groups (range: 18–86 years old, with one of these study including a middle-aged group (range: 39–60 years old), and 6 focused exclusively on older adults (range: 51–88 years old) (see Fig. 2.A). By summing the number of participants reported across all studies, a total of 1276 presumed unique participants were included (excluding participant samples enrolled in multiple studies). However, the sex of participants was not specified in two studies, resulting in a count of 56.89 % female among the 52 studies that reported this information (see Fig. 2B). Regarding handedness, 25 studies did not report participants' handedness, 21 included only right-handed participants and 8 included both right- and left-handed participants (with left-handed individuals being the minority).

3.2.2. Neurological Health and Cognitive Status

Neurological health information was absent in 26 studies, while 28 studies explicitly reported no neurological disorders. Of those, 14 were based on self-reporting. Similarly, speech or language disorders were not mentioned in 49 studies, while 5 studies mentioned the absence of such disorders, with 2 relying on self-reports. In terms of cognitive status, 40 studies lacked information on this criterion, while 14 explicitly mentioned no cognitive disorders, 2 of which were based on self-report (see Fig. 2C). A total of 11 studies used at least one cognitive screening test including the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) (cutoff score = 26: $n = 7$; cutoff score = 22: $n = 3$) or the Wechsler Abbreviated Scale of Intelligence (Zhu and Garcia, 1999) (cutoff score = 85: $n = 4$). Additionally, 12 studies included one or more cognitive measurements after meeting the inclusion criteria such as the Flanker Inhibitory Control and Attention Test (Eriksen and Eriksen, 1974; Weintraub et al., 2013) ($n = 5$), the Flemish version of the Reading Span Test (Van den Noort et al., 2008) ($n = 3$), an auditory or visual Stroop (Cohen and Martin, 1975; Hammes, 1978) ($n = 2$), a Digit Span Test (Blackburn and Benton, 1957; Drozdick et al., 2012) ($n = 4$), a sentence span task (Lewandowsky et al., 2010) ($n = 1$), the Conners Continuous Auditory Test of Attention (Conners, 2014) ($n = 1$), a visual scanning and divided attention tasks (Zimmermann and Fimm, 2002) ($n = 1$), a tonal sequence manipulation task (Albouy et al., 2017; Foster et al., 2013) ($n = 1$) and the Word Reading Test (Eén-Minuut Test) (Brus and Voeten, 1973) ($n = 1$).

3.2.3. Hearing Status

Regarding hearing status, 40 studies included only normal-hearing participants, 11 included both normal-hearing (NH) and hearing-impaired (HI) participants, and 1 study focused exclusively on HI participants. No information on hearing status was available in 2 studies (see Fig. 2D). Hearing was assessed by at least one method in 31 studies: 29 used pure-tone air conduction audiometry, with 4 of these also including bone conduction audiometry. One study used the Five Minutes Hearing Test (Koike et al., 1994), and another employed the Oldenburger Satztest, Hörtech, Germany (Wagener et al., 1999). In contrast, normal hearing was self-reported in 12 studies. A total of 6 studies explicitly compared NH and HI groups: in 5 of these, hearing impairment was determined using pure-tone audiometry criteria, while in one study it was based on SPiN difficulties. Additionally, 5 studies mentioned that HI participants used hearing aids. A clear definition of normal hearing, including threshold limits for specific frequencies, was provided in 20 studies, and 9 studies took interaural hearing threshold asymmetry into account.

3.3. Test Conditions, Stimuli, and SPiN Tasks

3.3.1. Test Conditions

The mode of presentation (i.e., monaural or binaural stimulation) was reported in most studies ($n = 52$). Binaural stimulation was employed in 46 studies, monaural stimulation in 4, and a combination of both in 2 studies. Transducers included earphones or headphones in 37 studies, free-field presentation in 8 studies, and a combination of both methods in 1 study. Hearing aids were used as transducers in 1 study, while 7 studies did not specify the transducer type. The presentation level (i.e., sound intensity in decibels) was specified in 41 studies: 25 studies used a fixed intensity, 13 studies applied a participant-adjusted comfortable level, and 6 studies adjusted the presentation level to account for hearing loss using frequency-specific amplification. Among

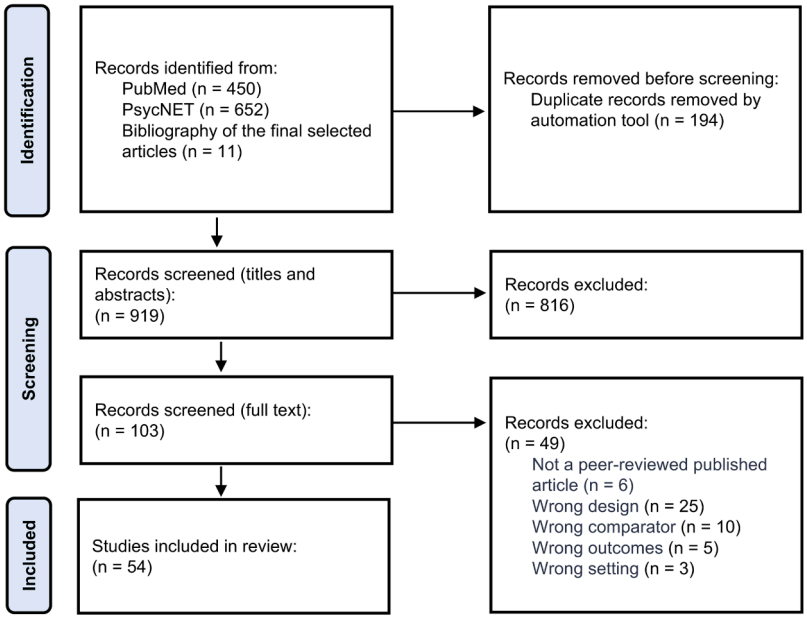

Fig. 1. Flow diagram illustrating the process of identification, screening, eligibility assessment, and inclusion of studies in the review.

Table 1
List of all studies included in the review, detailing participant and methodological key information.

Studies	Technique	N	Age	Hearing Status	Target Stimuli	Distractor Stimuli	NST Method
Brodbeck et al. (2018)	MEG	26	Y	NH	Audiobook	Competing Talker	FWD
Brodbeck et al. (2020)	MEG	26	Y	NH	Audiobook	Competing Talker	FWD
Brown and Bidelman (2022)	EEG	17	Y	NH	Audiobook	Music	FWD
Crosse, Di Liberto and Lalor (2016)	EEG	42	Y	NH	Video Speech	Noise	BWD
Das et al. (2018)	EEG	28	Y	NH	Audiobook	Competing Talker Multitalker Babble	BWD
Decruy et al. (2019)	EEG	54	Y/M/O	NH	Sentences Audiobook	Competing Talker Noise	BWD
Decruy, Lesenfants, et al. (2020)	EEG	13	Y	NH	Sentences	Noise	FWD
Decruy, Vanthornhout and Francart (2020)	EEG	28	Y + O	NH/HI	Sentences Audiobook	Competing Talker Noise	BWD
Desai et al. (2021)	EEG	17	Y	NH	Sentences Video Speech	Competing Talker Music Noise	FWD
Destoky et al. (2019)	MEG EEG	10	Y	NH	Audiobook	Multitalker Babble	BWD CHR
Ding and Simon (2012a)	MEG	20	Y	NH	Audiobook	Competing Talker	FWD BWD
Ding and Simon (2012b)	MEG	10	Y	NH	Audiobook	Competing Talker	FWD BWD
Ding and Simon (2013)	MEG	10	Y	NH	Audiobook	Noise	FWD BWD
Ding et al. (2014)	MEG	12	Y	NH	Audiobook	Noise	FWD
Etard and Reichenbach (2019)	EEG	12	Y	NH	Audiobook	Multitalker Babble	FWD BWD
Fiedler et al. (2019)	EEG	18	Y + O	NH	Audiobook	Competing Talker	FWD
Fuglsang et al. (2017)	EEG	26	Y	NH	Audiobook	Competing Talker Multitalker Babble	FWD BWD
Fuglsang et al. (2020)	EEG	44	O	NH/HI	Audiobook	Competing Talker	FWD BWD
Gillis, Decruy, et al. (2022)	EEG	28	Y + O	NH/HI	Audiobook	Noise	FWD
Hambrook and Tata (2019)	EEG	17	Y	NH	Sentences	Competing Talker	C-C
Har-Shai Yahav and Zion Golumbic (2021)	MEG	30	Y	n/a	Podcast Audiobook	Competing Talker	FWD
Har-Shai Yahav et al. (2024)	MEG	33	Y	n/a	Podcast	Competing Talker	FWD BWD
Hausfeld et al. (2021)	EEG	20	Y	NH	Audiobook	Competing Talker	FWD
Karunathilake et al. (2023)	MEG	34	Y/O	NH	Audiobook	Competing Talker Multitalker Babble	FWD BWD
Keitel et al. (2018)	MEG	20	Y	NH	Sentences	Noise	MI
Kong et al. (2014)	EEG	8	Y	NH	Audiobook	Competing Talker	C-C
Kurthen et al. (2021)	EEG	23	O	NH + HI	Audiobook	Competing Talker Multitalker Babble	FWD
Lesenfants et al. (2019)	EEG	19	Y	NH	Sentences Audiobook	Noise	FWD
McHaney et al. (2021)	EEG	19	O	NH + HI	Audiobook	Competing Talker	FWD
Mirkovic et al. (2019)	EEG	31	O	NH/HI	Audiobooks	Multitalker Babble	C-C
Mohammadi et al. (2023)	EEG	32	Y	NH	Sentences	Noise	PLV
Muncke et al. (2022)	EEG	18	Y + O	NH	Sentences	Competing Talker	FWD
Panela et al. (2024)	EEG	42	Y/O	NH + HI	Audiobook Podcast	Multitalker Babble	FWD
Petersen et al. (2017)	EEG	27	O	NH + HI	Audiobook	Competing Talker	C-C
Presacco et al. (2016a)	MEG	32	Y/O	NH	Audiobook	Competing Talker	BWD
Presacco et al. (2016b)	MEG	32	Y/O	NH	Audiobook	Competing Talker	BWD
Presacco et al. (2019)	MEG	46	Y/O	NH/HI	Audiobook	Competing Talker	BWD
Puschmann et al. (2019)	MEG	20	Y	NH	Audiobook	Competing Talker	BWD
Schmitt et al. (2022)	EEG	101	O	NH + HI	Sentences	Multitalker Babble Noise	C-C
Schubert et al. (2023)	MEG	49	Y	NH	Audiobook	Competing Talker	FWD
Synigal et al. (2020)	EEG	31	Y	NH	Audiobook	Competing Talker	FWD BWD
Vander Ghinst et al. (2016)	MEG	20	Y	NH	Audiobook	Multitalker Babble	CHR
Vander Ghinst et al. (2019)	MEG	20	Y	NH	Audiobook	Multitalker Babble	CHR
Vander Ghinst et al. (2021)	MEG	26	Y	NH/HI	Audiobook	Multitalker Babble	CHR
Vanthornhout et al. (2018)	EEG	24	Y	NH	Sentences	Noise	BWD
Verschueren et al. (2020)	EEG	19	Y	NH	Sentences Audiobook	Noise	FWD BWD
Wang et al. (2020)	EEG	20	Y	NH	Audiobook	Competing Talker	FWD BWD
Wang et al. (2023)	EEG	19	Y + O	HI	Audiobook	Noise	FWD BWD
Yasmin et al. (2023)	EEG	39	Y	NH	Podcast	Multitalker Babble	FWD
Zan et al. (2020)	MEG	32	Y/O	NH	Audiobook	Competing Talker	MI

(*continued on next page*)

Table 1 (*continued*)

Studies	Technique	N	Age	Hearing Status	Target Stimuli	Distractor Stimuli	NST Method
Zhang et al. (2023)	EEG	29	Y	NH	Audiobook	Noise	FWD BWD
Zinszer et al. (2022)	EEG	24	Y	NH	Audiobook	Noise	FWD
Zoefel and VanRullen (2016)	EEG	12	Y	NH	Audiobook	Noise	C-C
Zou et al. (2019)	EEG	32	Y	NH	Audiobook	Noise	FWD BWD

Note. The table reports the number of participants included in each NST study after exclusions (N), along with the associated categories for age group (Y = young adults, M = middle-aged adults, O = older adults, Y/O = young and older adults in separate groups, Y + O = young and older adults combined in one group) and hearing status (NH = normal hearing, HI = hearing-impaired, NH/HI = normal hearing and hearing-impaired in separate groups, NH+HI = normal hearing and hearing-impaired combined in one group). The "Audiobook" category refers to stories or narrative speech materials, while "Video Speech" refers to speech extracted from audiovisual content. The table also indicates the neurophysiological technique used (EEG = electroencephalography, MEG = magnetoencephalography), the NST method applied (FWD = forward modeling, BWD = backward modeling, C-C = cross-correlation, CHR = coherence, PLV = phase locking value, MI = mutual information). An extended version of this table including additional methodological details (the extracted speech features for NST analysis) is available in Supplementary Table S1.

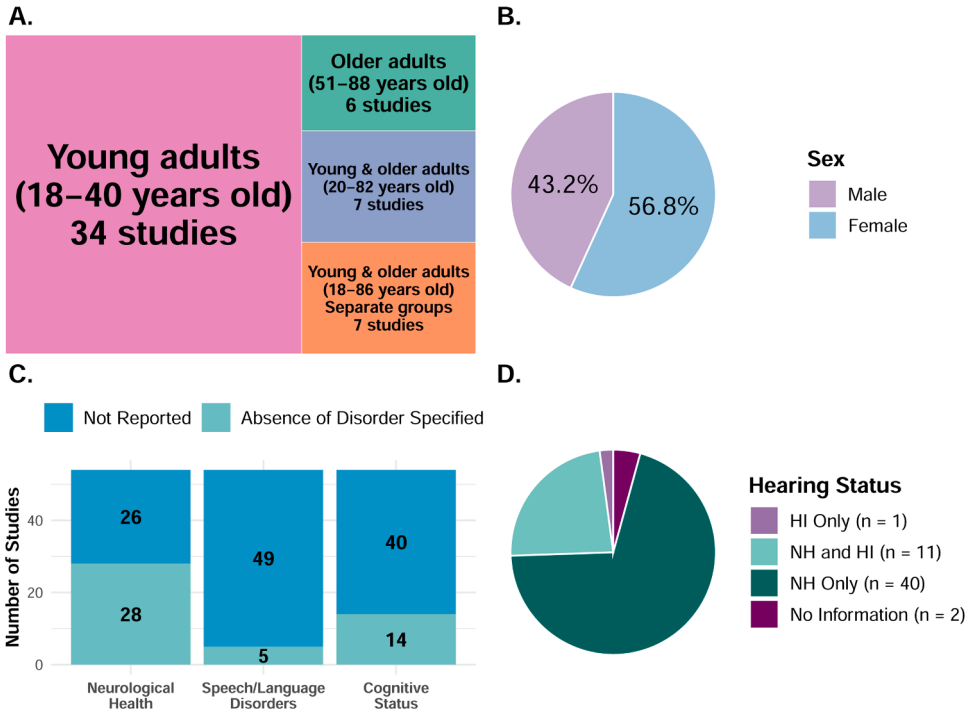

Fig. 2. Description of the participants. **(A)** Distribution of the age groups targeted in the studies included in the review. **(B)** Proportion of male and female participants based on available data from 52 studies. **(C)** Reporting of cognitive status, neurological health, and speech/language disorders across the 54 studies. The category 'Absence of Disorder Specified' indicates that the authors explicitly stated the absence of disorders among participants, including cases where this absence was self-reported. The category 'Not Reported' signifies that no information was provided regarding the presence or absence of disorders. **(D)** Hearing status of participants across the 54 studies. NH refers to normal-hearing participants, while HI refers to hearing-impaired participants. This includes both self-reported and measured hearing levels.

these, 3 studies reported different intensity levels for participants with normal hearing and those with hearing impairments. The presence of a soundproof booth was reported in 26 studies, while electromagnetic shielding for neurophysiological experiments (i.e., a Faraday cage) was mentioned in 33 studies.

3.3.2. *Target Stimuli*

The target continuous speech stimuli consisted of audiobooks (including stories and other narrative-type speech materials, which we grouped under the term 'Audiobook' due to their shared characteristics such as clear diction, well-structured narration, and consistent speech delivery) in 43 studies, isolated sentences in 12 studies, podcasts in 4 studies, and speech from video material (dialogue from a children's movie and conversational speech from a video featuring a man speaking) which we categorized as 'Video Speech,' in 2 studies. In 7 studies, two different types of speech material were used. A total of 40 studies included target speech stimuli presented in a quiet condition (i.e., in the absence of energetic or informational noise), and 14 studies did not have a quiet condition. Among these 40 studies, 37 provided information on duration of individual trials, while information was partial or absent in 2 studies. The duration of single-trial stimuli in quiet condition ranged from 10 s to 14 min (m = 2.67 min; med = 1.02 min) for audiobooks (n = 33), from 1.5 s to 12.35 s (m = 3.84 s; med = 3 s) for sentences (n = 6), from 30 s to 7 min (m = 3.75 min; med = 3.75) for podcasts (n = 2), and the data from only one study using video stimuli was available where the duration was 1 min. The number of presentations of target stimuli in quiet condition was reported in 32 studies but partially reported or absent in 8 studies. Fourteen studies included repetitions of the same speech material in their experiment where the number of repetitions ranged from 1 to 10 repetitions (m = 2.96; med = 2). The number of different presentations of stimuli in quiet condition ranged from 1 to 120 (m = 10.66; med = 4) for audiobooks (n = 30),

from 20 to 506 (m = 133.71; med = 40) for sentences (n = 7), and the data from only one study using podcasts and only one video stimuli was available where the number of presentations was 2 and 15 respectively. Fig. 3A illustrates the association between target and distractor types across all studies in this review.

3.3.3. Adverse Acoustic Conditions

As adverse acoustic condition, 29 studies used a competitive talker as the distractor (one or two competing talkers; otherwise, it was categorized as multi-talker speech babble), typically using the same speech material as the target (e.g., audiobooks, podcast, sentences). In 19 studies noise was used (speech-weighted noise: n = 15; environmental noise: n = 1; movie soundtrack: n = 1; phase-specific noise: n = 1; pink noise: n = 1). Fourteen studies used multi-talker speech babble; the number of talkers was specified in 12 studies and ranged from 3 to 12 (m = 6.18; med = 6). Music was used as a distractor in 2 studies. Two different types of distractors were used in 8 studies (vocoded or synthetic speech was excluded if present). In 9 studies, listening conditions included factors influencing the spatial presentation or perception of sounds, such as physical separation of the sound source (n = 3), presenting the target and masker to opposite ears (n = 4), virtual separation using head-related transfer functions (n = 2), interaural time differences (ITD) (n = 1), and the use of directional or omnidirectional microphone settings in hearing aids (n = 1). Specific SNRs values were found in 34 studies where the number of different SNR levels tested ranged from 1 to 9 (m = 3.35; med = 3) and SNR levels ranged from −12.5 to 12 dB SNR. In 9 studies individualized SNR values expressed in terms of speech reception threshold (SRT) or speech understanding percentage (SU %) were used where the number of conditions ranged from 1 to 6 (m = 3.44; med = 4) and the estimated SU % ranged from 6 % to 95 %. In 2 studies, both SNR and SU % values were used to define the level of difficulty in each condition. In 13 studies, no SNR or was specified. The duration of trials was specified in 52 studies, with partial or missing information in 2 studies. For audiobooks, the duration of single trial ranged from 10 s to 3 min (m = 87.5 s; med = 75 s) in noise (n = 11), from 16.69 s to 10 min (m = 3.77 min; med = 4.08 min) in multi-talker speech babble (n = 10), from 10 s to 4 min (m = 78.6 s; med = 78.6 s) in competing talker (n = 25) and was 2 min long in music (n = 1). For sentences, the duration of single trial ranged from 1.5 s to 10.34 s (m = 4.13 s; med = 3 s) in noise (n = 7), from 1.93 s to 3.6 s (m = 2.51 s; med = 2 s) in competing talker (n = 3), and was 10.34 s in multi-talker speech babble (n = 1). For podcasts, the duration of single trial ranged from 20.74 s to 44.53 s (m = 32.63 s) in competing talker (n = 2), and from 31.5 s to 7 min (m = 3.76 min) in multi-talker speech babble (n = 2). The number of presentations was specified in 46 studies but partially reported or absent in 8 studies. For audiobooks, the number of presentations (without repetition) ranged from 2 to 240 (m = 42.27; med = 9.5) in noise (n = 11), from 3 to 78 (m = 17; med = 8) in multi-talker babble (n = 10), from 3 to 360 (m = 36.64; med = 15) in competing talker paradigm (n = 25) and was 16 in music (n = 1). For sentences the number of presentations ranged from 30 to 680 (m = 218.8; med = 190) in noise (n = 8), from 10 to 672 (m = 215.5; med = 90) in competing talker paradigm (n = 4), and was 30 in multi-talker babble (n = 1). For podcast, the number of presentations ranged from 3 to 20 (m = 11.5) in multi-talker babble (n = 2), and from 15 to 72 (m = 43.5) in competing talker paradigm (n = 2). In 18 studies the same speech material in adverse acoustic condition was repeated where the number of repetitions of a single trail ranged from 1 to 14 (m = 2.83; med = 2).

3.3.4. SPiN Tasks

Although the terminology varies across studies, “speech intelligibility” typically refers to the recognition of individual words or short sentences, whereas “speech comprehension” or “understanding” generally involves answering questions about the meaning or semantic content of the speech material. In this review, 39 studies included tasks related to speech comprehension or understanding, while 29 studies focused on speech intelligibility or clarity (i.e., subjective ratings, number of words correctly reported). A word detection, sound detection, or word counting task was used in 9 studies. These tasks were designed to maintain participants’ attention on the stimuli or provide a

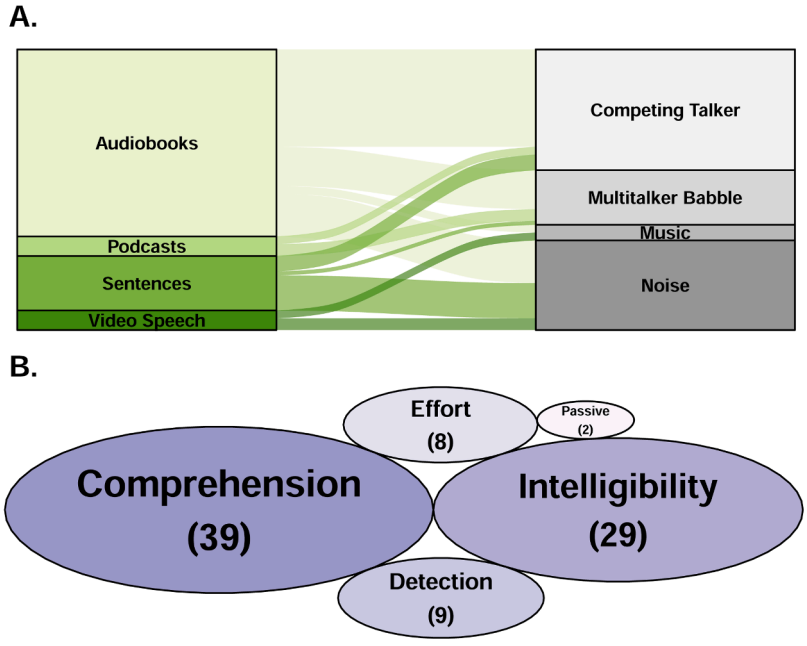

Fig. 3. **(A)** Association between target type (left) and distractor type (right) across all experiments in all the studies in this review. Speech from video material (Video Speech) was associated with noise (n = 2) and music (n = 1). Sentences were associated with noise (n = 9), multitalker babble (n = 1), and a competing talker (n = 4). Audiobooks were associated with noise (n = 11), multitalker babble (n = 10), a competing talker (n = 25), and music (n = 1). Podcasts were associated with multitalker babble (n = 2) and a competing talker (n = 2). **(B)** Distribution of different types of SPiN tasks performed in NST experiments across the studies in this review. "Comprehension" tasks are sometimes referred to as speech understanding tasks in certain studies. "Intelligibility" encompasses tasks assessing speech clarity, while “Effort” refers to listening effort tasks. “Detection” includes speech, word, or sound detection tasks, and “Passive” corresponds to passive listening tasks.

quantitative measure for correlation analysis with NST (see Section 3.5.4 for details on tasks analyzed in relation with the NST). Listening effort or difficulty ratings were assessed in 8 studies, and 2 studies employed a passive listening paradigm (see Fig. 3B). In total, 25 studies featured two tasks, with the most common combination being comprehension and intelligibility tasks ($n = 20$), and 3 studies included three tasks: comprehension, intelligibility, and listening effort or detection related tasks. Among all the studies, only 16 explicitly specified the exact response mode (i.e., verbal response or keyboard input). Additionally, 10 studies used the Matrix Sentences test (e.g., Luts et al. (2014); Wagener (1998)), 5 used the Quick Speech-in-Noise test (Killion et al., 2004), and 3 used the Hearing in Noise Test (e.g., Hällgren et al. (2006); Nielsen and Dau (2009); Nilsson et al. (1994)). As these tests involve word or sentence recognition, they were categorized as speech intelligibility tasks in our analysis.

3.4. NST Methods and extracted features of the stimuli

3.4.1. NST Methods

A total of 35 studies using EEG and 20 studies using MEG met our inclusion criteria for NST analysis with continuous speech in noise, with one study employing both EEG and MEG. We identified six NST methods: forward modeling, backward modeling, cross-correlation, coherence, phase locking value, and mutual information. Forward modeling was used in 32 studies. This method typically relies on temporal response functions (TRFs) to predict the neurophysiological response (e.g., EEG or MEG) based on the spectro-temporal characteristics of the speech stimulus (Crosse, Di Liberto, et al., 2016; Crosse et al., 2021). The accuracy of this prediction is typically assessed by quantifying how well the predicted neural response aligns with the recorded EEG or MEG signals. Various metrics can be derived from the TRF, including peak amplitude, power, and latency, which quantify the strength and temporal dynamics of NST. Backward modeling, applied in 24 studies, involves reconstructing the speech stimulus (e.g., the speech envelope) from the neurophysiological response (Crosse, Di Liberto, et al., 2016; Crosse et al., 2021). The strength of NST in this approach is reflected in the reconstruction accuracy, which quantifies how well the reconstructed stimulus correlate with the original speech signal. Forward and backward modeling were used together in 14 studies. Cross-correlation, employed in 6 studies, estimates the similarity between two time-series, such as the speech envelope and the neural responses (e.g., EEG), as a function of temporal lag (e.g., Schmitt et al. (2022)). While primarily used to quantify NST strength via a correlation coefficient, this method can also estimate neural response latency by identifying the time lag at which the correlation is maximal. Coherence, used in 4 studies, measures the coupling strength between the speech signal and the neural response within specific frequency bands (Halliday et al., 1995). Higher coherence values indicate stronger NST at specific frequencies. Phase locking value was applied in 2 studies to quantify the consistency of phase alignment between two signals (Lachaux et al., 1999), such as the speech envelope and neural oscillations, across different frequencies and time points (e.g., Mohammadi et al. (2023)). Higher phase locking values indicate stronger phase synchronization, reflecting a stronger NST. Mutual information, used in 2 studies, captures the extent to which two events are mutually constraining (e.g., Keitel et al. (2018)). Higher mutual information values indicate stronger statistical dependency between the speech signal and neural responses, reflecting a stronger NST. See Fig. 4A for a summary of the NST methods used in this review.

3.4.2. Extracted Features of the Stimuli

To assess NST through the methods outlined above, specific features of the speech signal, such as the envelope, the spectrogram, or the timing of distinct linguistic units, must be extracted to enable meaningful comparisons with neurophysiological responses. The speech envelope represents the amplitude modulations over time, with peaks typically corresponding to prominent linguistic features, such as syllabic rhythm and word onsets. The vast majority of studies included in this review used the amplitude envelope of the speech signal as the primary acoustic predictor for NST analyses. Specifically, 53 studies extracted the envelope of the target speech. Of these, 28 also extracted the envelope of the distractor signal, and 11 extracted the combined envelope of the target and distractor signals. Notably, 5 studies extracted both the target envelope and the combined target-distractor envelope, but not the distractor envelope alone. The temporal alignment of distinct linguistic units (e.g., phrases, words, syllables, vowels, consonants, phonemes) was extracted for speech tracking analysis in 5 studies. The spectrogram of a speech signal represents the distribution of acoustic energy across frequencies over time. By averaging energy across frequency bands, the spectrogram can be used to derive a broadband amplitude envelope, though this term also applies to envelopes extracted directly from the waveform using, for example, Hilbert transform. Alternatively, the

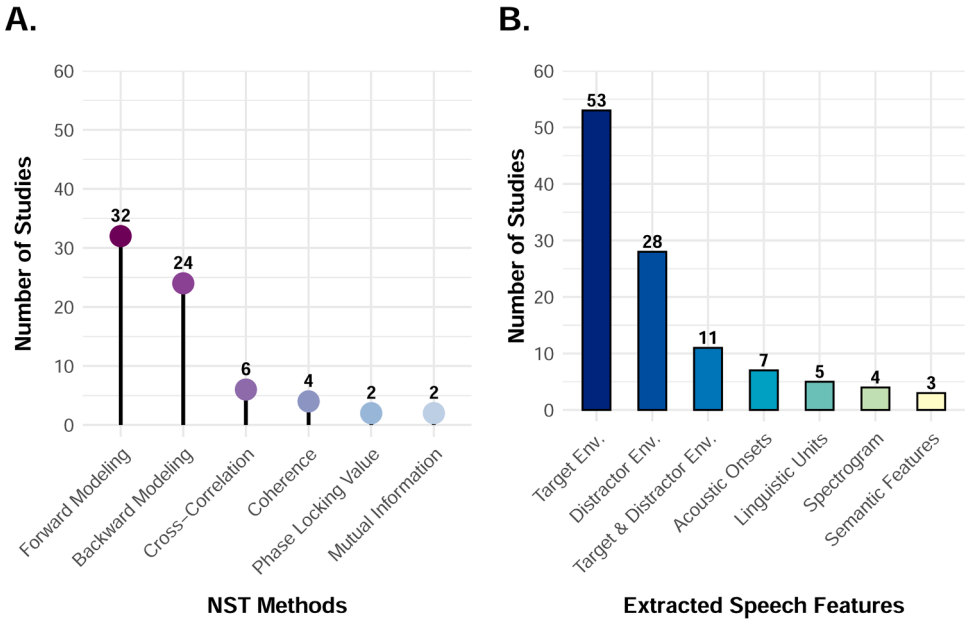

Fig. 4. **(A)** Distribution of NST analysis methods across the studies included in the review. **(B)** Speech features extracted for NST analysis across the 54 studies.

spectrogram itself can be used directly as a multiband predictor to compute NST. Four of the studies included in this review computed speech tracking directly from the spectrogram, enabling the model to track frequency-specific temporal modulations rather than relying on a single amplitude envelope. In 7 studies, acoustic onset representations were derived from either the speech envelope or the spectrogram, typically by rectifying the temporal derivative to isolate positive transitions, or by applying auditory models specifically designed to emphasize rapid increases in acoustic energy. Additionally, semantic features including word entropy, word surprisal, semantic violation, and semantic dissimilarity were incorporated in 3 studies (see Fig. 4B). Extracting the speech envelope generally involves low pass filtering the speech signal, while extracting the spectrogram often captures a broader range of frequencies. This extraction process aligns with established models of cochlear and subcortical auditory processing (e.g., Biesmans et al. (2017); Glasberg and Moore (1990); Irino and Patterson (2006); Patterson et al. (1988); Patterson et al. (1995); Yang et al. (1992)). This model-informed filtering was applied in 29 studies in this review. See Supplementary Table S1 for a detailed overview of the speech features extracted in each study.

3.5. Results from NST Methods

3.5.1. Age-Related Differences in NST

In this review, 14 studies included samples with an age range broad enough to examine age-related differences in NST. Of these, 7 studies included both young and older adults in a single group, 6 studies compared young and older adults in separate groups, and one study compared young, middle-aged, and older adults across three distinct groups. However, only 9 of these studies conducted specific analyses to investigate age-related differences in NST.

Five studies used backward modeling, all of which showed higher NST values for older participants across SNR conditions (Decruy et al., 2019; Decruy, Vanthornhout, and Francart, 2020; Presacco et al., 2016a, 2016b, 2019). In two of these studies, higher NST was found to be higher for both attended and unattended speech in older adults in most of the tested conditions (Presacco et al., 2016b, 2019). Notably, Decruy et al. (2019) reported that tracking remained stable from 17 to 50 years old, followed by a gradual increase with advancing age; this was the only study to include a distinct middle-aged group (aged between 39 and 60 years old). Two studies used a forward model; one reported higher NST for older adults in all conditions (Panela et al., 2024), while the other found higher NST with aging only in noisy conditions (Gillis, Decruy, et al., 2022). One study employed both forward and backward models, reporting higher NST (in both models) in older adults for both attended and unattended speech (Karunathilake et al., 2023). Regarding the latency of the neural response, Gillis, Decruy, et al. (2022) observed that the N1 peak latency of spectrogram tracking significantly decreased with age in quiet conditions, although the age effect was not robust in the presence of background noise. Additionally, Karunathilake et al. (2023) analyzed averaged latencies across noise conditions and found that, compared to younger adults, older adults exhibited a significantly earlier M50TRF peak, no significant latency difference for the middle M100TRF peak, and a significantly delayed late M200TRF peak (reflecting early, middle, and late responses of the MEG TRF at ~50, 100, and 200 ms, respectively). In one study using mutual information analysis, Zan et al. (2020) found that older adults exhibited significantly larger neural responses than younger adults across all three peaks of the temporal mutual information function MI50, MI100, and MI200 (occurring at approximately 50 ms, 100 ms, and 200 ms, respectively) in both clean and noisy speech conditions. This exaggerated tracking was observed for both foreground and background speech. In addition, age-related differences in hemispheric lateralization were observed. Younger adults showed a significant right-lateralized response for the MI100 peak and a trend toward right-lateralization for MI200, while older adults exhibited more bilateral patterns. For MI200, a significant age-related difference was found suggestive of a reduced right-hemisphere dominance in older listeners.

In sum, most studies reported higher NST in older adults compared to younger adults, with largely consistent findings across NST methods. The details of the studies supporting the analysis of age-related differences in NST are provided in Supplementary Table S2 and a summary of these findings is provided in Fig. 5A.

3.5.2. Hearing-Related Differences in NST

To investigate the effect of hearing impairment on NST, we focused on 11 studies that included both NH and HI participants. Among these, 5 included NH and HI adults in one group, while the remaining 6 compared NH and HI adults in separate groups. However, only 9 of these studies conducted specific analyses to examine the impact of hearing impairment on NST. An increase in NST associated with hearing loss was found in 5 studies: 1 used backward modeling (Decruy, Vanthornhout, and Francart, 2020), 1 used forward modeling (Gillis, Decruy, et al., 2022), 2 employed cross-correlation (Mirkovic et al., 2019; Schmitt et al., 2022), and 1 used both forward and backward modeling with consistent outcome (Fuglsang et al., 2020). Conversely, 3 studies (Kurthen et al., 2021; Petersen et al., 2017; Presacco et al., 2019), using forward, cross-correlation, and backward modeling respectively, reported no increase in NST related to hearing loss. Additionally, one study (Vander Ghinst et al., 2021) using coherence analysis characterized hearing impairment based on SPiN difficulties despite normal audiometric thresholds. This study found lower NST but a higher functional connectivity at 4–8 Hz between the auditory cortex and language/attention-related brain areas among SPiN-impaired participants compared to NH participants using a coherence NST method. Regarding neural response latency, Gillis, Decruy, et al. (2022) reported that TRFs were delayed in HI compared to NH participants in the quiet condition. Furthermore, while NH participants exhibited a latency decrease as speech understanding improved in noise for all peaks of the TRF, this effect was only observed for the P2 peak of the spectrogram in HI participants.

In sum, hearing impairment based on audiometric thresholds was associated with higher NST in most studies, while one study suggested that hearing impairment characterized by SPiN difficulty in the presence of normal audiometric thresholds could be associated with lower NST. The details of the studies supporting the analysis of hearing status differences in NST are provided in Supplementary Table S3 and a summary of these findings is provided in Fig. 5B

3.5.3. Impact of SNR on Neural Speech Tracking

To investigate the general impact of SNR on NST, we first examined the 34 studies that included only younger adults. From this group, we excluded 10 studies where SNR values were not specified and 3 study lacking specific information on the effect of SNR on NST which left 21 studies: 6 using a forward modeling (only); 3 using a backward modeling (only); 6 using forward and backward modeling; 1 using backward modeling and coherence (without information relative to SNR for the backward approach); 3 using coherence (only); 1 using cross-correlation; and 1 using phase locking value.

In studies using the forward modeling approach, 8 studies reported a general tendency of reduction in NST as noise increases (Decruy, Lesenfants, et al., 2020; Ding and Simon, 2013; Lesenfants et al., 2019; Verschueren et al., 2020; Wang et al., 2020; Yasmin et al., 2023; Zinszer et al., 2022; Zou et al., 2019). However, in three of these studies, NST increased at moderate noise levels but decreased at higher or lower noise intensities (Lesenfants et al., 2019; Yasmin et al., 2023; Zou et al., 2019). An absence of a clear relationship between SNR and NST was reported in 3 studies using the forward approach (Brown and Bidelman, 2022; Ding et al., 2014; Zhang et al., 2023). In one of the studies using forward modeling, an increase in NST was observed as SNR decreased only for stimuli presented in the first language (L1) of participants (Zinszer et al., 2022). Studies investigating the effect of SNR on the latency of forward

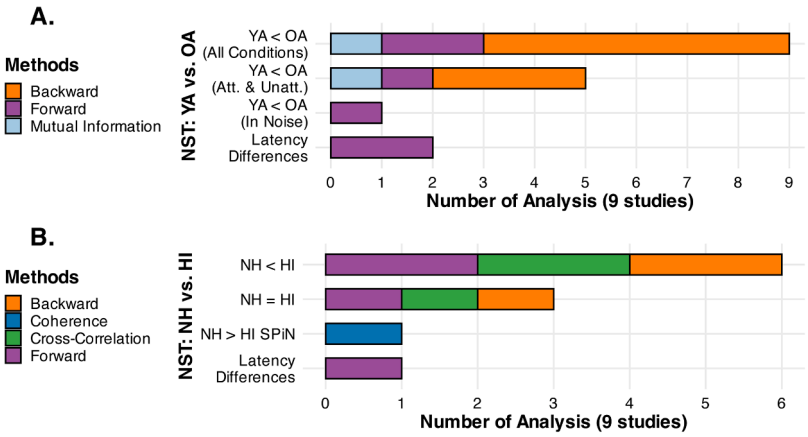

Fig. 5. **(A)** Summary of age-related differences in NST between younger adults (YA) and older adults (OA) as a function of NST methods and analysis across the 7 studies that examined this question. For further details, see Section 3.5.1 and Supplementary Tables S2. **(B)** Summary of hearing status-related differences in NST between normal-hearing (NH), hearing-impaired based on audiometric thresholds (HI), and hearing-impaired participants based on SPiN difficulties, as a function of NST methods and analysis across the 9 studies that examined this question. The symbols ⟨ and ⟩ indicate a lower or higher NST index for one population relative to the other, while = denotes no significant difference between the two groups. The abbreviations Att. and Unatt. refers to the NST of the attended and unattended speech respectively. The label "Latency Differences" refers to observed differences in the timing of neural responses as a function of age or hearing status. For further details, see Section 3.5.2 and Supplementary Tables S3.

modeling responses consistently show that higher noise levels tend to delay neural response latency (Brown and Bidelman, 2022; Ding and Simon, 2013; Verschueren et al., 2020; Wang et al., 2020; Yasmin et al., 2023; Zhang et al., 2023). However, Zhang et al. (2023) also reported that as SNR decreased, pre-onset semantic TRF latencies occurred progressively earlier.

Studies using backward modeling consistently reveal a trend of decreasing NST as SNR decreases (Crosse, Di Liberto, and Lalor, 2016; Verschueren et al., 2020; Wang et al., 2020; Zou et al., 2019). However, Ding and Simon (2013) found that reconstruction of the underlying speech envelope remained stable until −9 dB SNR, at which point it dropped. Earlier findings from the same group (Ding and Simon, 2012a) showed that, within a SNR range of ±8 dB, both forward models (using the spectrogram) and backward models (using the envelope) exhibited stable neural responses and reconstruction accuracy despite changes in relative intensity, supporting the existence of an object-based gain control mechanism. Zhang et al. (2023) showed that NST (reconstruction accuracy) decreased with decreasing SNR at the acoustic level, while NST at the semantic level exhibited no significant variation across SNR levels. Interestingly, one study found that NST in quiet conditions was lower than at intermediate SNR levels. In this study, Das et al. (2018) demonstrated that both attended and unattended speech correlations were lower in quiet conditions compared to when noise was present. Neural tracking of attended speech decreased with lower SNR, while unattended speech tracking remained relatively stable. However, the separation between attended and unattended speech correlations diminished as SNR decreased.

Studies using coherence also consistently showed a trend of decreasing NST as SNR decreased (Destoky et al., 2019; Vander Ghinst et al., 2019, 2016). Similarly, Vander Ghinst et al. (2021) found that cortical tracking of attended speech declined with lower SNR in the <1 Hz and 1–4 Hz bands (in NH participants). Interestingly, this study also revealed that, in the 4–8 Hz band, tracking peaked at intermediate SNRs (10 and 5 dB) and was weaker at both ends of the tested SNR continuum, such as quiet, 0 dB, and −5 dB SNR. Studies using cross-correlation and phase locking value provided additional insights into the effects of SNR on NST. Kong et al. (2014) found no significant effect of SNR on NST for either attended or unattended speech. However, for unattended speech, N1 and P2 latencies were significantly delayed, and amplitudes were reduced compared to attended speech. In contrast, Mohammadi et al. (2023) reported that phase locking values were significantly lower at −9 dB compared to 0 dB for sentence stimuli. In sum, most studies involving young adults found that NST decreases as SNR decreases. The details of the studies supporting the analysis of SNR-related differences in young adults are provided in Supplementary Table S4.

Next, we examined the 6 studies that included only older adults. Among those, two studies did not specify SNR values and were excluded, leaving four studies for analysis. Among these, two studies used the forward method: one observed a decrease in NST as SNR decreased (McHaney et al., 2021), while the other found no specific SNR-related decrease but noted differences in NST based on the number of background talkers (2 vs. 8) (Kurthen et al., 2021). The remaining two studies used the cross-correlation method, and both reported a decrease in NST as SNR decreased (Petersen et al., 2017; Schmitt et al., 2022). However, one study highlighted that this effect was significant only for babble noise, with no difference observed for pink noise (Schmitt et al., 2022). Overall, most studies including older adults found that NST decreases as SNR decreases, though the effects varied depending on the type of noise conditions. The details of the studies supporting the analysis of SNR-related differences in older adults are provided in Supplementary Table S5.

Next, we examined the 7 studies including both young and older adults (in one group). Of this set of study, 2 were excluded where SNR values were not specified NST which left 5 studies. Among those, 2 studies included both NH and HI participants (Decruy, Vanthornhout, and Francart, 2020; Gillis, Decruy, et al., 2022). Importantly, these two studies used the same participant sample and therefore cannot be considered independent. Among these 5 studies (4 independent samples), 1 study used backward modeling, 3 used forward modeling, and 1 applied both methods. Despite a relatively large participant age range across these studies (from the 20 s to 80 s), only 2 (Decruy, Vanthornhout, and Francart, 2020; Gillis, Decruy, et al., 2022) provided specific analyses on the interaction between SNR, age, and NST. Both studies (same participant sample) reported higher NST with increasing age and more favorable SNRs. The remaining studies did not specifically analyze the effect of age but generally observed a decrease in NST as SNR decreased (Fiedler et al., 2019; Muncke et al., 2022; Wang et al., 2023).

Finally, in the 7 studies that analyzed young and older adults in separate groups, only one explicitly included older adults with hearing impairment (Presacco et al., 2019). Using forward modeling, Panela et al. (2024) found no significant interaction between age, SNR, and NST. In this study, TRF amplitudes decreased in the early time window (0.03–0.06 s) in the presence of noise, while an increase was observed in

a later time window (0.09–0.13 s). In one study combining forward and backward modeling, Karunathilake et al. (2023) found that the M50TRF amplitude (an early prominent peak occurring around 50 ms in the TRF measured via MEG) decreased from quiet to noise for both age groups, with older adults showing a stronger reduction. The M100TRF amplitude (a peak occurring around 100 ms in the TRF) increased, while the M200TRF amplitude (a peak occurring around 200 ms in the TRF) decreased in older adults under noisy conditions, with no significant changes observed in younger adults. For both age groups, noise delayed the peak latencies of all three TRF components. Reconstruction accuracy of the attended talker envelope also declined significantly from quiet to noisy conditions in both groups. In the 4 studies using backward modeling, a stronger decline in NST with decreasing SNR was observed for older adults compared to younger adults (Decruy et al., 2019; Presacco et al., 2016a, 2016b, 2019). Presacco et al. (2019) also noted that older adults with hearing impairment demonstrated a significant decline in reconstruction values with decreasing SNR, which was not observed in younger or NH older adults. Decruy et al. (2019) showed that the age-related increase in envelope tracking was greater than a linear trend would predict, and that this increase interacted with speech understanding levels derived from individualized SNRs. This age-related increase was steeper for Matrix sentences than for audiobook materials and was significantly reduced at 0 dB SNR compared to the no-noise condition. The effect of age on envelope tracking was attenuated at lower speech understanding levels, corresponding to more challenging SNRs. Similarly, Zan et al. (2020) found that older adults exhibited significantly higher mutual information levels at all three peaks (MI50, MI100, MI200), across all SNRs, for both foreground and background speech. However, age modulated how MI200 responded to changes in SNR: in younger adults, MI200 remained stable for foreground speech but declined for background speech, whereas in older adults, MI200 declined for foreground speech and increased for background speech.

In sum, most studies reported a decline in NST with decreasing SNR, with a stronger impact observed in older adults compared to younger adults. The details of the studies supporting the analysis of SNR-related differences between young adults and older adults are provided in Supplementary Tables S6 and S7.

3.5.4. Relation Between NST and Behavioral SPiN Tasks

Another important aim of this review was to examine the relationship between NST and behavioral outcomes from SPiN tasks. Of the 54 studies included in the review, 25 studies did not explore this relationship, while the remaining 29 studies did. Not all SPiN task outcomes included in the studies were analyzed in relation to NST, as some tasks were only intended to maintain participants' attention to the ongoing stimulus. The outcomes from these studies are detailed in Supplementary Table S8. The tasks in these studies included speech intelligibility ($n = 15$), speech comprehension or understanding ($n = 14$), listening effort ($n = 4$), speech or sound detection ($n = 3$), and speech clarity ($n = 2$).

A positive association between speech intelligibility tasks and NST has also been found in many studies (Das et al., 2018; Ding and Simon, 2013; Lesenfants et al., 2019; Vander Ghinst et al., 2016; Vanthornhout et al., 2018; Verschueren et al., 2020), with the effect being more pronounced in older adults, as shown by Karunathilake et al. (2023). In contrast, Presacco et al. (2016a, 2016b) found no significant correlations between QuickSIN scores and NST in either younger or older adults. Notably, both studies, along with Zan et al. (2020), were based on the same participant sample. Using a mutual information approach, Zan et al. (2020) reported a negative association between MI200 levels and QuickSIN performance in older adults. An inverse relationship has also been observed, where latency increases as intelligibility decreases, in several studies (Karunathilake et al., 2023; Yasmin et al., 2023). This was true only for sentences compared to stories in Verschueren et al. (2020), and the effect became more pronounced as noise levels increased, according to Zhang et al. (2023). In contrast, some studies found no significant relationships between NST and intelligibility (Kurthen et al., 2021; Mohammadi et al., 2023; Schmitt et al., 2022).

Several studies found a positive association between NST and speech comprehension/understanding (Decruy, Lesenfants, et al., 2020; Keitel et al., 2018; Kurthen et al., 2021; McHaney et al., 2021; Verschueren et al., 2020), with similar findings reported by Decruy et al. (2019), who observed that envelope tracking increased with speech understanding in middle-aged and older NH adults, and that this increase was steeper in older adults. This pattern was observed across both NH and HI participants (Decruy, Vanthornhout, and Francart, 2020; Gillis, Decruy, et al., 2022; Schmitt et al., 2022), and stronger for native language than second language (Zinszer et al., 2022). When considering the peak latency of the NST (TRF) in relation with comprehension, an inverse relationship was often observed where latency increased as comprehension decreased (Brown and Bidelman, 2022; Zhang et al., 2023). While this mechanism remained true for NH participants, Gillis, Decruy, et al. (2022) suggested that it may only apply to one peak (P2 of the spectrogram tracking) in HI participants. Additionally, Etard and Reichenbach (2019), reported that the maximum correlation between NST and comprehension occurs in the delta band between −100 to 240 ms following stimulation. However, Har-Shai Yahav and Zion Golumbic (2021) found no significant relationship between comprehension performance and neural data.

When examining the relation between listening effort and NST, most studies found no association (Decruy, Vanthornhout, and Francart, 2020; Schubert et al., 2023), or only a marginally significant negative relationship as reported by Mohammadi et al. (2023). However, Decruy, Lesenfants, et al. (2020) reported that when participants understood the sentences well, NST decreased with increasing self-reported effort and reaction times in a dual-task paradigm. Conversely, when speech was harder to understand, NST increased with greater effort.

In studies examining the relation between speech clarity and NST, Etard and Reichenbach (2019) showed that at 160 ms latency following stimulation, NST correlated most strongly with speech clarity. On the other hand, Zhang et al. (2023), (in supplementary analyses) revealed significant negative correlations between NST and clarity ratings, although these findings were not further elaborated in the main text.

In studies examining the relation between speech detection and NST, Crosse, Di Liberto and Lalor (2016) and Puschmann et al. (2019) observed a significant positive relationship between NST and detection performance, although Crosse, Di Liberto and Lalor (2016) used a task that included visual stimuli. In contrast, Zou et al. (2019) found no significant relation between behavioral sound detection accuracy and NST under any condition.

In sum, most studies examining the relationship between behavioral performance and NST reported a positive association, where higher NST was linked to better task outcomes. Additionally, when response latencies were analyzed, many studies observed increased delays in neural responses corresponding to poorer behavioral performance.

4. Discussion

This review aimed to analyze the literature on NST during continuous speech processing in adverse acoustic conditions. Specifically, we focused on NST in relation to age, hearing loss, SNR, and SPiN performance. We found that most studies primarily focused on young adults without hearing impairments, while information on participants' neurological and cognitive health was inconsistently reported. Audiobooks were the most frequently used target material, and competing talkers were the most common adverse acoustic condition, with specific SNR values ranging from −12.5 to 12 dB. SPiN tasks were predominantly focused on speech intelligibility and comprehension. Most studies employed EEG to measure NST, with forward modeling being the most widely used method among the six identified approaches. NST primarily relied on the speech envelope, and more than half of the studies applied model-informed filtering to align with established auditory processing models. In sum, most studies comparing young and older adults reported

higher NST in older adults. Hearing impairment, as determined by audiometric thresholds, was associated with increased NST in most studies comparing NH and HI participants. Across studies, NST generally decreased as SNRs became less favorable, while the latency of the neural response tended to increase under the same conditions. Finally, most studies examining the relationship between behavioral performance and NST reported a positive association, where higher NST was linked to better SPiN task outcomes.

4.1. Underlying Mechanism of Age Differences in NST

The review highlights an underrepresentation of older adults in studies examining NST. While 34 studies focused exclusively on young adults, only 6 included older adults as the sole focus, 7 directly compared young and older cohorts, and another 7 included both young and older adults in a unique sample. However, in the 9 studies where the relation between age and NST was analyzed, all studies tended to demonstrate higher NST in older adults.

While Gillis, Decruy, et al. (2022), who used the same participant sample as Decruy, Vanthornhout and Francart (2020), were cautious in interpreting the enhanced NST in older adults (suggesting that the observed age effects might have been influenced by the three younger age-matched pairs) other studies have proposed various hypotheses to explain the NST enhancement associated with aging.

First, NST enhancement could be linked to an age-related imbalance between excitatory/inhibitory neural mechanisms mediated by gamma-aminobutyric acid (GABA) in the central auditory pathway, as suggested by Karunathilake et al. (2023), Panela et al. (2024), Presacco et al. (2016a, 2016b, 2019), and Zan et al. (2020). This hypothesis is supported by findings from both animal studies (e.g., de Villers-Sidani et al. (2010), Hughes et al. (2010), Juarez-Salinas et al. (2010)) and human studies (e.g., Dobri and Ross (2021), Harris et al. (2022), Lalwani et al. (2019), Ross et al. (2020)). The resulting higher neural activity may contribute to diminishing temporal, spectral, and spatial precision, which in turn, could increase difficulties in SPiN tasks (Herrmann and Butler, 2021). This excitatory/inhibitory imbalance hypothesis is further supported by findings showing that older adults exhibit higher NST (Decruy et al., 2019) even for unattended speech, as reported by Decruy, Vanthornhout and Francart (2020), Karunathilake et al. (2023), Presacco et al. (2016b, 2019) and Zan et al. (2020) reinforcing the notion that reduced neural inhibitory mechanism efficiency may underlie age-related changes in auditory neural processing.

Secondly, age-related NST enhancement could be explained by redundant neural processing across cortical regions involved in speech comprehension, which might serve to compensate for reduced cortical connectivity, as suggested by Karunathilake et al. (2023), Presacco et al. (2016a, 2016b, 2019) and Zan et al. (2020). According to Peelle et al. (2010) this compensation involves increased activity in frontal regions outside the core sentence-processing network, which helps offset the reduced coherence between activated cortical areas in older adults.

Finally, the amplified age-related NST might reflect recruitment of additional cortical regions through top-down compensatory mechanisms as suggested by Karunathilake et al. (2023), Panela et al. (2024), Presacco et al. (2016a), and Zan et al. (2020) based on findings of prior research (Pichora-Fuller et al., 2016; Rumschlag et al., 2022; Vaden et al., 2015; Wild et al., 2012; Wong et al., 2010). Presacco et al. (2019) emphasized that age, rather than peripheral hearing loss, appears to be the primary factor driving these enhanced cortical responses. The minimal differences observed between HI and NH older participants, compared to younger NH individuals, suggest that the exaggerated NST in older adults cannot be solely attributed to hearing impairment.

Regarding the latency of neural responses, only two studies have reported findings. Both studies observed a decrease in the latency of early responses in older adults, specifically the N1 of the spectrogram in Gillis, Decruy, et al. (2022) and the M50TRF in Karunathilake et al. (2023). Karunathilake et al. (2023) suggested that this earlier response in older adults reflects an excitation/inhibition imbalance favoring excitation compared to younger adults. They also proposed that this early response might result from an increased mid-latency response amplitude (M100TRF), which could shorten the latency of the earlier peak. Additionally, Karunathilake et al. (2023) found that older adults exhibited a delayed mid-latency response (M100TRF) and a significantly delayed late-latency response (M200TRF) compared to younger adults. Supported by Parthasarathy et al. (2020), they suggested that older adults may compensate for degraded afferent input by over-relying on middle- and late-stage processing mechanisms, engaging additional cortical regions and compensatory strategies at later stages. Although Zan et al. (2020) reported enhanced mutual information amplitudes across all latencies (50, 100, and 200 ms) in older adults compared to younger ones, they did not report any latency differences between age groups. Visual inspection of the mutual information function time courses presented in the figures suggests similar peak timing across groups, indicating that the observed age-related effects are primarily driven by response magnitude rather than latency shifts. Future studies could examine whether the timing of mutual information peaks differs with age, as has been shown in TRF-based studies.

To sum up, this review underscores the underrepresentation of older adults in NST research, despite consistent findings that NST is higher in older adults. Higher NST in older adults has been linked to age-related excitatory/inhibitory imbalances, compensatory neural recruitment, and increased cortical activity outside core processing networks. These findings suggest that aging, rather than hearing loss alone, plays a key role in driving NST enhancement, highlighting the need for further research to disentangle the mechanisms underlying these changes and their implications for auditory processing and SPiN in older populations.

4.2. Underlying Mechanism of Hearing-Related Differences in NST

In addition to highlighting the underrepresentation of older adults in NST research, this review also reveals a significant scarcity of HI participants in studies examining NST. Only 11 studies included both NH and HI participants, and only 9 conducted specific analyses to examine the impact of hearing impairment on NST.

Hearing loss was associated with higher NST in a slight majority of these studies ($n = 5$). Decruy, Vanthornhout and Francart (2020) highlighted that hearing loss has been associated with lower amplitude detection thresholds (Fullgrabe et al., 2003; Wallaert et al., 2017) and better sensitivity to suprathreshold modulations (Moore et al., 1996) in previous behavioral studies. They suggested that these findings may be linked to the enhanced NST observed in individuals with hearing loss in their study. They also specified that difference in participant-specific SNRs between NH and HI could not explain the effects since the same result was observed when SNRs were similar in both groups (e.g., fixed SNR condition). Fuglsang et al. (2020) pointed out that enhanced NST could be a peripheral inherited characteristic since damage to outer hair cells in presbycusis reduces cochlear compression (Ruggero and Rich, 1991), resulting in steeper level-response functions and loudness recruitment (Moore and Oxenham, 1998) and may, in turn, enhance envelope synchrony in auditory nerve fibers (Kale and Heinz, 2010).

Most authors who observed enhanced NST in HI participants have interpreted this phenomenon as a potential compensatory mechanism, although each proposed different underlying explanations. Based on animal studies (Chambers et al., 2016; Hughes et al., 2010; Salvi et al., 2016), Fuglsang et al. (2020) suggested that increased excitability of central neurons could help minimize sensitivity loss but could also lead to hyperactivity for mid- and high-level sounds. As for Decruy, Vanthornhout and Francart (2020), they suggested that enhanced NST in HI adults might reflect compensatory cortical processing to offset a potentially degraded subcortical response since hearing impairment have been associated with a higher interdependence between subcortical and cortical levels of the auditory system (Bidelman et al., 2014; Presacco et al., 2019).

Based on previous findings suggesting that adults with hearing loss perform worse on cognitive tests and need more effort compared to NH individuals (Humes et al., 2013; Lin et al., 2011; Ohlenforst et al., 2017), Decruy, Vanthornhout and Francart (2020) proposed a compensatory cognitive hypothesis where HI adults might recruit additional neural resources to separate target speech from background noise. This hypothesis, however, was not directly supported by cognitive test results or self-reported effort ratings in the study, possibly due to limitations in the sensitivity of the measures used.

Gillis, Decruy, et al. (2022) proposed that HI listeners recruit additional brain regions, such as frontal areas, to process degraded auditory input, an hypothesis supported by findings from Bidelman, Price, et al. (2019) and Campbell and Sharma (2013). Drawing on functional connectivity analyses by Bidelman, Mahmud, et al. (2019), which revealed that HI listeners exhibit "higher eccentricity" networks characterized by chain-like rather than star-like neural communication pathways, they suggested that this network structure may underlie the altered neural topographies and prolonged latencies observed in HI listeners. While these recruitment-based compensatory mechanisms offer some resilience, they appear to have limits under increasing noise conditions. Unlike NH listeners, whose response latencies increase as task demands grow, Gillis, Decruy, et al. (2022) observed that HI listeners' latencies remain consistently delayed, indicating that their compensatory mechanisms may already be operating at maximum capacity even under less challenging conditions.

Schmitt et al. (2022) found that enhanced NST as a function of hearing loss was primarily driven by responses occurring in an earlier time window (around 100 ms), which were linked to slow acoustic features, such as prosody and rhythm, that are particularly essential for older adults (Giroud et al., 2019; Meyer et al., 2018). However, since comprehension processes occur at a later stage and the association between NST and comprehension became stronger as hearing loss progressed, Schmitt et al. (2022) propose a compensatory mechanism where HI individuals increasingly rely on low-level acoustic features to support speech segmentation, which indirectly aids comprehension by mitigating the reduced audibility caused by hearing loss.

The hypothesis of a GABA-mediated inhibitory neurotransmission deficit leading to increased NST has also been linked to hearing loss, in addition to aging. Fuglsang et al. (2020) highlights Caspary's studies (Caspary et al., 2013, 2008) which suggest this mechanism is associated with cochlear damage as well as aging. Similarly, normal aging is linked to auditory nerve fiber degeneration (Sergeyenko et al., 2013; Wu et al., 2019), which can also increase central auditory system activity. This phenomenon, known as cochlear neuropathy (Viana et al., 2015), complicates the identification of a specific cause for increased NST (aging or hearing loss), as it can occur without clinical hearing loss in early stages.

Although age- and hearing-related neural changes such as reduced GABAergic inhibition and auditory nerve fiber loss have been well documented in animal models and post-mortem human studies, these mechanisms have not yet been directly linked to increased NST in living human participants. Future research should go beyond descriptive findings by testing mechanistic hypotheses using direct neurophysiological measures, for example by combining EEG or MEG with metabolic imaging techniques such as magnetic resonance spectroscopy (MRS) to estimate GABA concentrations (Puts and Edden, 2012), or positron emission tomography (PET) imaging of inhibitory neurotransmission (Andersson et al., 2019), although the latter remains costly and less accessible in most research settings.

4.3. SNR Modulates the NST Strength and Latency

When considering the effect of a decreasing SNR (increasing competing noise) in a SPiN task, a reasonable assumption is that the addition of noise progressively disrupts the synchrony between brain activity and the target speech signal. This is, in fact, what has been observed in most studies, regardless of the NST method used. However, in five studies involving NH younger adults, the relationship between SNR and NST appeared to be less straightforward. Lesenfants et al. (2019) identified an S-shaped relationship between SNR and NST in the theta band (4–8 Hz) for models based on low-level speech features (e.g., broadband envelope, spectrogram, and a combination of the time-aligned sequence of phonetic features and spectrogram). The maximum correlation was observed around 0 dB SNR, with reduced correlations at both higher and lower SNRs. In contrast, models based on higher-level speech features (e.g., time-aligned sequence of phonemes or binary phonetic features) in the delta (0.5–4 Hz) and theta bands exhibited a strictly monotonic increase in correlation as SNR improved. Similarly, Das et al. (2018) reported comparable results and suggested that the effort associated with top-down processes involved in SPiN tasks may enhance neural entrainment to the attended speech stream, resulting in stronger NST, particularly at mid-range SNRs. The reduction in NST observed in no-noise conditions likely reflects the absence of a need for such processes, whereas the lowest SNR conditions may overwhelm these mechanisms, also leading to reduced NST. Yasmin et al. (2023) provided a similar reasoning to explain the observed U-shape relationship between SNR and NST (TRF amplitude at the acoustic level) and stated that the increased negative amplitude toward mid-range SNRs may reflect increased attention or cognitive control for moderately masked, still intelligible speech whereas NST is reduced when masking reduces speech intelligibility beyond some point, at which point listeners essentially give up (Pichora-Fuller et al., 2016; Picou and Ricketts, 2018). This U-shape relationship, however, was not found when considering the semantic tracking where amplitude decreased only at low SNRs suggesting that the semantic TRF response is relatively robust to changes in babble-noise level as long as 50 % to 80 % of words are intelligible.

Zou et al. (2019) reported a divergence in the relationship between SNR and NST depending on both the stimulus representation and the TRF metric used. A U-shaped pattern emerged in the total power of the TRF weights (i.e., the waveform amplitude) only when the TRF was derived from the raw, non-normalized speech-noise mixture. This was interpreted as evidence for a contrast-dependent gain mechanism, whereby moderate noise levels amplify neural activity to compensate for reduced stimulus contrast. When the stimulus was amplitude-normalized or when the clean speech envelope was used, TRF amplitude decreased monotonically with decreasing SNR, reflecting a more direct effect of masking on neural encoding. The predictive power of the TRF (i.e., prediction accuracy) also declined monotonically with SNR across all stimulus types. Vander Ghinst et al. (2021) reported significantly higher coherence (mean across hemispheres) to the attended speech stream at intermediate SNRs as well compared to extreme SNRs. This pattern was observed in the theta band for NH participants and some SPiN-impaired participants. The authors highlighted the critical role of theta band activity in speech perception, particularly in parsing continuous speech into discrete syllabic units and chunking acoustic features at the syllable timescale to support auditory stream formation (e.g., Ding et al. (2016); Hyafil et al. (2015); Riecke et al. (2015); Teng et al. (2018)). Although not explicitly mentioned by the authors, we speculate that the SNR-related phenomenon observed in this study may share similarities with the findings of Lesenfants et al. (2019) discussed earlier, as both occur in the theta band and are based on low-level speech features. Overall, most studies converge toward a decreasing NST as SNR decrease but few studies demonstrate that the quieter conditions can also be associated with lower NST compared to intermediate SNR levels resulting in a U-shaped pattern of response (Das et al., 2018; Lesenfants et al., 2019; Vander Ghinst et al., 2021; Yasmin et al., 2023; Zou et al., 2019). This pattern does not appear to be tied to a specific NST method, as it was observed across diverse approaches including forward modeling (TRF amplitude (Yasmin et al., 2023; Zou et al., 2019) or prediction accuracy (Lesenfants et al., 2019)), backward modeling (Das et al., 2018), and coherence (Vander Ghinst et al., 2021).

It also emerged with both noisy and clean speech predictors (i.e., envelope including noise or envelope of the clean speech) as well as in the presence of various distractor type such as noise (Lesenfants et al., 2019; Zou et al., 2019), multitalker babble (Vander Ghinst et al., 2021; Yasmin et al., 2023) or competing talker (Das et al., 2018) (see Table 1 and Supplementary Table S1 for details on the predictors under "Extracted Features"). It is important to note that using clean or noisy speech predictors does not alter the auditory stimulation itself (i.e., participants always hear the noisy mixture). Rather, it influences the features used to model the neural response. Thus, when increased NST is observed at intermediate SNR levels even with clean predictors, as in Lesenfants et al. (2019), it suggests that mechanisms beyond low-level acoustic contrast, such as attention or listening effort, may also contribute. Although not included in the present review due to their use of synthetic speech, Herrmann (2025) also observed a similar U-shaped relationship at moderate to high SNRs under both attentive and passive listening conditions. This finding indicates that attention alone may not fully account for the observed pattern, and that additional mechanisms may be involved. However, the studies included in this review that reported a U-shaped relationship between NST and SNR share several common features: they all examined younger adults with normal hearing, employed a wide range of SNR conditions, and analyzed NST in the delta and theta frequency bands. A unique finding however is the increased NST as SNR decreased for stimuli presented in the L1 of participants in Zinszer et al. (2022). Since this study used a limited number of conditions (Q, 5, 0 dB SNR) a plausible scenario would be that the NST would decrease at lower NST, thus producing the U-shaped pattern discussed above.

Another important impact of SNR on NST is the delayed signal latency observed as SNR decreases. In younger adults, this phenomenon was identified by Brown and Bidelman (2022) who linked it to early processes in speech perception (also described by Crosse, Di Liberto and Lalor (2016), and Verschueren et al. (2020)) and theories of attentional load (Bendixen, 2014). Verschueren et al. (2020) suggested that the increase in latency with decreasing SNR could reflect either an enhancement of top-down processing or a decline in attention and listening effort. In Wang et al. (2020), the increased latency of the P2 component with decreasing SNR was proposed to indicate greater difficulty in perceiving attended speech amidst background noise (Ding and Simon, 2012a, 2014), as SNR had no effect on the ignored stream. Yasmin et al. (2023) observed that semantic tracking latencies were more robust to changes in SNR than acoustic tracking, suggesting that contextual information is more resilient to challenging listening conditions compared to acoustics and intelligibility. Zhang et al. (2023) found consistent results regarding the latency of acoustic TRFs, proposing that latency could serve as a sensitive marker for studying noisy speech processing, supported by findings from Ding and Simon (2013), Kaplan-Neeman et al. (2006), and Whiting et al. (1998). Additionally, they reported a unique finding related to pre-onset responses to word entropy at the semantic level, showing that peak latencies occurred earlier as noise levels increased. The authors suggested that this mechanism might compensate for noise interference, drawing on semantic prediction studies (Mattys et al., 2012; Miller et al., 1951; Obleser and Kotz, 2010; Pickering and Gambi, 2018; Zekveld et al., 2011). Initiating pre-onset responses earlier and extending their duration would provide the brain with additional time to prepare for upcoming speech information, thereby enhancing the chances of successful comprehension. In studies including younger and older adults, Gillis, Decruy, et al. (2022) observed that latency decreased with favorable SNRs in NH adults (as in Muncke et al. (2022)), but this effect was limited to one peak (P2 of the spectrogram) in HI adults. While the authors did not elaborate on this specific finding, they suggested that latency remains elevated across SNRs in HI listeners, as shown in Bidelman, Mahmud, et al., (2019), because they may already recruit extensive brain regions for speech understanding. As a result, their neural response latency does not further increase with additional noise, suggesting a limited compensatory capacity. Karunathilake et al. (2023) demonstrated that all TRF peaks (M50, M100, and M200) were significantly delayed under all conditions compared to quiet, for both younger and older adults. However, aging contributed more substantially to the peak delay from quiet to 0 dB in both the M100TRF and M200TRF. Supported by findings from Parthasarathy et al. (2020), the authors suggested that older adults may rely more heavily on middle and late processing mechanisms to compensate for degraded afferent input.

A complementary perspective is offered by Brodbeck et al. (2020), who examined the effect of acoustic masking on response latency without explicitly manipulating SNR levels (i.e., in a competing talker paradigm) in NH young adults. They observed that masked onsets in continuous speech were associated with delayed cortical responses, regardless of whether the masked speech stream was attended or ignored. Importantly, they argue that such latency shifts do not result from a uniform degradation due to reduced SNR, but rather from a dynamic, feature-specific delay reflecting the difficulty of recovering masked auditory objects. Moreover, they suggest that SNR cannot serve as a causal explanatory variable, since it presupposes prior stream segregation. This interpretation challenges traditional views of SNR as a direct determinant of neural response quality and instead supports a model in which the auditory system adaptively recovers relevant features, even from ignored streams, depending on their perceptual accessibility. These findings reinforce the idea that cortical latency shifts may reflect active object-level processing rather than passive signal degradation (Ding and Simon, 2012a). In a related line of research, Brodbeck et al. (2018) examined the NST at the acoustic and lexical levels (same participants as in Brodbeck et al. (2020)). Their results show that while responses to acoustic onsets were observed for both attended and unattended speech streams, lexical predictors such as word onsets accounted for neural responses only in the attended stream. This dissociation highlights a distinction between early auditory processing and higher-level linguistic analysis. It suggests that low-level acoustic features can be tracked automatically, even in unattended speech, whereas the extraction of lexical-semantic information is modulated by attention. These findings are consistent with the ELU model (Ronnberg et al., 2013), which posits that early speech processing involves RAMBPHO, a Rapid, Automatic, Multimodal Binding of PHOnological information that operates pre-attentively to match incoming input with phonological representations in long-term memory. However, when speech input is degraded or masked, as in a competing talker paradigm, RAMBPHO alone may not be sufficient to achieve lexical access. In such cases, attention and object-level integration may be required to support higher-level, lexico-semantic processing.

To conclude, the relationship between SNR and NST is complex, though most studies, regardless of the NST method, report lower NST as SNR declines, though exceptions highlight potential U-shaped or task-dependent patterns influenced by cognitive and neural processing demands. Further, SNR impacts response latency, with increasing noise generally causing delayed neural responses. Future research should further explore these effects to better understand auditory and speech processing in challenging listening conditions across different populations, where evidence is more limited and more complex.

4.4. Behavioral SPiN Outcome is Likely to be Reflected in the NST

The data summarized in this review suggest that behavioral SPiN performance is generally reflected in the quality of NST, regardless of the method used. Only a minority of studies examining this relationship ($n = 10$) reported no significant effects with respect to one or more behavioral tasks. This lack of association does not appear to be driven by the type of SPiN task, as it was observed across comprehension tasks (Har-Shai Yahav and Zion Golumbic, 2021) , intelligibility tasks (Mohammadi et al., 2023; Presacco et al., 2016a, 2016b; Schmitt et al., 2022; Vander Ghinst et al., 2016; Verschueren et al., 2020), tasks related to listening effort (Decruy, Vanthornhout, and Francart, 2020; Schubert

et al., 2023), and speech detection tasks (Zou et al., 2019). When discussed, these null findings are often attributed to limitations in the behavioral measures themselves, such as lack of sensitivity or misalignment with the speech material used in the neural task. For instance, the absence of correlation with QuickSIN performance in Presacco et al. (2016a, 2016b) may stem from the fact that the behavioral test and the neural tracking analysis were based on different speech materials.

Furthermore, in some studies, behavioral performance data were embedded within the SNR values. This occurs when subject-specific SNRs are derived from SRT values, which can then be used to estimate a percentage of speech understanding, comprehension or intelligibility (Decruy, Lesenfants, et al., 2020; Decruy et al., 2019; Decruy, Vanthornhout, and Francart, 2020; Gillis, Decruy, et al., 2022; Keitel et al., 2018; Lesenfants et al., 2019; Mirkovic et al., 2019; Muncke et al., 2022; Petersen et al., 2017; Vanthornhout et al., 2018; Verschueren et al., 2020; Wang et al., 2023). In such cases, the relationship between behavioral performance and NST likely mirrors the relationship between SNR and NST. However, specific analysis addressing this question concluded that NST is related to changes in speech understanding that cannot be solely explained by changes in SNR, thereby reinforcing evidence of a genuine relationship between NST and SPiN capacity (Decruy et al., 2019; Decruy, Vanthornhout, and Francart, 2020). It is also worth noting that several studies using subject-specific SNRs based on SRTs refer to their outcome measures as reflecting "speech understanding" or "comprehension". However, since SRTs are typically derived from intelligibility tasks (i.e., word recognition or recall), we believe that such outcomes more accurately reflect speech intelligibility.

Additionally, we observed similar general trends across studies (i.e., increased NST associated with positive outcomes in behavioral tasks and increased latency associated with negative outcomes) when comparing results from younger and older adults, with and without hearing loss. However, Karunathilake et al. (2023) identified finer-grained differences between younger and older adults. Specifically, they found that stronger early M50TRF amplitudes were associated with better speech intelligibility scores, and this effect was more pronounced in older adults. Conversely, smaller late M200TRF amplitudes were linked to poorer speech intelligibility scores exclusively in older adults. In line with the findings of Wingfield and Grossman (2006) the authors proposed an age-related compensatory mechanism for speech intelligibility, specifying that it is primarily active during the late stages of speech processing. Interestingly, these results align with the findings of Kurthen et al. (2021) where the late response TRF300 significantly predicted accuracy in a comprehension task. Specifically, larger amplitudes (more positive values) were associated with better accuracy in a comprehension task in older adults (mean age = 70.96 years). Also consistent with this finding are the results of Schmitt et al. (2022) who suggested that the effect of cross-correlation was primarily observed in the later time window (around 200 ms). Specifically, a one-unit increase in cross-correlation was associated with a 15 % increase in the odds of responding correctly in a comprehension task, also in older adults (mean age = 70.52 years). However, this narrative does not align with Zan et al. (2020) who found a negative association between NST (mutual information late peak MI200) and speech intelligibility (QuickSIN) scores in older adults using the same participant sample and dataset as Presacco et al. (2016a, 2016b). The behavioral and neural measures relied on different speech and distractor materials, creating a mismatch that may have limited the comparability between NST and behavioral outcomes.

Decruy et al. (2019) reported that, on an individual level, envelope tracking increased with better speech understanding, suggesting that NST reflects the quality of speech encoding within listeners. However, the idea that NST increases with age and hearing loss, two factors known to reduce SPiN abilities, while also being associated with better behavioral performance may appear paradoxical. In the same study, the authors reported that envelope tracking was higher in older adults even when speech understanding or SNR was matched with younger participants. They suggested that this enhancement may not reflect improved speech encoding, but rather the degree of difficulty experienced during processing, indicating that stronger neural responses can also signify increased listening effort rather than better perceptual outcomes. This interpretation was partially corroborated in a subsequent study by the same group (Decruy, Lesenfants, et al., 2020), which found that the relationship between NST and effort depended on the level of speech understanding, whereas another study from the same authors (Decruy, Vanthornhout, and Francart, 2020) found no significant association between NST and self-reported effort.

Taken together, the reviewed studies indicate a consistent relationship between SPiN capacity and NST, with stronger tracking generally associated with better SPiN performance. Although a few studies reported null effects, these were often attributed to methodological limitations associated with the task. Evidence suggests that NST reflects not only SNR-driven changes but also unique aspects of speech understanding, particularly in older adults, where late-stage neural responses (e.g., M200TRF and TRF300) appear critical for compensatory processing. These findings highlight the complex interplay between neural and behavioral measures, emphasizing the importance of considering both early and late neural dynamics in understanding speech intelligibility across the lifespan.

5. Conclusion

This review highlights the interplay between NST and speech processing in adverse listening conditions, emphasizing its modulation by age, hearing loss, and SNR. The evidence points to a robust relationship between NST and behavioral SPiN performance, with most studies finding that better NST is associated with improved task outcomes. Age-related differences in NST are consistently observed and may reflect compensatory mechanisms, such as greater reliance on late-stage and top-down processing, increased listening effort, and recruitment of additional cortical regions. However, the absence of longitudinal evidence limits the robustness of these interpretations. Similarly, hearing loss appears to amplify NST, potentially as a peripheral or central compensation for degraded auditory input, but these conclusions, being derived from cross-sectional data, should also be approached with caution. Lastly, while NST generally decreases as SNR worsens, specific patterns such as U-shaped or monotonic relationships highlight the complexity of these effects and their potential dependence on the task, methodology, and participant characteristics. Together, these findings underline the potential of NST as a sensitive marker of auditory and linguistic processing in challenging acoustic environments, offering a foundation for future research to further explore its mechanisms and implications.

CRediT authorship contribution statement

David Ratelle: Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Pascale Tremblay:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Conceptualization.

Funding

Pascale Tremblay holds a Canada Research Chair in the Neurobiology of Speech and Hearing (CRC #2022–00090).

David Ratelle holds scholarships from Fonds de Recherche du Québec en Santé (https://doi.org/10.69777/344801) and Canadian Institutes of Health Research. None of the funding sources had any involvement in the study design, collection, analysis and interpretation of data, writing of the report and decision to submit the article for publication.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.heares.2025.109367.

Data availability

The data used is provided in the supplementary material and in the text

References

Albouy, P., Weiss, A., Baillet, S., Zatorre, R.J., 2017. Selective entrainment of theta oscillations in the dorsal stream causally enhances auditory working memory performance. Neuron 94 (1), 193–206 e195.

Andersson, J.D., Matuskey, D., Finnema, S.J., 2019. Positron emission tomography imaging of the gamma-aminobutyric acid system. Neurosci. Lett. 691, 35–43. https://doi.org/10.1016/j.neulet.2018.08.010.

Arbogast, T.L., Mason, C.R., Kidd Jr, G., 2002. The effect of spatial separation on informational and energetic masking of speech. J. Acoust. Soc. Am. 112 (5), 2086–2098.

Beck, D., Larsen, D., Bush, E., 2018. Speech in noise: Hearing loss, neurocognitive disorders, aging, traumatic brain injury and more. J Otolaryngol ENT Res 10 (4), 199–205.

Bendixen, A., 2014. Predictability effects in auditory scene analysis: a review. Front. Neurosci. 8, 60. https://doi.org/10.3389/fnins.2014.00060.

Berezutskaya, J., Baratin, C., Freudenburg, Z.V., Ramsey, N.F., 2020. High-density intracranial recordings reveal a distinct site in anterior dorsal precentral cortex that tracks perceived speech. Hum. Brain Mapp. 41 (16), 4587–4609.

Bidelman, G.M., Mahmud, M.S., Yeasin, M., Shen, D., Arnott, S.R., Alain, C., 2019. Age-related hearing loss increases full-brain connectivity while reversing directed signaling within the dorsal-ventral pathway for speech. Brain Struct. Funct. 224 (8), 2661–2676. https://doi.org/10.1007/s00429-019-01922-9.

Bidelman, G.M., Price, C.N., Shen, D., Arnott, S.R., Alain, C., 2019. Afferent-efferent connectivity between auditory brainstem and cortex accounts for poorer speech-in-noise comprehension in older adults. Hear. Res. 382, 107795. https://doi.org/10.1016/j.heares.2019.107795.

Bidelman, G.M., Villafuerte, J.W., Moreno, S., Alain, C., 2014. Age-related changes in the subcortical-cortical encoding and categorical perception of speech. Neurobiol. Aging 35 (11), 2526–2540. https://doi.org/10.1016/j.neurobiolaging.2014.05.006.

Biesmans, W., Das, N., Francart, T., Bertrand, A., 2017. Auditory-Inspired Speech Envelope Extraction Methods for Improved EEG-Based Auditory Attention Detection in a Cocktail Party Scenario. IEEe Trans. Neural Syst. Rehabil. Eng. 25 (5), 402–412. https://doi.org/10.1109/TNSRE.2016.2571900.

Blackburn, H.L., Benton, A.L., 1957. Revised administration and scoring of the digit span test. J. Consult. Psychol. 21 (2), 139.

Brodbeck, C., Hong, L.E., Simon, J.Z., 2018. Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech. Curr. Biol. 28 (24), 3976–3983. https://doi.org/10.1016/j.cub.2018.10.042 e3975.

Brodbeck, C., Jiao, A., Hong, L.E., Simon, J.Z., 2020. Neural speech restoration at the cocktail party: Auditory cortex recovers masked speech of both attended and ignored speakers. PLoS. Biol. 18 (10), e3000883. https://doi.org/10.1371/journal.pbio.3000883.

Brodbeck, C., Simon, J.Z., 2020. Continuous speech processing. Curr. Opin. Physiol. 18, 25–31.

Brown, J.A., Bidelman, G.M., 2022. Familiarity of Background Music Modulates the Cortical Tracking of Target Speech at the "Cocktail Party". Brain Sci. 12 (10). https://doi.org/10.3390/brainsci12101320.

Brown, W.S., Marsh, J.T., Smith, J.C., 1973. Contextual meaning effects on speech-evoked potentials. Behav. Biol. 9 (6), 755–761. https://doi.org/10.1016/s0091-6773(73)80136-1.

Brungart, D.S., 2001. Informational and energetic masking effects in the perception of two simultaneous talkers. J. Acoust. Soc. Am. 109 (3), 1101–1109.

Brungart, D.S., Simpson, B.D., Ericson, M.A., Scott, K.R., 2001. Informational and energetic masking effects in the perception of multiple simultaneous talkers. J. Acoust. Soc. Am. 110 (5), 2527–2538.

Brus, B.T., & Voeten, M. (1973). Eén-minuut-test [One Minute Test]. Nijmegen, The Netherlands: Berkhout.

Buchsbaum, M., Fedio, P., 1970. Hemispheric differences in evoked potentials to verbal and nonverbal stimuli in the left and right visual fields. Physiol. Behav. 5 (2), 207–210.

Campbell, J., Sharma, A., 2013. Compensatory changes in cortical resource allocation in adults with hearing loss. Front. Syst. Neurosci. 7, 71. https://doi.org/10.3389/fnsys.2013.00071.

Caspary, D.M., Hughes, L.F., Ling, L.L., 2013. Age-related GABAA receptor changes in rat auditory cortex. Neurobiol. Aging 34 (5), 1486–1496. https://doi.org/10.1016/j.neurobiolaging.2012.11.009.

Caspary, D.M., Ling, L., Turner, J.G., Hughes, L.F., 2008. Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. J. Exp. Biol. 211 (Pt 11), 1781–1791. https://doi.org/10.1242/jeb.013581.

CHABA, 1988. Speech understanding and aging. Working Group on Speech Understanding and Aging. Committee on Hearing, Bioacoustics, and Biomechanics, Commission on Behavioral and Social Sciences and Education, National Research Council [Review]. J. Acoust. Soc. Am. 83 (3), 859–895. http://www.ncbi.nlm.nih.gov/pubmed/3281988.

Chambers, A.R., Resnik, J., Yuan, Y., Whitton, J.P., Edge, A.S., Liberman, M.C., Polley, D. B., 2016. Central Gain Restores Auditory Processing following Near-Complete Cochlear Denervation. Neuron 89 (4), 867–879. https://doi.org/10.1016/j.neuron.2015.12.041.

Cherry, E.C., 1953. Some experiments on the recognition of speech, with one and with two ears. J. Acoust. Soc. Am. 25 (5), 975–979. https://doi.org/10.1121/1.1907229.

Cohen, G., Martin, M., 1975. Hemisphere differences in an auditory Stroop test. Percept. Psychophys. 17 (1), 79–83. https://doi.org/10.3758/bf03204002.

Cohen, R., 1971. Differential cerebral processing of noise and verbal stimuli. Science (1979) 172 (3983), 599–601. https://doi.org/10.1126/science.172.3983.599.

Conners, C., 2014. Conners Continuous Auditory Test of Attention. Multi-Health Systems, North Tonawanda, NY.

Covidence. Covidence Systematic Review Software. In V. H. Innovation (Ed.). Melbourne, Australia: www.covidence.org.

Crosse, M.J., Di Liberto, G.M., Bednar, A., Lalor, E.C., 2016. The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. Front. Hum. Neurosci. 10, 604.

Crosse, M.J., Di Liberto, G.M., Lalor, E.C., 2016. Eye Can Hear Clearly Now: Inverse Effectiveness in Natural Audiovisual Speech Processing Relies on Long-Term Crossmodal Temporal Integration. J. Neurosci. 36 (38), 9888–9895. https://doi.org/10.1523/JNEUROSCI.1396-16.2016.

Crosse, M.J., Zuk, N.J., Di Liberto, G.M., Nidiffer, A.R., Molholm, S., Lalor, E.C., 2021. Linear Modeling of Neurophysiological Responses to Speech and Other Continuous Stimuli: Methodological Considerations for Applied Research. Front. Neurosci. 15, 705621. https://doi.org/10.3389/fnins.2021.705621.

Culling, J.F., Stone, M.A., 2017. Energetic masking and masking release. The auditory system at the cocktail party 41–73.

Das, N., Bertrand, A., Francart, T., 2018. EEG-based auditory attention detection: boundary conditions for background noise and speaker positions. J. Neural Eng. 15 (6), 066017. https://doi.org/10.1088/1741-2552/aae0a6.

Davis, A.C., 1989. The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. Int. J. Epidemiol. 18 (4), 911–917.

de Villers-Sidani, E., Alzghoul, L., Zhou, X., Simpson, K.L., Lin, R.C., Merzenich, M.M., 2010. Recovery of functional and structural age-related changes in the rat primary auditory cortex with operant training. Proc. Natl. Acad. Sci. u S. a 107 (31), 13900–13905. https://doi.org/10.1073/pnas.1007885107.

Decruy, L., Lesenfants, D., Vanthornhout, J., Francart, T., 2020. Top-down modulation of neural envelope tracking: The interplay with behavioral, self-report and neural measures of listening effort. Eur. J. Neurosci. 52 (5), 3375–3393. https://doi.org/10.1111/ejn.14753.

Decruy, L., Vanthornhout, J., Francart, T., 2019. Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. J. Neurophysiol. 122 (2), 601–615. https://doi.org/10.1152/jn.00687.2018.

Decruy, L., Vanthornhout, J., Francart, T., 2020. Hearing impairment is associated with enhanced neural tracking of the speech envelope. Hear. Res. 393, 107961.

Desai, M., Holder, J., Villarreal, C., Clark, N., Hoang, B., Hamilton, L.S., 2021. Generalizable EEG Encoding Models with Naturalistic Audiovisual Stimuli. J. Neurosci. 41 (43), 8946–8962. https://doi.org/10.1523/JNEUROSCI.2891-20.2021.

Destoky, F., Philippe, M., Bertels, J., Verhasselt, M., Coquelet, N., Vander Ghinst, M., Wens, V., De Tiege, X., Bourguignon, M., 2019. Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal envelope. Neuroimage 184, 201–213. https://doi.org/10.1016/j.neuroimage.2018.09.006.

Ding, N., Chatterjee, M., Simon, J.Z., 2014. Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. Neuroimage 88, 41–46. https://doi.org/10.1016/j.neuroimage.2013.10.054.

Ding, N., Melloni, L., Zhang, H., Tian, X., Poeppel, D., 2016. Cortical tracking of hierarchical linguistic structures in connected speech. Nat. Neurosci. 19 (1), 158–164. https://doi.org/10.1038/nn.4186.

Ding, N., Simon, J.Z., 2012a. Emergence of neural encoding of auditory objects while listening to competing speakers. Proc. Natl. Acad. Sci. u S. a 109 (29), 11854–11859. https://doi.org/10.1073/pnas.1205381109.

Ding, N., Simon, J.Z., 2012b. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. J. Neurophysiol. 107 (1), 78–89. https://doi.org/10.1152/jn.00297.2011.

Ding, N., Simon, J.Z., 2013. Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. J. Neurosci. 33 (13), 5728–5735. https://doi.org/10.1523/JNEUROSCI.5297-12.2013.

Ding, N., Simon, J.Z., 2014. Cortical entrainment to continuous speech: functional roles and interpretations. Front. Hum. Neurosci. 8, 311. https://doi.org/10.3389/fnhum.2014.00311.

Dobri, S.G.J., Ross, B., 2021. Total GABA level in human auditory cortex is associated with speech-in-noise understanding in older age. Neuroimage 225, 117474. https://doi.org/10.1016/j.neuroimage.2020.117474.

Drozdick, L., Wahlstrom, D., Zhu, J., & Weiss, L. (2012). The Wechsler Adult Intelligence Scale—Fourth Edition and the Wechsler Memory Scale—Fourth Edition. In.

Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2016. Increased activity in frontal motor cortex compensates impaired speech perception in older adults. Nat. Commun. 7, 12241. https://doi.org/10.1038/ncomms12241.

Eriksen, B.A., Eriksen, C.W., 1974. Effects of noise letters upon the identification of a target letter in a nonsearch task. Percept. Psychophys. 16 (1), 143–149. https://doi.org/10.3758/bf03203267.

Etard, O., Reichenbach, T., 2019. Neural Speech Tracking in the Theta and in the Delta Frequency Band Differentially Encode Clarity and Comprehension of Speech in Noise. J. Neurosci. 39 (29), 5750–5759. https://doi.org/10.1523/JNEUROSCI.1828-18.2019.
Fiedler, L., Wostmann, M., Herbst, S.K., Obleser, J., 2019. Late cortical tracking of ignored speech facilitates neural selectivity in acoustically challenging conditions. Neuroimage 186, 33–42. https://doi.org/10.1016/j.neuroimage.2018.10.057.
Foster, N.E., Halpern, A.R., Zatorre, R.J., 2013. Common parietal activation in musical mental transformations across pitch and time. Neuroimage 75, 27–35.
Freyman, R.L., Balakrishnan, U., Helfer, K.S., 2001. Spatial release from informational masking in speech recognition. J. Acoust. Soc. Am. 109 (5), 2112–2122.
Freyman, R.L., Balakrishnan, U., Helfer, K.S., 2004. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. J. Acoust. Soc. Am. 115 (5), 2246–2256.
Freyman, R.L., Helfer, K.S., Balakrishnan, U., 2007. Variability and uncertainty in masking by competing speech. J. Acoust. Soc. Am. 121 (2), 1040–1046.
Freyman, R.L., Helfer, K.S., McCall, D.D., Clifton, R.K., 1999. The role of perceived spatial separation in the unmasking of speech. J. Acoust. Soc. Am. 106 (6), 3578–3588.
Frisina, D.R., Frisina, R.D., 1997. Speech recognition in noise and presbycusis: relations to possible neural mechanisms. Hear. Res. 106 (1-2), 95–104.
Fuglsang, S.A., Dau, T., Hjortkjaer, J., 2017. Noise-robust cortical tracking of attended speech in real-world acoustic scenes. Neuroimage 156, 435–444. https://doi.org/10.1016/j.neuroimage.2017.04.026.
Fuglsang, S.A., Marcher-Rorsted, J., Dau, T., Hjortkjaer, J., 2020. Effects of Sensorineural Hearing Loss on Cortical Synchronization to Competing Speech during Selective Attention. J. Neurosci. 40 (12), 2562–2572. https://doi.org/10.1523/JNEUROSCI.1936-19.2020.
Fullgrabe, C., Meyer, B., Lorenzi, C., 2003. Effect of cochlear damage on the detection of complex temporal envelopes. Hear. Res. 178 (1-2), 35–43. https://doi.org/10.1016/s0378-5955(03)00027-3.
Gillis, M., Decruy, L., Vanthornhout, J., Francart, T., 2022. Hearing loss is associated with delayed neural responses to continuous speech. European Journal of Neuroscience 55 (6), 1671–1690.
Gillis, M., Van Canneyt, J., Francart, T., Vanthornhout, J., 2022. Neural tracking as a diagnostic tool to assess the auditory pathway. Hear. Res. 426, 108607.
Gilliver, M., Beach, E.F., Williams, W., 2013. Noise with attitude: Influences on young people's decisions to protect their hearing. Int. J. Audiol. 52 (sup1), S26–S32.
Gilliver, M., Beach, E.F., Williams, W., 2015. Changing beliefs about leisure noise: Using health promotion models to investigate young people's engagement with, and attitudes towards, hearing health. Int. J. Audiol. 54 (4), 211–219.
Giroud, N., Keller, M., Hirsiger, S., Dellwo, V., Meyer, M., 2019. Bridging the brain structure-brain function gap in prosodic speech processing in older adults. Neurobiol. Aging 80, 116–126. https://doi.org/10.1016/j.neurobiolaging.2019.04.017.
Glanz, O., Derix, J., Kaur, R., Schulze-Bonhage, A., Auer, P., Aertsen, A., Ball, T., 2018. Real-life speech production and perception have a shared premotor-cortical substrate. Sci. Rep. 8 (1), 8898.
Glasberg, B.R., Moore, B.C., 1990. Derivation of auditory filter shapes from notched-noise data. Hear. Res. 47 (1-2), 103–138. https://doi.org/10.1016/0378-5955(90)90170-t.
Greenberg, H.J., Graham, J.T., 1970. Electroencephalographic changes during learning of speech and nonspeech stimuli. J. Verbal. Learning. Verbal. Behav. 9 (3), 281–284.
Hällgren, M., Larsby, B., Arlinger, S., 2006. A Swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition: Una versión sueca de la Prueba de Audición en Ruido (HINT) para evaluar el reconocimiento del lenguaje. Int. J. Audiol. 45 (4), 227–237.
Halliday, D.M., Rosenberg, J.R., Amjad, A.M., Breeze, P., Conway, B.A., Farmer, S.F., 1995. A framework for the analysis of mixed time series/point process data–theory and application to the study of physiological tremor, single motor unit discharges and electromyograms. Prog. Biophys. Mol. Biol. 64 (2-3), 237–278. https://doi.org/10.1016/s0079-6107(96)00009-0.
Hambrook, D.A., Tata, M.S., 2019. The effects of distractor set-size on neural tracking of attended speech. Brain Lang. 190, 1–9. https://doi.org/10.1016/j.bandl.2018.12.005.
Hammes, J.G.W., 1978. De Stroop kleur-Woord Test. Harcourt Test Publ.
Hannula, S., Bloigu, R., Majamaa, K., Sorri, M., Mäki-Torkko, E., 2011. Self-reported hearing problems among older adults: prevalence and comparison to measured hearing impairment. J. Am. Acad. Audiol. 22 (08), 550–559.
Har-Shai Yahav, P., Sharaabi, A., Zion Golumbic, E., 2024. The effect of voice familiarity on attention to speech in a cocktail party scenario. Cereb. Cortex. 34 (1). https://doi.org/10.1093/cercor/bhad475.
Har-Shai Yahav, P., Zion Golumbic, E., 2021. Linguistic processing of task-irrelevant speech at a cocktail party. Elife 10. https://doi.org/10.7554/eLife.65096.
Harris, K.C., Dias, J.W., McClaskey, C.M., Rumschlag, J., Prisciandaro, J., Dubno, J.R., 2022. Afferent Loss, GABA, and Central Gain in Older Adults: Associations with Speech Recognition in Noise. J. Neurosci. 42 (38), 7201–7212. https://doi.org/10.1523/JNEUROSCI.0242-22.2022.
Hausfeld, L., Shiell, M., Formisano, E., Riecke, L., 2021. Cortical processing of distracting speech in noisy auditory scenes depends on perceptual demand. Neuroimage 228, 117670. https://doi.org/10.1016/j.neuroimage.2020.117670.
Healy, E.W., & Yoho, S.E. (2016). Difficulty understanding speech in noise by the hearing impaired: Underlying causes and technological solutions. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).
Heine, C., Browning, C., 2002. Communication and psychosocial consequences of sensory loss in older adults: overview and rehabilitation directions. Disabil. Rehabil. 24 (15), 763–773.
Helfer, K.S., Freyman, R.L., 2008. Aging and speech-on-speech masking. Ear Hear. 29 (1), 87–98.
Herrmann, B., 2025. Enhanced neural speech tracking through noise indicates stochastic resonance in humans. Elife 13. https://doi.org/10.7554/eLife.100830.
Herrmann, B., Butler, B.E., 2021. Hearing loss and brain plasticity: the hyperactivity phenomenon. Brain Struct. Funct. 226 (7), 2019–2039. https://doi.org/10.1007/s00429-021-02313-9.
Holt-Lunstad, J., Smith, T.B., Baker, M., Harris, T., Stephenson, D., 2015. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspectives on psychological science 10 (2), 227–237.
Hughes, L.F., Turner, J.G., Parrish, J.L., Caspary, D.M., 2010. Processing of broadband stimuli across A1 layers in young and aged rats. Hear. Res. 264 (1-2), 79–85. https://doi.org/10.1016/j.heares.2009.09.005.
Humes, L.E., Kidd, G.R., Lentz, J.J., 2013. Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. Front. Syst. Neurosci. 7, 55. https://doi.org/10.3389/fnsys.2013.00055.
Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., Giraud, A.L., 2015. Speech encoding by coupled cortical theta and gamma oscillations. Elife 4, e06213. https://doi.org/10.7554/eLife.06213.
Irino, T., Patterson, R.D., 2006. A Dynamic Compressive Gammachirp Auditory Filterbank. IEEe Trans. Audio Speech. Lang. Process. 14 (6), 2222–2232. https://doi.org/10.1109/TASL.2006.874669.
Jokinen, K., 1973. Presbyacusis: VI. Masking of speech. Acta Otolaryngol. 76 (1-6), 426–430.
Juarez-Salinas, D.L., Engle, J.R., Navarro, X.O., Recanzone, G.H., 2010. Hierarchical and serial processing in the spatial auditory cortical pathway is degraded by natural aging. J. Neurosci. 30 (44), 14795–14804. https://doi.org/10.1523/JNEUROSCI.3393-10.2010.
Kale, S., Heinz, M.G., 2010. Envelope coding in auditory nerve fibers following noise-induced hearing loss. J. Assoc. Res. Otolaryngol. 11 (4), 657–673. https://doi.org/10.1007/s10162-010-0223-6.
Kaplan-Neeman, R., Kishon-Rabin, L., Henkin, Y., Muchnik, C., 2006. Identification of syllables in noise: electrophysiological and behavioral correlates. J. Acoust. Soc. Am. 120 (2), 926–933. https://doi.org/10.1121/1.2217567.
Karunathilake, I.M.D., Dunlap, J.L., Perera, J., Presacco, A., Decruy, L., Anderson, S., Kuchinsky, S.E., Simon, J.Z., 2023. Effects of aging on cortical representations of continuous speech. J. Neurophysiol. 129 (6), 1359–1377. https://doi.org/10.1152/jn.00356.2022.
Keitel, A., Gross, J., Kayser, C., 2018. Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. PLoS. Biol. 16 (3), e2004473. https://doi.org/10.1371/journal.pbio.2004473.
Killion, M.C., Niquette, P.A., Gudmundsen, G.I., Revit, L.J., Banerjee, S., 2004. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 116 (4), 2395–2405.
King, P., 1954. Psychogenic deafness. The Journal of Laryngology & Otology 68 (9), 623–635.
Koike, K.J., Hurst, M.K., Wetmore, S.J., 1994. Correlation between the American Academy of Otolaryngology-Head and Neck Surgery five-minute hearing test and standard audiologic data. Otolaryngol. Head. Neck. Surg. 111 (5), 625–632. https://doi.org/10.1177/019459989411100514.
Kong, Y.Y., Mullangi, A., Ding, N., 2014. Differential modulation of auditory responses to attended and unattended speech in different listening conditions. Hear. Res. 316, 73–81. https://doi.org/10.1016/j.heares.2014.07.009.
Kopetzky, S.J., 1948. Deafness, Tinnitus, and Vertigo. Thomas Nelson & Sons.
Kurthen, I., Galbier, J., Jagoda, L., Neuschwander, P., Giroud, N., Meyer, M., 2021. Selective attention modulates neural envelope tracking of informationally masked speech in healthy older adults. Hum. Brain Mapp. 42 (10), 3042–3057.
Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. Hum. Brain Mapp. 8 (4), 194–208. https://doi.org/10.1002/(sici)1097-0193(1999)8:4<194::aid-hbm4>3.0.co;2-c.
Lalwani, P., Gagnon, H., Cassady, K., Simmonite, M., Peltier, S., Seidler, R.D., Taylor, S.F., Weissman, D.H., Polk, T.A., 2019. Neural distinctiveness declines with age in auditory cortex and is associated with auditory GABA levels. Neuroimage 201, 116033. https://doi.org/10.1016/j.neuroimage.2019.116033.
Lavie, L., Banai, K., Attias, J., Karni, A., 2014. How difficult is difficult? Speech perception in noise in the elderly hearing impaired. J. Basic Clin. Physiol. Pharmacol. 25 (3), 313–316.
Lesenfants, D., Vanthornhout, J., Verschueren, E., Decruy, L., Francart, T., 2019. Predicting individual speech intelligibility from the cortical tracking of acoustic- and phonetic-level speech representations. Hear. Res. 380, 1–9. https://doi.org/10.1016/j.heares.2019.05.006.
Lewandowsky, S., Oberauer, K., Yang, L.X., Ecker, U.K., 2010. A working memory test battery for MATLAB. Behav. Res. Methods 42 (2), 571–585.
Lin, F.R., Ferrucci, L., Metter, E.J., An, Y., Zonderman, A.B., Resnick, S.M., 2011. Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. Neuropsychology. 25 (6), 763–770. https://doi.org/10.1037/a0024238.
Liu, J., Stohl, J., Overath, T., 2024. Hidden hearing loss: Fifteen years at a glance. Hear. Res., 108967
Luts, H., Jansen, S., Dreschler, W., & Wouters, J. (2014). Development and normative data for the Flemish/Dutch Matrix test.

Marrufo-Perez, M.I., Lopez-Poveda, E.A., 2022. Adaptation to noise in normal and impaired hearing. J. Acoust. Soc. Am. 151 (3), 1741. https://doi.org/10.1121/10.0009802.
Matsumiya, Y., Tagliasco, V., Lombroso, C.T., Goodglass, H., 1972. Auditory evoked response: meaningfulness of stimuli and interhemispheric asymmetry. Science (1979) 175 (4023), 790–792. https://doi.org/10.1126/science.175.4023.790.
Mattys, S.L., Davis, M.H., Bradlow, A.R., Scott, S.K., 2012. Speech recognition in adverse conditions: a review. Lang. Cogn. Process. 27 (7-8), 953–978. https://doi.org/10.1080/01690965.2012.705006.
McClaskey, C.M., 2024. Neural hyperactivity and altered envelope encoding in the central auditory system: Changes with advanced age and hearing loss. Hear. Res. 442, 108945. https://doi.org/10.1016/j.heares.2023.108945.
McHaney, J.R., Gnanateja, G.N., Smayda, K.E., Zinszer, B.D., Chandrasekaran, B., 2021. Cortical Tracking of Speech in Delta Band Relates to Individual Differences in Speech in Noise Comprehension in Older Adults. Ear Hear. 42 (2), 343–354. https://doi.org/10.1097/AUD.0000000000000923.
Mesgarani, N., Chang, E.F., 2012. Selective cortical representation of attended speaker in multi-talker speech perception. Nature 485 (7397), 233–236.
Meyer, M., Keller, M., & Giroud, N. (2018). Suprasegmental speech prosody and the human brain. In.
Miller, G.A., Heise, G.A., Lichten, W., 1951. The intelligibility of speech as a function of the context of the test materials. J. Exp. Psychol. 41 (5), 329–335. https://doi.org/10.1037/h0062491.
Mirkovic, B., Debener, S., Schmidt, J., Jaeger, M., Neher, T., 2019. Effects of directional sound processing and listener's motivation on EEG responses to continuous noisy speech: Do normal-hearing and aided hearing-impaired listeners differ? Hear. Res. 377, 260–270. https://doi.org/10.1016/j.heares.2019.04.005.
Mohammadi, Y., Graversen, C., Ostergaard, J., Andersen, O.K., Reichenbach, T., 2023. Phase-locking of Neural Activity to the Envelope of Speech in the Delta Frequency Band Reflects Differences between Word Lists and Sentences. J. Cogn. Neurosci. 35 (8), 1301–1311. https://doi.org/10.1162/jocn_a_02016.
Moore, B.C., Oxenham, A.J., 1998. Psychoacoustic consequences of compression in the peripheral auditory system. Psychol. Rev. 105 (1), 108–124. https://doi.org/10.1037/0033-295x.105.1.108.
Moore, B.C.J., Wojtczak, M., Vickers, D.A., 1996. Effect of loudness recruitment on the perception of amplitude modulation. J. Acoust. Soc. Am. 100 (1), 481–489. https://doi.org/10.1121/1.415861.
Morrell, L.K., Salamy, J.G., 1971. Hemispheric asymmetry of electrocortical responses to speech stimuli. Science (1979) 174 (4005), 164–166. https://doi.org/10.1126/science.174.4005.164.
Muncke, J., Kuruvila, I., Hoppe, U., 2022. Prediction of Speech Intelligibility by Means of EEG Responses to Sentences in Noise. Front. Neurosci. 16, 876421. https://doi.org/10.3389/fnins.2022.876421.
Nasreddine, Z.S., Phillips, N.A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H., 2005. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J. Am. Geriatr. Soc. 53 (4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x.
Nielsen, J.B., Dau, T., 2009. Development of a Danish speech intelligibility test. Int. J. Audiol. 48 (10), 729–741.
Nilsson, M., Soli, S.D., Sullivan, J.A., 1994. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J. Acoust. Soc. Am. 95 (2), 1085–1099.
Nourski, K.V., Steinschneider, M., Rhone, A.E., Berger, J.I., Dappen, E.R., Kawasaki, H., Howard III, M.A., 2024. Intracranial electrophysiology of spectrally degraded speech in the human cortex. Front. Hum. Neurosci. 17, 1334742.
Obleser, J., Kayser, C., 2019. Neural Entrainment and Attentional Selection in the Listening Brain. Trends. Cogn. Sci. 23 (11), 913–926. https://doi.org/10.1016/j.tics.2019.08.004.
Obleser, J., Kotz, S.A., 2010. Expectancy constraints in degraded speech modulate the language comprehension network. Cereb. Cortex. 20 (3), 633–640. https://doi.org/10.1093/cercor/bhp128.
Ohlenforst, B., Zekveld, A.A., Jansma, E.P., Wang, Y., Naylor, G., Lorens, A., Lunner, T., Kramer, S.E., 2017. Effects of Hearing Impairment and Hearing Aid Amplification on Listening Effort: a Systematic Review. Ear Hear. 38 (3), 267–281. https://doi.org/10.1097/AUD.0000000000000396.
Panela, R.A., Copelli, F., Herrmann, B., 2024. Reliability and generalizability of neural speech tracking in younger and older adults. Neurobiol. Aging 134, 165–180. https://doi.org/10.1016/j.neurobiolaging.2023.11.007.
Pang, J., Beach, E.F., Gilliver, M., Yeend, I., 2019. Adults who report difficulty hearing speech in noise: an exploration of experiences, impacts and coping strategies. Int. J. Audiol. 58 (12), 851–860. https://doi.org/10.1080/14992027.2019.1670363.
Parthasarathy, A., Hancock, K.E., Bennett, K., DeGruttola, V., Polley, D.B., 2020. Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. Elife 9. https://doi.org/10.7554/eLife.51419.
Patterson, R., Nimmo-Smith, I., Holdsworth, J., & Rice, P. (1988). An efficient auditory filterbank based on the gammatone function.
Patterson, R.D., Allerhand, M.H., Giguere, C., 1995. Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform. J. Acoust. Soc. Am. 98 (4), 1890–1894. https://doi.org/10.1121/1.414456.
Peelle, J.E., Troiani, V., Wingfield, A., Grossman, M., 2010. Neural processing during older adults' comprehension of spoken sentences: age differences in resource allocation and connectivity. Cereb. Cortex. 20 (4), 773–782. https://doi.org/10.1093/cercor/bhp142.
Perron, M., Theaud, G., Descoteaux, M., Tremblay, P., 2021. The frontotemporal organization of the arcuate fasciculus and its relationship with speech perception in young and older amateur singers and non-singers. Hum. Brain Mapp. 42 (10), 3058–3076. https://doi.org/10.1002/hbm.25416.
Petersen, E.B., Wostmann, M., Obleser, J., Lunner, T., 2017. Neural tracking of attended versus ignored speech is differentially affected by hearing loss. J. Neurophysiol. 117 (1), 18–27. https://doi.org/10.1152/jn.00527.2016.
Pichora-Fuller, M.K., Kramer, S.E., Eckert, M.A., Edwards, B., Hornsby, B.W., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C.L., Naylor, G., Phillips, N.A., Richter, M., Rudner, M., Sommers, M.S., Tremblay, K.L., Wingfield, A., 2016. Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). Ear Hear. 37 (Suppl 1), 5S–27S. https://doi.org/10.1097/AUD.0000000000000312.
Pickering, M.J., Gambi, C., 2018. Predicting while comprehending language: a theory and review. Psychol. Bull. 144 (10), 1002–1044. https://doi.org/10.1037/bul0000158.
Picou, E.M., Ricketts, T.A., 2018. The relationship between speech recognition, behavioural listening effort, and subjective ratings. Int. J. Audiol. 57 (6), 457–467. https://doi.org/10.1080/14992027.2018.1431696.
Presacco, A., Simon, J.Z., Anderson, S., 2016a. Effect of informational content of noise on speech representation in the aging midbrain and cortex. J. Neurophysiol. 116 (5), 2356–2367.
Presacco, A., Simon, J.Z., Anderson, S., 2016b. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. J. Neurophysiol. 116 (5), 2346–2355.
Presacco, A., Simon, J.Z., Anderson, S., 2019. Speech-in-noise representation in the aging midbrain and cortex: Effects of hearing loss. PLoS. One 14 (3), e0213899. https://doi.org/10.1371/journal.pone.0213899.
Pryce, H., 2006. The process of coping in King-Kopetzky syndrome. Audiol. Med. 4 (2), 60–67.
Puschmann, S., Baillet, S., Zatorre, R.J., 2019. Musicians at the Cocktail Party: Neural Substrates of Musical Training During Selective Listening in Multispeaker Situations. Cereb. Cortex. 29 (8), 3253–3265. https://doi.org/10.1093/cercor/bhy193.
Puts, N.A., Edden, R.A., 2012. In vivo magnetic resonance spectroscopy of GABA: a methodological review. Prog. Nucl. Magn. Reson. Spectrosc. 60, 29–41. https://doi.org/10.1016/j.pnmrs.2011.06.001.
Quist-Hanssen, S., Thorud, E., Aasand, G., 1978. Noise-induced hearing loss and the comprehension of speech in noise. Acta Otolaryngol. 86 (sup360), 90–95.
Riecke, L., Sack, A.T., Schroeder, C.E., 2015. Endogenous Delta/Theta Sound-Brain Phase Entrainment Accelerates the Buildup of Auditory Streaming. Curr. Biol. 25 (24), 3196–3201. https://doi.org/10.1016/j.cub.2015.10.045.
Ronnberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., Dahlstrom, O., Signoret, C., Stenfelt, S., Pichora-Fuller, M.K., Rudner, M., 2013. The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. Front. Syst. Neurosci. 7, 31. https://doi.org/10.3389/fnsys.2013.00031.
Ross, B., Dobri, S., Schumann, A., 2020. Speech-in-noise understanding in older age: The role of inhibitory cortical responses. Eur. J. Neurosci. 51 (3), 891–908. https://doi.org/10.1111/ejn.14573.
Roth, W.T., Kopell, B.S., Bertozzi, P.E., 1970. The effect of attention on the average evoked response to speech sounds. Electroencephalogr. Clin. Neurophysiol. 29 (1), 38–46. https://doi.org/10.1016/0013-4694(70)90077-5.
Ruggero, M.A., Rich, N.C., 1991. Furosemide alters organ of corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane. J. Neurosci. 11 (4), 1057–1067. https://doi.org/10.1523/JNEUROSCI.11-04-01057.1991.
Rumschlag, J.A., McClaskey, C.M., Dias, J.W., Kerouac, L.B., Noble, K.V., Panganiban, C., Lang, H., Harris, K.C., 2022. Age-related central gain with degraded neural synchrony in the auditory brainstem of mice and humans. Neurobiol. Aging 115, 50–59. https://doi.org/10.1016/j.neurobiolaging.2022.03.014.
Salvi, R., Sun, W., Ding, D., Chen, G.D., Lobarinas, E., Wang, J., Radziwon, K., Auerbach, B.D., 2016. Inner Hair Cell Loss Disrupts Hearing and Cochlear Function Leading to Sensory Deprivation and Enhanced Central Auditory Gain. Front. Neurosci. 10, 621. https://doi.org/10.3389/fnins.2016.00621.
Saunders, G.H., Haggard, M.P., 1989. The clinical assessment of obscure auditory dysfunction—1. Auditory and psychological factors. Ear Hear. 10 (3), 200–208.
Saunders, G.H., Haggard, M.P., 1992. The clinical assessment of C "Obscure Auditory Dysfunction"(OAD) 2. Case control analysis of determining factors. Ear Hear. 13 (4), 241–254.
Schmitt, R., Meyer, M., Giroud, N., 2022. Better speech-in-noise comprehension is associated with enhanced neural speech tracking in older adults with hearing impairment. Cortex 151, 133–146.
Schubert, J., Schmidt, F., Gehmacher, Q., Bresgen, A., Weisz, N., 2023. Cortical speech tracking is related to individual prediction tendencies. Cereb. Cortex. 33 (11), 6608–6619. https://doi.org/10.1093/cercor/bhac528.
Sergeyenko, Y., Lall, K., Liberman, M.C., Kujawa, S.G., 2013. Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. J. Neurosci. 33 (34), 13686–13694. https://doi.org/10.1523/JNEUROSCI.1783-13.2013.
Shapiro, M.T., Melnick, W., VerMeulen, V., 1972. Effects of modulated noise on speech intelligibility of people with sensorineural hearing loss. Annals of Otology, Rhinology & Laryngology 81 (2), 241–248.
Synigal, S.R., Teoh, E.S., Lalor, E.C., 2020. Including Measures of High Gamma Power Can Improve the Decoding of Natural Speech From EEG. Front. Hum. Neurosci. 14, 130. https://doi.org/10.3389/fnhum.2020.00130.
Teng, X., Tian, X., Doelling, K., Poeppel, D., 2018. Theta band oscillations reflect more than entrainment: behavioral and neural evidence demonstrates an active chunking process. Eur. J. Neurosci. 48 (8), 2770–2782. https://doi.org/10.1111/ejn.13742.
Tremblay, K.L., Pinto, A., Fischer, M.E., Klein, B.E., Klein, R., Levy, S., Tweed, T.S., Cruickshanks, K.J., 2015. Self-reported hearing difficulties among adults with normal audiograms: The Beaver Dam Offspring Study. Ear Hear. 36 (6), e290–e299.

Tremblay, P., Brisson, V., Deschamps, I., 2021. Brain aging and speech perception: Effects of background noise and talker variability. Neuroimage 227, 117675. https://doi.org/10.1016/j.neuroimage.2020.117675.
Tremblay, P., Perron, M., Deschamps, I., Kennedy-Higgins, D., Houde, J.C., Dick, A.S., Descoteaux, M., 2018. The role of the arcuate and middle longitudinal fasciculi in speech perception in noise in adulthood. Hum. Brain Mapp. https://doi.org/10.1002/hbm.24367.
Vaden Jr., K.I., Kuchinsky, S.E., Ahlstrom, J.B., Dubno, J.R., Eckert, M.A, 2015. Cortical activity predicts which older adults recognize speech in noise and when. J. Neurosci. 35 (9), 3929–3937. https://doi.org/10.1523/JNEUROSCI.2908-14.2015.
Van den Noort, M., Bosch, P., Haverkort, M., Hugdahl, K., 2008. A standard computerized version of the reading span test in different languages. European Journal of Psychological Assessment 24 (1), 35–42.
Vander Ghinst, M., Bourguignon, M., Niesen, M., Wens, V., Hassid, S., Choufani, G., Jousmaki, V., Hari, R., Goldman, S., De Tiege, X., 2019. Cortical Tracking of Speech-in-Noise Develops from Childhood to Adulthood. J. Neurosci. 39 (15), 2938–2950. https://doi.org/10.1523/JNEUROSCI.1732-18.2019.
Vander Ghinst, M., Bourguignon, M., Op de Beeck, M., Wens, V., Marty, B., Hassid, S., Choufani, G., Jousmaki, V., Hari, R., Van Bogaert, P., Goldman, S., De Tiege, X., 2016. Left Superior Temporal Gyrus Is Coupled to Attended Speech in a Cocktail-Party Auditory Scene. J. Neurosci. 36 (5), 1596–1606. https://doi.org/10.1523/JNEUROSCI.1730-15.2016.
Vander Ghinst, M., Bourguignon, M., Wens, V., Naeije, G., Ducene, C., Niesen, M., Hassid, S., Choufani, G., Goldman, S., De Tiege, X., 2021. Inaccurate cortical tracking of speech in adults with impaired speech perception in noise. Brain Commun. 3 (3), fcab186. https://doi.org/10.1093/braincomms/fcab186.
Vanthornhout, J., Decruy, L., Wouters, J., Simon, J.Z., Francart, T., 2018. Speech Intelligibility Predicted from Neural Entrainment of the Speech Envelope. J. Assoc. Res. Otolaryngol. 19 (2), 181–191. https://doi.org/10.1007/s10162-018-0654-z.
Verschueren, E., Vanthornhout, J., Francart, T., 2020. The Effect of Stimulus Choice on an EEG-Based Objective Measure of Speech Intelligibility. Ear Hear. 41 (6), 1586–1597. https://doi.org/10.1097/AUD.0000000000000875.
Viana, L.M., O'Malley, J.T., Burgess, B.J., Jones, D.D., Oliveira, C.A., Santos, F., Merchant, S.N., Liberman, L.D., Liberman, M.C., 2015. Cochlear neuropathy in human presbycusis: Confocal analysis of hidden hearing loss in post-mortem tissue. Hear. Res. 327, 78–88. https://doi.org/10.1016/j.heares.2015.04.014.
Wagener, K. (1998). Entwicklung und Evaluation eines Satztests für die deutsche Sprache.
Wagener, K., Brand, T., Kollmeier, B., 1999. Entwicklung und Evaluation eines Satztests für die deutsche Sprache. I-III: Design, Optimierung und Evaluation des Oldenburger Satztests (Development and evaluation of a sentence test for the German language. I-III: design, optimization and evaluation of the Oldenburg sentence test). Zeitschrift für Audiologie (Audiological Acoustics) 38, 4–15.
Wallaert, N., Moore, B.C., Ewert, S.D., Lorenzi, C., 2017. Sensorineural hearing loss enhances auditory sensitivity and temporal integration for amplitude modulation. J. Acoust. Soc. Am. 141 (2), 971. https://doi.org/10.1121/1.4976080.
Wang, B., Xu, X., Niu, Y., Wu, C., Wu, X., Chen, J., 2023. EEG-based auditory attention decoding with audiovisual speech for hearing-impaired listeners. Cereb. Cortex. 33 (22), 10972–10983. https://doi.org/10.1093/cercor/bhad325.
Wang, L., Wu, E.X., Chen, F., 2020. Robust EEG-Based Decoding of Auditory Attention With High-RMS-Level Speech Segments in Noisy Conditions. Front. Hum. Neurosci. 14, 557534. https://doi.org/10.3389/fnhum.2020.557534.
Wang, X., Xu, L., 2021. Speech perception in noise: Masking and unmasking. J. Otol. 16 (2), 109–119.
Weintraub, S., Bauer, P.J., Zelazo, P.D., Wallner-Allen, K., Dikmen, S.S., Heaton, R.K., Tulsky, D.S., Slotkin, J., Blitz, D.L., Carlozzi, N.E., Havlik, R.J., Beaumont, J.L., Mungas, D., Manly, J.J., Borosh, B.G., Nowinski, C.J., Gershon, R.C., 2013. I. NIH Toolbox Cognition Battery (CB): introduction and pediatric data. Monogr. Soc. Res. Child Dev. 78 (4), 1–15. https://doi.org/10.1111/mono.12031.
Whiting, K.A., Martin, B.A., Stapells, D.R., 1998. The Effects of Broadband Noise Masking on Cortical Event-Related Potentials to Speech Sounds /ba/and /da. Ear Hear. 19 (3), 218–231. https://journals.lww.com/ear-hearing/fulltext/1998/06000/the_effects_of_broadband_noise_masking_on_cortical.5.aspx.
Wild, C.J., Yusuf, A., Wilson, D.E., Peelle, J.E., Davis, M.H., Johnsrude, I.S., 2012. Effortful listening: the processing of degraded speech depends critically on attention. J. Neurosci. 32 (40), 14010–14021. https://doi.org/10.1523/JNEUROSCI.1528-12.2012.
Wingfield, A., Grossman, M., 2006. Language and the aging brain: patterns of neural compensation revealed by functional brain imaging. J. Neurophysiol. 96 (6), 2830–2839. https://doi.org/10.1152/jn.00628.2006.
Wong, P.C., Ettlinger, M., Sheppard, J.P., Gunasekera, G.M., Dhar, S., 2010. Neuroanatomical characteristics and speech perception in noise in older adults. Ear Hear. 31 (4), 471–479. https://doi.org/10.1097/AUD.0b013e3181d709c2.
Wong, P.C., Jin, J.X., Gunasekera, G.M., Abel, R., Lee, E.R., Dhar, S., 2009. Aging and cortical mechanisms of speech perception in noise. Neuropsychologia 47 (3), 693–703. https://doi.org/10.1016/j.neuropsychologia.2008.11.032.
Wong, P.C., Uppunda, A.K., Parrish, T.B., Dhar, S., 2008. Cortical mechanisms of speech perception in noise. J. Speech. Lang. Hear. Res. 51 (4), 1026–1041. https://doi.org/10.1044/1092-4388(2008/075), 51/4/1026 [pii].
Wu, P.Z., Liberman, L.D., Bennett, K., de Gruttola, V., O'Malley, J.T., Liberman, M.C., 2019. Primary Neural Degeneration in the Human Cochlea: Evidence for Hidden Hearing Loss in the Aging Ear. Neuroscience 407, 8–20. https://doi.org/10.1016/j.neuroscience.2018.07.053.
Yang, X., Wang, K., Shamma, S.A., 1992. Auditory representations of acoustic signals. IEEe Trans. Inf. Theory. 38 (2), 824–839. https://doi.org/10.1109/18.119739.
Yasmin, S., Irsik, V.C., Johnsrude, I.S., Herrmann, B., 2023. The effects of speech masking on neural tracking of acoustic and semantic features of natural speech. Neuropsychologia 186, 108584. https://doi.org/10.1016/j.neuropsychologia.2023.108584.
Zan, P., Presacco, A., Anderson, S., Simon, J.Z., 2020. Exaggerated cortical representation of speech in older listeners: mutual information analysis. J. Neurophysiol. 124 (4), 1152–1164. https://doi.org/10.1152/jn.00002.2020.
Zekveld, A.A., Rudner, M., Johnsrude, I.S., Festen, J.M., van Beek, J.H., Ronnberg, J., 2011. The influence of semantically related and unrelated text cues on the intelligibility of sentences in noise. Ear Hear. 32 (6), e16–e25. https://doi.org/10.1097/AUD.0b013e318228036a.
Zhang, X., Li, J., Li, Z., Hong, B., Diao, T., Ma, X., Nolte, G., Engel, A.K., Zhang, D., 2023. Leading and following: Noise differently affects semantic and acoustic processing during naturalistic speech comprehension. Neuroimage 282, 120404. https://doi.org/10.1016/j.neuroimage.2023.120404.
Zhu, J., Garcia, E., 1999. The Wechsler Abbreviated Scale of Intelligence (WASI). Psychological Corporation, New York.
Zimmermann, P., Fimm, B., 2002. A test battery for attentional performance. Applied Neuropsychology of Attention: Theory. Diagnosis and Rehabilitation 110–151.
Zinszer, B.D., Yuan, Q., Zhang, Z., Chandrasekaran, B., Guo, T., 2022. Continuous speech tracking in bilinguals reflects adaptation to both language and noise. Brain Lang. 230, 105128. https://doi.org/10.1016/j.bandl.2022.105128.
Zoefel, B., VanRullen, R., 2016. EEG oscillations entrain their phase to high-level features of speech sound. Neuroimage 124 (Pt A), 16–23. https://doi.org/10.1016/j.neuroimage.2015.08.054.
Zou, J., Feng, J., Xu, T., Jin, P., Luo, C., Zhang, J., Pan, X., Chen, F., Zheng, J., Ding, N., 2019. Auditory and language contributions to neural encoding of speech features in noisy environments. Neuroimage 192, 66–75. https://doi.org/10.1016/j.neuroimage.2019.02.047.