Cervical transcutaneous vagus nerve stimulation enhances speech recognition in noise: A crossover, placebo-controlled study

Michael Jigo[1], Jason B. Carmel[1,2], Qi Wang[1,3], & Charles Rodenkirch[1,4]

[1]Sharper Sense, Inc., New York, NY
[2]Department of Neurology and Orthopedics, Columbia University Medical Center, New York, NY
[3]Department of Biomedical Engineering, Columbia University, New York, NY
[4]The Jacobs Technion-Cornell Institute at Cornell Tech, New York, NY

Keywords. speech in noise, aging, vagus nerve stimulation, locus coeruleus, norepinephrine, sensory perception, hearing loss

Abstract

Understanding speech in noisy environments is a widespread challenge, even for people with clinically normal hearing and users of hearing aids and cochlear implants. While these devices improve speech audibility, they do not optimize the brain's ability to isolate relevant speech from background noise. Tonic vagus nerve stimulation (VNS) has been shown to rapidly improve sensory processing in the brain and enhance auditory perception in humans. Here, we investigated whether tonic cervical transcutaneous VNS (ctVNS) can enhance speech-in-noise recognition. Two cohorts of older human adults (60-84 years) completed audiological speech-in-noise assessments, QuickSIN (n=16) or AzBio (n=14), while receiving tonic ctVNS or placebo stimulation to the neck-shoulder junction in a counterbalanced, crossover fashion. Participants repeated sentences spoken amid various levels of background babble. Relative to placebo stimulation, tonic ctVNS improved the signal-to-noise ratio for reliable speech perception (SNR-50) by an average of 0.72 dB in QuickSIN (p=0.033) and 0.38 dB in AzBio (p=0.045), and enhanced word recognition accuracy by 4% (QuickSIN: p=0.0066, AzBio: p=0.000025). Notably, ctVNS benefits exceeded minimum clinically important differences in 39% of participants on average. Improvements by tonic ctVNS were most pronounced in individuals with speech-in-noise deficits (p=0.0089), peaked at noise levels typical of real-world environments (p=0.029), and did not depend on hearing loss severity (p=0.97) or age (p=0.88). Our findings demonstrate that tonic ctVNS delivers immediate, clinically meaningful improvements in speech-in-noise recognition, highlighting its potential to complement traditional assistive hearing technologies and inform novel therapies for sensory processing disorders.

Introduction

Speech-in-noise (SIN) recognition poses a significant challenge, particularly for individuals in high-noise environments, older adults, and those with hearing loss [1, 2]—a group comprising over 1.5 billion people worldwide [3]. Even among individuals with clinically normal hearing, 10-12% report difficulties understanding speech in noisy settings [4-6]**.** These widespread challenges underscore the critical need for methods that enhance SIN recognition.

SIN perception requires an interplay between the auditory periphery's ability to detect incoming sounds and the brain's capacity to isolate speech amidst competing noise [2, 7, 8]. While hearing aids, cochlear implants, and over-the-counter (OTC) sound amplifiers improve speech audibility by enhancing peripheral sensitivity, most users continue to face significant challenges in noisy environments [9-11]. Neuromodulation techniques that directly target auditory processing in the brain, therefore, present a potential alternative to traditional approaches.

Neuromodulation via tonic vagus nerve stimulation (VNS) can enhance central auditory processing [12-14]. Our previous research in rats demonstrated that tonic VNS rapidly improved somatosensory acuity by activating the locus coeruleus (LC)-norepinephrine (NE) system through continuous stimulation of the nucleus tractus solitarius pathway [12, 13]. This tonic activation elevated NE concentration, which optimized intrathalamic circuits for the transmission of sensory information and improved perceptual sensitivity [12, 15]. These sensory enhancements occurred through NE-mediated suppression of T-type calcium channels in thalamic sensory relay neurons, which mediate burst spiking activity that normally impairs central sensory processing [12, 15]. These preclinical findings translated successfully to humans, where tonic cervical transcutaneous VNS (ctVNS), applied to the left side of the neck, enhanced auditory performance in a gap discrimination task [14], which required the same precise temporal processing in the brain essential for SIN recognition [16-18].

Here, we tested the hypothesis that tonic ctVNS enhances SIN recognition. We leveraged age-related changes in older adults' auditory processing to examine whether ctVNS interacts with age, the severity of hearing loss, or SIN deficits. Participants listened to and repeated hundreds of pre-recorded words in standard SIN assessments—QuickSIN [19] or AzBio [20]—while receiving tonic ctVNS or placebo stimulation in a counterbalanced, crossover fashion.

To preview our results, tonic ctVNS improved the signal-to-noise ratio (SNR) for reliable speech recognition by 12%, increased word recognition accuracy by 4%, and achieved or surpassed the minimum clinically important difference (MCID) in 39% of participants on average.

Methods

All participants passed the abbreviated Montreal Cognitive Assessment [21], provided written informed consent, and were compensated $30/hour. Protocols complied with the Declaration of Helsinki and were approved by Cornell University's Institutional Review Board for Human Participant Research.

Protocol

Stimulation sites and equipment. Hydrogel electrodes (diameter: 1"; PALS, Axelgaard Manufacturing, CA, USA) were positioned on the skin's surface with a center-to-center distance constrained by participants' neck/shoulder morphology. For ctVNS, electrodes were placed within the left carotid triangle—lateral to the larynx, medial to and oriented parallel to the sternocleidomastoid muscle. For off-target stimulation, electrodes were positioned on the neck-shoulder junction over the trapezius muscle where the vagus nerve is absent. Hypoallergenic sanitary wipes cleaned sites before electrode placement.

Study design. A placebo-controlled, single-blind, crossover design was followed. Participants received ctVNS and off-target stimulation in a counterbalanced order, separated by a washout period of 28.64 minutes (95% CI: [25.25, 32.04]). Sessions were completed on a single day in the morning (9am-12pm) or afternoon (1pm-6pm).

Stimulation parameters. Tonic, non-burst, stimulation was delivered continuously (100% duty cycle; constant current) using an FDA-cleared neuromuscular stimulator (Mettler Sys*Stim 240, Mettler Electronics Corporation, CA, USA). Biphasic square pulses (100 μs/phase; 200 μs total) were delivered at 30 Hz with current amplitudes determined separately for each site and participant. Current was set to the maximum level that was perceptible, but did not elicit pain or muscle contractions.

Intensity rating. At the end of each stimulation block, and before stimulation offset, participants responded to the prompt: "At this moment, how intense does the stimulation sensation feel?" with options: 1=not intense, 2=slightly, 3=moderately, 4=very, 5=extremely intense.

Sample

Demographics. One cohort completed QuickSIN [19] (n=16, 11 women; age=69.50 years; SD, 7.87), another completed AzBio [20] (n=14, 9 women; age, 71.19 years; SD, 6.70). Participants were naïve to the study's purpose and were asked to refrain from stimulants (caffeine, nicotine) for three hours prior. Two individuals in the QuickSIN cohort wore hearing aids. One left-handed individual participated in the AzBio cohort. Inclusion criteria were: 60-85 years old, no history of neurological or psychiatric disorders, no history of substance abuse, not taking beta adrenergic blockers, no medical implants, no severe cardiac disorders, not currently pregnant, and no prior abnormalities or surgeries involving the neck or vagus nerve. No adverse events occurred.

Power analysis. While a sample size of 12 was sufficient to demonstrate statistically significant ctVNS-evoked sensory enhancements in younger adults [14], we used a simulation-based approach [22] to determine sample sizes with ≥95% power for SIN perception. The smallest assumed change (SAC) we could observe was informed by studies deploying SIN assessments (QuickSIN or AzBio) before and during sound amplification. For QuickSIN's SNR-50, the SAC was 1.14 dB, the median improvement after two weeks of hearing aid use [23]. For AzBio, SAC for accuracy was 2.5%, based on improvements by an OTC sound amplifier [24]. To estimate the variance, six younger adults (5 women; age, 27.5 years; SD, 6.12) completed QuickSIN in a pilot that mimicked the present study in design and analyses. After 10,000 Monte Carlo simulations and assuming a 5% false positive rate, a sample of 16 had 95% power to detect the SAC for SNR-50 in QuickSIN, and 14 had 97% power to detect the SAC for Accuracy in AzBio.

Speech-in-noise measures

QuickSIN. One practice list of sentences was completed by participants. Then, 5.4 unique lists (SD=0.81; 30 keywords/list [25]) were played in a randomized order per stimulation site. A pre-recorded female talker read sentences at six signal-to-noise ratios (SNR) with four-talker babble (25 to 0 dB, 5 dB steps). Target speech was presented at 70 dBA.

AzBio. We used AzBio Ver. 2.0 [26], which included 15 equivalent lists, each with 20 unique sentences spoken by two male and two female talkers at 60 dBA amid ten-talker babble (-5 to 15 dB SNR, 5 dB steps). List 15 was used for practice and the remaining lists were ordered using a balanced Latin square design to control for first-order carryover effects during the main

study. SNRs were fixed within and randomized between lists. Recommended SNRs (5, 10 dB) [27] were tested twice, totaling seven lists per stimulation site.

Accuracy & SNR-50. Word recognition accuracy was measured as the proportion of words repeated correctly. Following the procedure used in standard audiological testing [28, 29], we used Spearman-Kärber equations to transform these proportions into an estimate of SNR-50. QuickSIN is designed to use a simplified equation [25] whereas for AzBio, we used a generalized form [28].

Psychometric function. Three-parameter logistic functions were fit to accuracy as a function of SNR using quickpsy [30]: $p(SNR)=(1-\lambda) / (1+exp(-\beta[SNR-SNR\alpha])$. SNRα defined its midpoint at low SNRs, λ defined its upper asymptote at high SNRs, and β defined its steepness [31]. Pearson's χ^2 determined goodness-of-fit [32].

Apparatus

Sound field. Participants sat alone in a sound-treated room. Psychtoolbox-3 [33] delivered stimuli using MATLAB running on Windows 10. Speech-in-noise audio were delivered via a single loudspeaker (308P MkII, JBL) connected to a digital-to-analogue converter and amplifier (Atom DAC+ and Atmp Amp+, JDS Labs) calibrated for a flat frequency response (0.25-8 kHz) using a measurement microphone (UMIK-2, miniDSP) and calibration routines in Room EQ Wizard (www.roomeqwizard.com). The loudspeaker faced participants one meter away (0° azimuth). A microphone (FIFINE K669B) connected to a virtual audio mixer (VB-Audio VoiceMeeter) streamed responses to an experimenter seated outside the sound field.

Audiometry. A MAICO MA 40 audiometer (MAICO Diagnostics, Eden Prairie, MN) determined monaural thresholds at 0.5, 1, 2, and 4 kHz using the modified Hughson-Westlake procedure [34] and their average in the better ear (PTA4) defined hearing loss [3]. For QuickSIN participants, a customized procedure initially measured binaural thresholds in the sound field. A single-interval, yes-no, detection task was performed at the four frequencies; participants reported when they heard a tone. The QUEST method [35] adjusted tone volume (calibrated to dB HL [36]) to converge on 50% hit rates. The four-frequency average defined binaural thresholds (PTA-Binaural). To estimate PTA4 from PTA-Binaural, seven QuickSIN participants completed a separate pure tone audiometry session with the audiometer. Their PTA4 and PTA-Binaural thresholds were highly correlated ($r=0.91$, $p=0.0049$), and we used linear regression ($R^2=0.82$) to estimate PTA4 for the remaining QuickSIN participants: $PTA4 = 13.67 + 1.05(PTA\text{-}Binaural)$.

Statistics

All analyses were performed in R 4.4.0 [37]. Model formulas and coefficients are in Supplementary Materials.

Hierarchical modeling. Linear mixed-effects regression models (LME), implemented in lmerTest [38], modeled the within-participant change in performance between stimulation sites. Generally, LMEs modeled differences between stimulation sites as a fixed effect, and the variance among participants, sentence lists [39], SNRs, and/or SIN tests as random effects. Generalized LMEs with a binomial link function modeled changes in the percentage of correct responses. Conditional and marginal R^2 were estimated using MuMIn [40]. Bootstrapping estimated confidence intervals. P-values for these analyses were adjusted to control the false discovery rate using the Benjamini & Hochberg method [41] implemented in p.adjust [37].

A three-way multivariable LMEs evaluated whether SNR-50 during off-target stimulation, hearing loss (PTA4), or age predicted ctVNS effects; all variables were standardized to z-scores before multivariable modeling. Accuracy was not included as a predictor as it was highly correlated with SNR-50. ctVNS-evoked effects were quantified as the difference from off-target stimulation in the psychometric function's upper asymptote, which summarized changes at moderate-to-high SNRs. Permutation analyses corrected for regression to the mean (RTM) [42]: labels were shuffled between stimulation sites 10,000 times within cohorts and LMEs were fit to shuffled data, generating null distributions of t-values for each main effect and interaction. We additionally corrected for multiple comparisons in this analysis using the maxT method [43], wherein the maximum t-value across all comparisons in an iteration were concatenated into a single null distribution. Corrected p-values were derived as the proportion of null t-values larger than the observed values in the original, unshuffled dataset.

Sex. Post-hoc analyses of sex differences were performed. SIN metrics were pooled across cohorts to create a larger sample of males, then LMEs evaluated whether ctVNS efficacy interacted with sex.

Results

Two cohorts completed pure tone audiometry (PTA) followed by either QuickSIN [19, 25] (n=16) or AzBio [20] (n=14). During the SIN assessments, tonic ctVNS was delivered continuously on the left side of the neck and, after a washout period, placebo stimulation was delivered to an off-target site; the order of sites was counterbalanced (**Figure 1**). Biphasic square pulses were delivered continuously at 30 Hz, at a current intensity tailored to participants' comfort, and its perceived intensity was rated as "not intense" or "slightly intense" on average (**Table 1**). Participants repeated pre-recorded sentences read by a female talker (QuickSIN, 161 words/site, SD=24.20 words) or both male and female talkers (AzBio, 983 words/site, SD=14.52 words) amid babble (-5 to 25 dB SNR). SIN performance was assessed by two primary outcome measures, standards in audiological testing: SNR-50 and word recognition accuracy. Additionally, we examined two secondary measures, λ and SNRα, which reflect SIN performance at low and high noise levels, respectively.

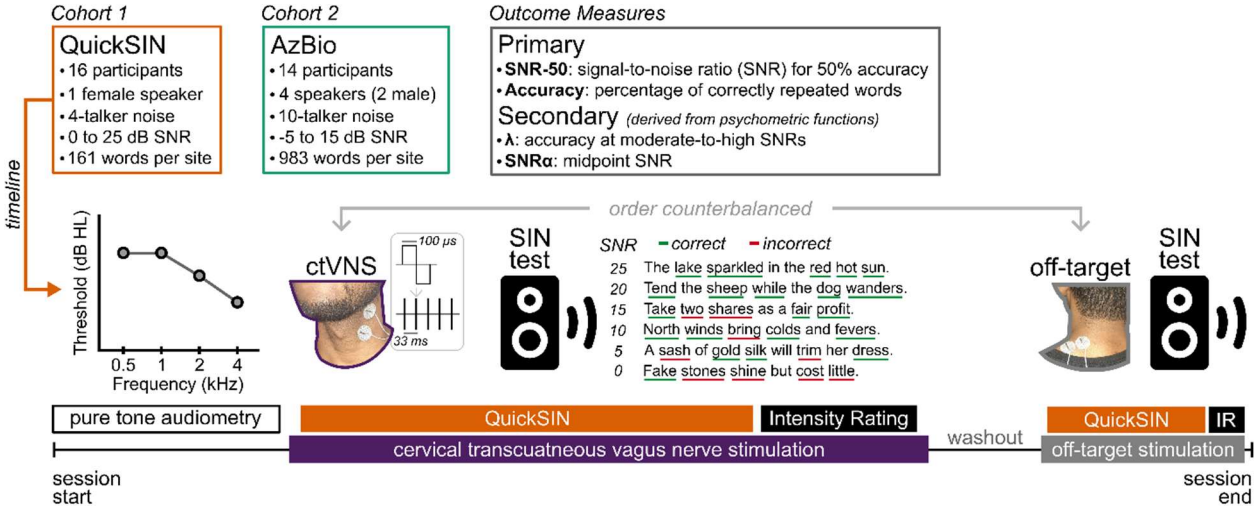

Figure 1. Placebo-controlled, crossover study design and outcome measures. (***top***) Characteristics of each cohort (number of participants, speech-in-noise parameters), along with the study's primary and secondary outcome measures. (***bottom***) Timeline showing events during a testing session for the QuickSIN cohort; an identical timeline was followed for the AzBio cohort. Pure tone audiometry was completed without electrical stimulation, followed by tonic ctVNS or off-target stimulation in a counterbalanced order; the stimulation waveform and frequency are shown beside "ctVNS". After each speech-in-noise assessment, participants rated the perceived stimulation intensity.

Table 1. Stimulation parameters and ratings of perceived intensity. Average values (bolded entries) and 95% confidence intervals (square brackets) are displayed. "Center-to-center" is the distance between electrodes. Intensity ratings were measured at the end of each stimulation block using a 5-point Likert scale (1=not intense; 5=extremely intense) to the prompt: "At this moment, how intense does the stimulation sensation feel?"

Test	Site	Current (mA)	Center-to-center (cm)	Duration (min)	Intensity rating (1=low, 5=high)
QuickSIN	ctVNS	**7.71** [6.80, 8.62]	**3.84** [3.55, 4.12]	**8.00** [7.29, 8.72]	**1.69** [1.44 2.00]
	off-target	**8.15** [7.40, 8.90]	**3.04** [2.97, 3.11]	**7.89** [7.23, 8.55]	**1.63** [1.19 2.19]
AzBio	ctVNS	**8.59** [7.97, 9.21]	**4.00** [3.76, 4.24]	**20.95** [19.73, 22.17]	**1.43** [1.07 1.86]
	off-target	**8.86** [8.18, 9.55]	**3.04** [3.00, 3.09]	**20.93** [19.95, 21.91]	**1.71** [1.29 2.21]

ctVNS improved speech-in-noise recognition thresholds

SNR-50, the signal-to-noise ratio threshold for 50% word recognition, was estimated with the Spearman-Kärber procedure, a standard method in SIN assessment [28, 29]. Lower values indicate better SIN perception. Linear mixed-effects models (; **Table S1**) showed that, relative to off-target stimulation, ctVNS reduced SNR-50 by 0.72 dB in QuickSIN (**Figure 2A**; SE=0.30, t(152.70)=-2.42, p=0.033) and by 0.38 dB in AzBio (**Figure 2B**; SE=0.17, t(13)=-2.22, p=0.045). These changes corresponded to a 11.51% and 13.27% improvement in SNR-50, respectively. There was no interaction between ctVNS-evoked improvements and participant sex (t(174.27)=0.13, p=0.90, **Table S2**).

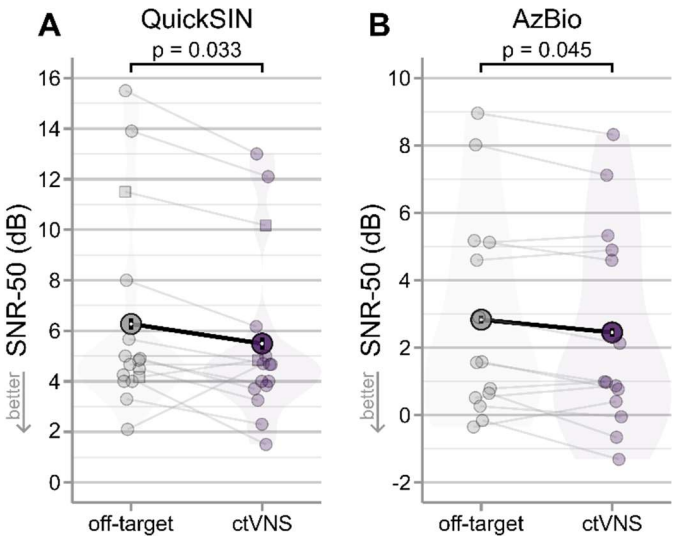

Figure 2. ctVNS improved speech-in-noise recognition thresholds. SNR-50 estimated with (**A**) QuickSIN and (**B**) AzBio. Lower values indicate heightened speech-in-noise recognition. Large data points connected by black lines show group means, with errors ±1 within-subject SEM [44]. Square data markers indicate the two individuals that wore hearing aids during QuickSIN. P-values are corrected for multiple comparisons using the Benjamini-Hochberg procedure.

ctVNS increased word recognition accuracy

The percentage of correctly repeated words was assessed with generalized LMEs (**Table S3**). Compared to off-target stimulation, ctVNS increased the log odds of accurate recognition by 0.24 in QuickSIN (**Figure 3A**; SE=0.086, z=2.81, p=0.0066) and by 0.15 in AzBio (**Figure 3B**; SE=0.035, z=4.39, p=0.000025), equating to a 3.92% (6 words) and 3.64% (36 words) increase in the number of words correctly recalled, respectively. These ctVNS-evoked enhancements did not interact with sex (z=1.49, p=0.14, **Table S2**). For AzBio, performance at audiologist-recommended SNRs (5, 10 dB) [27] were assessed separately (**Table S3**). ctVNS enhanced recognition accuracy by 0.14 log odds at 5 dB (**Figure 3C**; SE=0.055, z=2.52, p=0.019) and by 0.33 at 10 dB SNR (**Figure 3D**; SE=0.076, z=4.37, p=0.000025), corresponding to improvements of 2.42% and 2.14%, respectively.

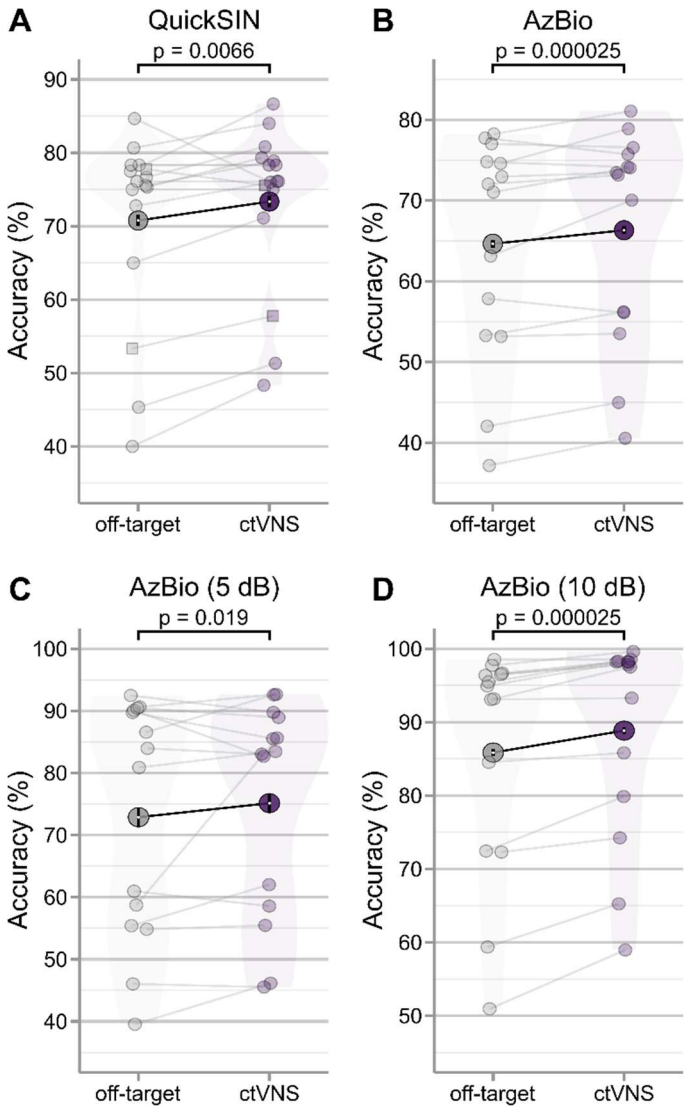

Figure 3. ctVNS increased speech-in-noise recognition accuracy. Mean percentage of words correctly recognized across all signal-to-noise ratios in (**A**) QuickSIN and (**B**) AzBio. Accuracy is shown separately for a subset of clinically relevant signal-to-noise ratios in the AzBio test, (**C**) 5 dB and (**D**) 10 dB. Plotting conventions follow those in Figure 2. All p-values are corrected for multiple comparisons using the Benjamini-Hochberg procedure.

ctVNS-evoked enhancements surpassed clinically meaningful benchmarks

The minimum clinically important differences (MCIDs) for SNR-50 and word recognition accuracy, based on benchmarks from FDA-approved cochlear implants, are 1 dB [45] and 4.9% [46], respectively. While the MCID for word recognition accuracy is specifically defined for a 5 dB SNR, **Table 2** reports all SNRs. In the QuickSIN cohort, ctVNS-evoked improvements met and exceeded the MCID for SNR-50 in 9 of 16 participants (56%) and for word recognition accuracy in 56% of participants at 5 dB SNR. In the AzBio cohort, 14% of participants had improvements that exceeded the MCID for SNR-50, while 29% exceeded the MCID for accuracy at 5 dB SNR. On average, tonic ctVNS produced clinically meaningful enhancements in SIN perception in 39% of participants.

Table 2. Tonic ctVNS-evoked enhancements surpassed clinically meaningful benchmarks. The percentage (%MCID) and number (#MCID) of participants from each cohort that met or exceeded the MCID for each primary outcome measure. Shading indicates the measures of primary interest.

Cohort	Measure		%MCID	#MCID
QuickSIN	SNR-50		56	9
	Accuracy	0 dB	38	6
		5 dB	56	9
		10 dB	38	6
		15 dB	50	8
		20 dB	31	5
		25 dB	44	7
AzBio	SNR-50		14	2
	Accuracy	-5 dB	7.1	1
		0 dB	36	5
		5 dB	29	4
		10 dB	21	3
		15 dB	14	2

ctVNS-evoked enhancements increased with SNR

SIN accuracy follows a sigmoidal psychometric function with SNR (**Figure 4**) defined by two key parameters [31]: SNRα, the signal-to-noise ratio at its midpoint, and λ, its upper asymptote at high SNRs. Logistic functions approximated the functions well (QuickSIN: $\chi^2>12$, $p>0.067$; AzBio: all $\chi^2=20$, $p=1$). On average, ctVNS increased the upper asymptote in QuickSIN by 3.74% (**Figure 4A**; SE=1.43, $t(15)=2.62$, $p=0.029$) and in AzBio by 2.43% (**Figure 4B**; SE=0.99, $t(13)=-2.45$, $p=0.029$). This translated to ctVNS-evoked enhancements that interacted with SNR (**Figure 4C**, **Table S5**), with word recognition accuracy increasing by 0.030 log odds per SNR in QuickSIN (SE=0.012, $z=2.50$, $p=0.012$) and by 0.027 log odds in AzBio (SE=0.0067, $z=3.96$, $p=0.00015$). In contrast, ctVNS did not alter SNRα in QuickSIN ($t(15)=-0.066$, $p=0.95$) or AzBio ($t(13)=0.15$, $p=0.95$; **Table S4**).

To link ctVNS efficacy to SNRs encountered in real-world settings [47-49], we analyzed performance according to categories derived from older adults' ratings of noise intensities [50] (**Figure 4C**). SNRs were pooled within noise categories and across SIN tests, revealing that tonic ctVNS significantly improved accuracy at moderate-to-high SNRs rated as "noisy" (5-10 dB SNR; z=4.62, p=0.000012) and "quiet" (>10 dB SNR; z=3.41, p=0.00099), but not at "very noisy" SNR levels (≤0 dB; z=0.64, p=0.53, **Table S6**).

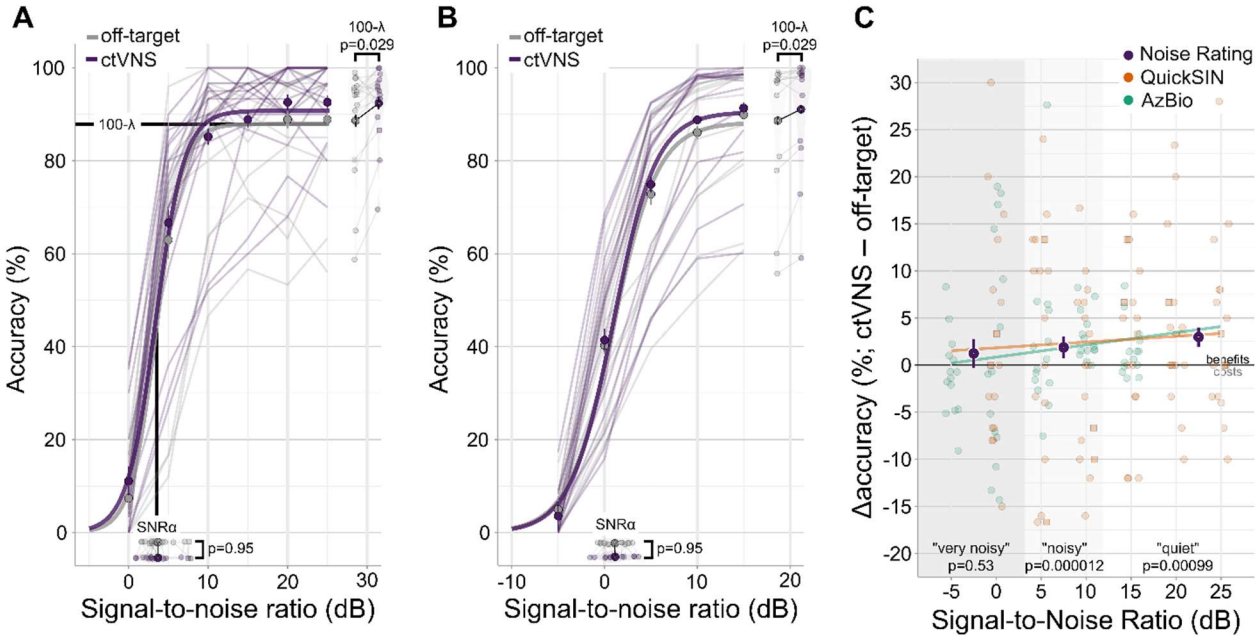

Figure 4. ctVNS-evoked enhancements increased with SNR. Psychometric functions for (**A**) QuickSIN and (**B**) AzBio cohorts. Smooth lines depict the fitted logistic functions to each stimulation site, while the transparent, jagged lines show the raw accuracy for individual participants. Parameter estimates for SNRα and λ are displayed beside the x- and y-axes, respectively, with large points connected by black lines indicating the group average and error bars of ±1 within-subject SEM[44]. P-values correspond to the result of linear mixed-effects modeling. (**C**) ctVNS-evoked performance changes as a function of SNR. Orange and green points depict the accuracy difference between off-target and ctVNS in QuickSIN and AzBio, respectively, while the lines show the best-fitting LME model. Purple points show the ctVNS effect pooled within noise rating categories and across tests. The shading corresponds to a noise category "very noisy" (≤0 dB), "noisy" (5-10 dB), or "quiet" (>10 dB). All p-values were corrected for multiple comparisons using the Benjamini-Hochberg procedure.

Enhancements grew with the severity of SIN deficits, but neither hearing loss nor age

Baseline SNR-50 (during off-target stimulation) varied between participants. Using normative SNR-50 data for QuickSIN [25], 69% of QuickSIN participants had normal or near-normal SIN recognition, 12% were mildly impaired, and 19% were moderately impaired. Although equivalent categories are less established for AzBio, a higher SNR-50 indicates greater SIN deficits. Similarly, hearing loss—defined as the pure tone average (PTA4) in the better ear across four frequencies (0.5, 1, 2, and 4 kHz) according to World Health Organization guidelines [3]—varied between participants: 27% had normal hearing, 60% had mild, 10% had moderate, and 3% had moderately severe hearing loss. SIN deficits have been shown to correlate with hearing loss and age, which we replicated (**Figure S1** and **Tables S8-S9**).

In an exploratory analysis, we examined the extent to which these variables (baseline SNR-50, PTA4, Age) predicted the magnitude of ctVNS-evoked enhancements using multivariable LME modeling. Given the significant enhancements to SIN performance at moderate-to-high SNRs, ctVNS-evoked changes were quantified as the change in the psychometric function's upper asymptote (λ). Tonic ctVNS-evoked enhancements grew by 3.59% (R^2=0.49, SE=1.01, t(21.66)=3.55, p=0.028) for each standardized unit of change in baseline SNR-50 (**Figure 5A**). However, neither hearing loss (t(21.03)=-0.63, p=0.97; **Figure 5B**), age (t(21.16)=0.89, p=0.88; **Figure 5C**) nor any two-way or three-way interactions significantly predicted ctVNS-evoked enhancements.

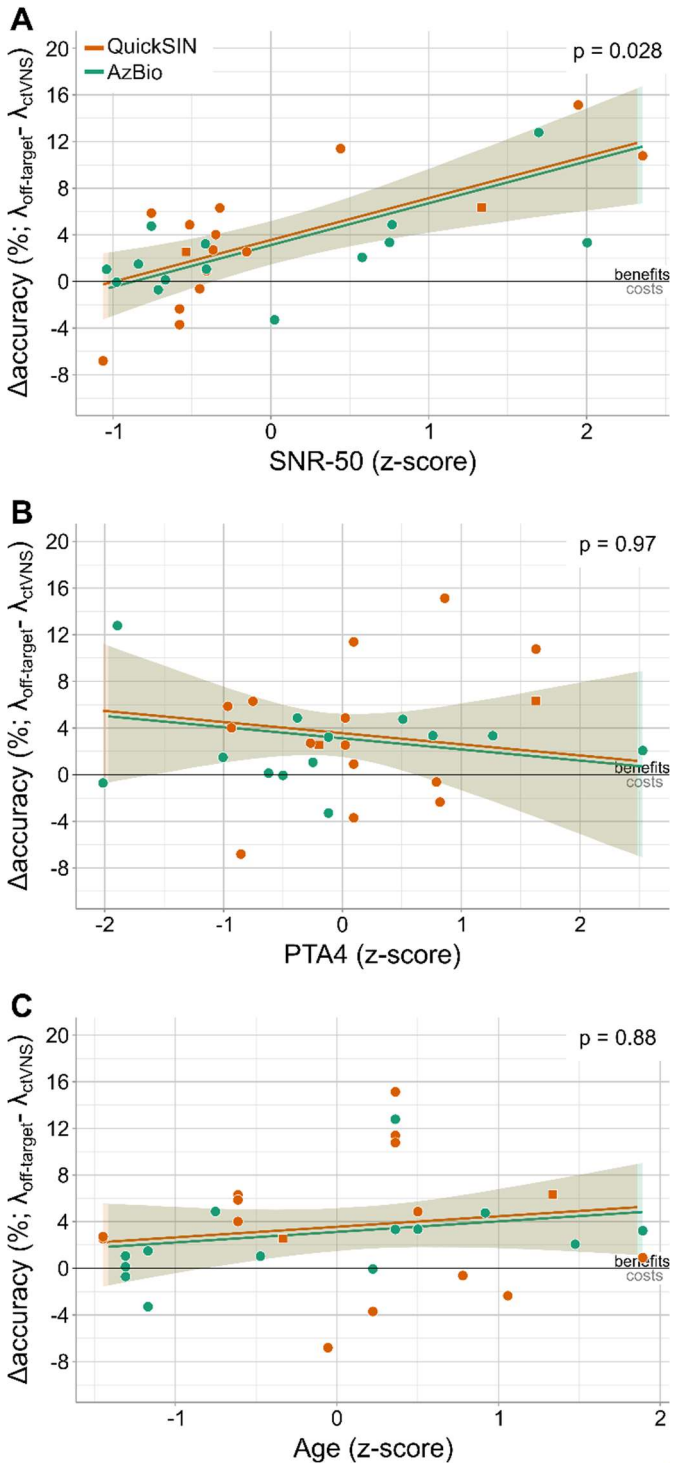

Figure 5. ctVNS-evoked enhancements grew with the severity of speech-in-noise deficits, but neither hearing loss nor age. ctVNS-evoked changes in the upper asymptote of SIN perception (λ), are displayed as function of (**A**) baseline SNR-50 measured during off-target stimulation (**B**) hearing loss (PTA4), and (**D**) age. In all panels, each dot depicts individual participants in a cohort, lines depict the marginal regression for each independent variable determined using multivariable LME modeling, with shaded areas indicating 95% confidence intervals. All p-values were corrected for regression to the mean and for multiple comparisons using the maxT method [43]**.**

Discussion

We tested the hypothesis that tonic ctVNS enhances SIN recognition. Across two cohorts of older adults performing standard audiological SIN assessments, tonic ctVNS improved clinically relevant metrics that guide audiologists' recommendations for hearing aid or cochlear implant use [27, 51, 52]. Compared to placebo stimulation, the average participant experienced a 12% enhancement to their SNR-50 threshold for reliable SIN recognition and a 4% increase to their word recognition accuracy during tonic ctVNS. Moreover, ctVNS-evoked enhancements met or exceeded minimum clinically important differences (MCID) in 39% of participants, on average.

Improvements by tonic ctVNS paralleled those by hearing aids, OTC sound amplification devices, and cochlear implants. A clinical trial found that individuals with mild-to-moderate hearing loss experienced median SNR-50 improvements of 0.47 dB or 1.14 dB after two weeks of using FDA-cleared hearing aids (Lexie Lumen) that were self-fitted or audiologist-fitted, respectively [23]. In contrast, tonic ctVNS produced an average SNR-50 improvement of 0.72 dB—surpassing self-fitted hearing aids and achieving over half the benefit of audiologist-fitted aids. Similarly, an OTC sound amplifier (Tweak Focus) increased SIN recognition accuracy in AzBio by 4.90%, on average, when speech and noise sources were separated by 180°. A different OTC sound amplifier (MSA 30X) impaired accuracy by 11.2% [24]. Despite our more challenging setup with speech and noise co-located at 0° [53], tonic ctVNS achieved a comparable 4% improvement, on average. Lastly, a retrospective review of 863 adult cochlear implant patients identified a 4.90% improvement in word recognition accuracy as their MCID [46]. Similarly, an FDA-approved cochlear implant (Nucleus 24 Cochlear Implant System) used an MCID of 1 dB for SNR-50 to determine its effectiveness [45]. Even though our participants were not selected based on their severity of hearing loss—a prerequisite for hearing aid or cochlear implant candidacy [23, 24, 27]—and most had normal hearing or mild hearing loss, tonic ctVNS delivered clinically meaningful enhancements in SIN perception. Improvements in 35% of participants exceeded the MCID for SNR-50 and 42% for word recognition accuracy. Thus, in a single session, tonic ctVNS achieved SIN enhancements comparable to those by established assistive hearing technologies, even in a cohort predominantly characterized by near-normal hearing.

Tonic ctVNS produced the greatest enhancements at moderate-to-high SNRs, which are common during conversations in real-world environments. Large-scale recordings from various noisy settings (e.g., commercial transportation) showed that SNRs are typically 5-15 dB [47-49], with older adults rating SNRs below 0 dB as "very noisy", between 5-10 dB as "noisy", and above

10 dB as “quiet” [50]. We found that tonic ctVNS enhanced SIN recognition in “noisy” and “quiet” conditions, without affecting performance in “very noisy” settings. Based on a prominent model of SIN recognition [54], this enhancement pattern suggests tonic ctVNS improved SIN performance by enhancing auditory processing in the brain. This model attributes SIN deficits to two components: 1) an impaired sensory periphery with difficulties detecting auditory signals and 2) impaired sensory processing in the brain that distorts the detected auditory information. In this model, impairments in the auditory periphery are hypothesized to appear as deficits at low SNRs, reflected in the psychometric function’s midpoint [31]. Indeed, the midpoint correlated with hearing loss severity (**Figure S1C**), supporting its link to peripheral sensitivity. In contrast, impaired central auditory processing is hypothesized to more prominently influence performance at moderate-to-high SNRs, represented by the function’s upper asymptote [31]. Therefore, the ctVNS-evoked enhancements at the upper asymptote not only demonstrate enhanced SIN recognition under real-world settings but also suggest improved auditory processing in the brain.

We hypothesize that tonic ctVNS enhanced central auditory processing by optimizing thalamocortical circuits for sensory processing via LC-NE activation. Our prior research in younger adults supports this hypothesis, demonstrating that tonic ctVNS can enhance the ability to detect brief gaps in audio during a gap discrimination task requiring precise temporal processing in the brain [14]. As accurate temporal processing is vital for speech recognition [16-18], these earlier findings showed that ctVNS can enhance aspects of central auditory processing critical for SIN recognition. Our present findings provide converging support for this hypothesis. Evidence from invasive VNS [12, 13] and direct LC-NE stimulation in rats [15, 55] demonstrate these sensory enhancements stem from a tonic activation of the LC-NE system. This activation increases NE concentration, optimizing thalamocortical circuits for accurate sensory processing by suppressing disruptive burst spiking activity [12, 15].

A noteworthy finding was that enhancements were most pronounced in individuals with SIN deficits but were similar regardless of hearing loss severity or age. Thus, tonic ctVNS depended neither on peripheral sensitivity nor age, but uniquely alleviated SIN difficulties. Similar compensatory benefits have been observed in younger adults performing psychophysics tasks that rely on sensory processing in the brain [14]. This consistent pattern across diverse age groups and sensory tasks suggests that tonic ctVNS may alleviate sensory deficits that are neural in origin, benefiting those with the greatest need.

Other neuromodulation approaches have shown limited success in improving SIN recognition. Transcranial direct current stimulation (tDCS) over the left superior temporal gyrus

improved SIN performance in the CUNY Sentences Test [56], but used a sham condition without adequate control for the tDCS sensation [57]. In contrast, tonic ctVNS and placebo stimulation were perceived as equivalently intense in our study (**Table 1**). Intermittent theta-burst transcranial magnetic stimulation was initially found to improve accuracy of SIN perception when delivered to the left ventral premotor cortex [58]. However, a follow-up study could not replicate these effects [59]. Instead, benefits occurred only in individuals with poor baseline performance, like the compensatory benefits of tonic ctVNS we have found in younger adults [14] and, now, in older adults. Lastly, pairing brief VNS bursts with auditory tones or sensory training can accelerate and strengthen recovery from sensory disruption [60-65]. However, in a rat model of hearing loss, pairing VNS with training did not improve speech perception beyond training alone [66]. In our previous work with rats, tonic VNS rapidly enhanced central sensory processing, with effects dissipating quickly after stimulation ended [12], suggesting that ongoing tonic stimulation is critical for sustained sensory enhancement. This contrast between the SIN enhancements we demonstrated in humans with and without hearing loss, and the lack thereof in an animal model of hearing loss highlights the critical role stimulation parameters play for effective neuromodulation.

Our stimulation protocol was designed using prior research. First, off-target stimulation on the neck-shoulder junction served as placebo because it has been shown to have no effect on sensory performance, matching the impact of forearm stimulation or a no-stimulation baseline in previous work [14]. Second, the electrodes' spatial geometry used during ctVNS has been shown to enhance auditory performance in humans [14] and is supported by computational modeling indicating that our parameters (**Table 1**) can drive action potentials in afferent and efferent fibers of the vagus nerve's cervical branch [67]. Third, our ctVNS protocol has been shown to modulate heart rate variability [14], a marker of efferent vagal activation [68, 69]. Lastly, given that VNS-evoked thalamocortical optimization dissipates within seconds after stimulation ends [12], our approximately 30-minute washout period ensured a return to baseline between ctVNS and placebo conditions—a shorter duration (20 minutes) still yielded robust ctVNS effects in prior work [14].

A limitation of our study is that two participants in the QuickSIN cohort wore hearing aids during SIN testing, introducing a potential confound. To alleviate this concern, we repeated all statistical analyses without these participants and found identical results, so both were included in the final dataset (marked with square symbols in all plots). Notably, tonic ctVNS improved performance at moderate-to-high SNRs for both individuals (**Figure 5**), suggesting that ctVNS

complemented their hearing aids to enhance SIN perception. Future studies are needed to explore this interaction.

In summary, tonic ctVNS provided immediate, clinically meaningful improvements in speech-in-noise recognition, particularly amid typical conversational noise levels and for individuals with SIN deficits. This study adds converging evidence for an LC-NE-mediated enhancement to auditory processing in the brain, triggered on demand by tonic ctVNS. By establishing tonic ctVNS as a rapid-acting solution for SIN difficulties, this study highlights its potential to complement traditional assistive hearing technologies and offers promising implications for novel treatments of sensory disorders.

Glossary

LC: locus coeruleus

MCID: minimum clinically important difference

NE: norepinephrine

VNS: vagus nerve stimulation

SIN: speech-in-noise

ctVNS: cervical transcutaneous vagus nerve stimulation

λ: the psychometric function's upper asymptote of SIN accuracy

SNR-50: signal-to-noise ratio that attains 50% word recognition accuracy

SNRα: signal-to-noise ratio that bisects the psychometric function, its midpoint

QuickSIN: the quick speech-in-noise test

AzBio: the AzBio sentence test

PTA: pure tone audiometry

PTA4: the average volume of pure tones at 0.5, 1, 2, 4 kHz that defines hearing loss, following World Health Organization guidelines

LME: linear mixed-effects model

OTC: over the counter

RTM: regression to the mean

Washout period: a break between ctVNS and off-target stimulation

Declarations of interest. All authors have financial interest in Sharper Sense, Inc., a company developing methods for enhancing sensory processing with vagus nerve stimulation. Jason B. Carmel is a Founder and stockholder in BackStop Neural and has received honoraria from Pacira, Motric Bio, and Restorative Therapeutics.

Funding. This work was funded by Sharper Sense, Inc. and The Jacobs Technion-Cornell Institute at Cornell Tech. Qi Wang was supported by the National Institutes of Health (R01NS119813, R01AG075114, R21MH125107), National Science Foundation (CBET 1847315, TI 2232149) and the Air Force Office of Scientific Research (FA9550-22-1-0337). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Air Force.

CRediT authorship contribution statement. **Michael Jigo:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Visualization, Project administration, Writing–Original Draft, Review & Editing. **Jason B. Carmel:** Conceptualization, Supervision, Writing–Review & Editing. **Qi Wang:** Conceptualization, Supervision, Writing – Review & Editing. **Charles Rodenkirch:** Conceptualization, Methodology, Supervision, Resources, Funding acquisition, Writing–Review & Editing.

Acknowledgements. We thank Dr. Frank R. Lin, MD, PhD, Dr. Nicholas S. Reed, Ph.D, and Dr. Joseph J. Montano, Ed.D for their helpful input, the Weill Cornell Medicine-Hearing and Speech Center for providing us with an audiometer, and all volunteer research participants.

References

[1] Le Prell CG, Clavier OH. Effects of noise on speech recognition: Challenges for communication by service members. Hearing Research 2017;349:76-89.

[2] Van Hedger SC, Johnsrude IS. Speech perception under adverse listening conditions. Speech perception: Springer; 2022, p. 141-71.

[3] World report on hearing. World Health Organization; 2021.

[4] Tremblay KL, Pinto A, Fischer ME, Klein BE, Klein R, Levy S, et al. Self-reported hearing difficulties among adults with normal audiograms: The Beaver Dam Offspring Study. Ear and Hearing 2015;36(6):e290-e9.

[5] Spankovich C, Gonzalez VB, Su D, Bishop CE. Self reported hearing difficulty, tinnitus, and normal audiometric thresholds, the National Health and Nutrition Examination Survey 1999–2002. Hearing Research 2018;358:30-6.

[6] Zheng Y, Guan J. Cochlear Synaptopathy: A Review of Hidden Hearing Loss. Journal of Otorhinolaryngology Disorders and Treatments 2018;1(1).

[7] Bregman AS. Auditory scene analysis: The perceptual organization of sound. MIT press; 1994.

[8] Middlebrooks JC, Simon JZ, Popper AN, Fay RR. The auditory system at the cocktail party. Springer; 2017.

[9] Kochkin S. MarkeTrak VI: Consumers rate improvements sought in hearing instruments. Hearing Review 2002;9(11):18-22.

[10] Lopez-Poveda EA, Johannesen PT, Pérez-González P, Blanco JL, Kalluri S, Edwards B. Predictors of hearing-aid outcomes. Trends in Hearing 2017;21.

[11] Zaltz Y, Bugannim Y, Zechoval D, Kishon-Rabin L, Perez R. Listening in noise remains a significant challenge for cochlear implant users: Evidence from early deafened and those with progressive hearing loss compared to peers with normal hearing. Journal of Clinical Medicine 2020;9(5).

[12] Rodenkirch C, Wang Q. Rapid and transient enhancement of thalamic information transmission induced by vagus nerve stimulation. Journal of Neural Engineering 2020;17(2).

[13] Rodenkirch C, Carmel JB, Wang Q. Rapid effects of vagus nerve stimulation on sensory processing through activation of neuromodulatory systems. Frontiers in Neuroscience 2022;16.

[14] Jigo M, Carmel JB, Wang Q, Rodenkirch C. Transcutaneous cervical vagus nerve stimulation improves sensory performance in humans: a randomized controlled crossover pilot study. Scientific Reports 2024;14(1).

[15] Rodenkirch C, Liu Y, Schriver BJ, Wang Q. Locus coeruleus activation enhances thalamic feature selectivity via norepinephrine regulation of intrathalamic circuit dynamics. Nature Neuroscience 2019;22(1):120-33.

[16] Tallal P, Miller S, Fitch RH. Neurobiological basis of speech: a case for the preeminence of temporal processing. Annals-New York Academy of Sciences 1993;682.

[17] Schneider BA, Pichora-Fuller MK. Age-related changes in temporal processing: implications for speech perception. *Seminars in hearing.* 22. Thieme Medical Publishers, Inc.; 2001:227-40.

[18] Anderson S, Karawani H. Objective evidence of temporal processing deficits in older adults. Hearing Research 2020;397.

[19] Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America 2004;116(4):2395-405.

[20] Spahr AJ, Dorman MF, Litvak LM, Van Wie S, Gifford RH, Loizou PC, et al. Development and validation of the AzBio sentence lists. Ear and Hearing 2012;33(1):112-7.

[21] Wong A, Nyenhuis D, Black SE, Law LS, Lo ES, Kwan PW, et al. Montreal Cognitive Assessment 5-minute protocol is a brief, valid, reliable, and feasible cognitive screen for telephone administration. Stroke 2015;46(4):1059-64.
[22] Green P, MacLeod CJ. SIMR: An R package for power analysis of generalized linear mixed models by simulation. Methods in Ecology and Evolution 2016;7(4):493-8.
[23] De Sousa KC, Manchaiah V, Moore DR, Graham MA, Swanepoel DW. Effectiveness of an over-the-counter self-fitting hearing aid compared with an audiologist-fitted hearing aid: A randomized clinical trial. JAMA Otolaryngology–Head & Neck Surgery 2023;149(6):522-30.
[24] Reed NS, Betz J, Kendig N, Korczak M, Lin FR. Personal sound amplification products vs a conventional hearing aid for speech understanding in noise. JAMA 2017;318(1):89-90.
[25] QuickSIN™ Speech-in-Noise Test (Version 1.3) [Audio CD and Test Manual]. Elk Grove Village, IL: Etymotic Research; 2006.
[26] AzBio Sentence Test v2.0. Goodyear, Az: Auditory Potential LLC; 2022.
[27] Dunn CC, Zwolan TA, Balkany TJ, Strader HL, Biever A, Gifford RH, et al. A Consensus to Revise the Minimum Speech Test Battery–Version 3. American Journal of Audiology 2024;33(3):624-47.
[28] Miller J, Ulrich R. On the analysis of psychometric functions: The Spearman-Kärber method. Perception & Psychophysics 2001;63:1399-420.
[29] Wilson RH, McArdle RA, Smith SL. An Evaluation of the BKB-SIN, HINT, QuickSIN, and WIN Materials on Listeners With Normal Hearing and Listeners With Hearing Loss. Journal of Speech, Language, and Hearing Research 2007;50(4):844-56.
[30] Linares D, López-Moliner J. quickpsy: An R package to fit psychometric functions for multiple groups. The R Journal, 2016, vol 8, num 1, p 122-131 2016.
[31] Ross B, Dobri S, Schumann A. Psychometric function for speech-in-noise tests accounts for word-recognition deficits in older listeners. The Journal of the Acoustical Society of America 2021;149(4):2337-52.
[32] Wichmann FA, Hill NJ. The psychometric function: I. Fitting, sampling, and goodness of fit. Perception & Psychophysics 2001;63(8):1293-313.
[33] Kleiner M, Brainard D, Pelli D. What's new in Psychtoolbox-3? 2007.
[34] Carhart R, Jerger JF. Preferred method for clinical determination of pure-tone thresholds. Journal of speech and hearing disorders 1959;24(4):330-45.
[35] Watson AB, Pelli DG. QUEST: A Bayesian adaptive psychometric method. Perception & Psychophysics 1983;33(2):113-20.
[36] ISO 389-7: Acoustics—Reference Zero for the Calibration of Audiometric Equipment—Part. 7: Reference Threshold of Hearing under Free-Field and Diffuse-Field Listening Conditions. International Organization for Standardization: Geneva, Switzerland; 2019.
[37] R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2024.
[38] Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. Journal of Statistical Software 2017;82(13).
[39] McArdle RA, Wilson RH. Homogeneity of the 18 QuickSIN™ lists. Journal of the American Academy of Audiology 2006;17(03):157-67.
[40] Barton K, Barton MK. Package 'mumin'. Version 2015;1(18):439.
[41] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological) 1995;57(1):289-300.
[42] Furrow RE. Regression to the Mean in Pre–Post Testing: Using Simulations and Permutations to Develop Null Expectations. CBE—Life Sciences Education 2019;18(2):le2.
[43] Westfall PH, Young SS. Resampling-based multiple testing: Examples and methods for p-value adjustment. John Wiley & Sons; 1993.

[44] Morey RD. Confidence intervals from normalized data: A correction to Cousineau (2005). Tutorials in Quantitative Methods for Psychology 2008;4(2):61-4.
[45] Summary of Safety and Effectiveness Data for Nucleus 24 Cochlear Implant System. U.S. Food and Drug Administration; 2022.
[46] Patro A, Moberly AC, Freeman MH, Perkins EL, Jan TA, Tawfik KO, et al. Investigating the Minimal Clinically Important Difference for AzBio and CNC Speech Recognition Scores. Otology & Neurotology 2024;45(9):e639-e43.
[47] Pearsons KS, Bennett RL, Fidell SA. Speech levels in various noise environments. Office of Health and Ecological Effects; 1977.
[48] Olsen WO. Average Speech Levels and Spectra in Various Speaking/Listening Conditions. American Journal of Audiology 1998;7(2):21-5.
[49] Smeds K, Wolters F, Rung M. Estimation of signal-to-noise ratios in realistic sound scenarios. Journal of the American Academy of Audiology 2015;26(02):183-96.
[50] Wu Y-H, Stangl E, Chipara O, Hasan SS, Welhaven A, Oleson J. Characteristics of real-world signal to noise ratios and speech listening situations of older adults with mild to moderate hearing loss. Ear and Hearing 2018;39(2):293-304.
[51] Mueller HG. Signia expert series: Speech-in-noise testing for selection and fitting of hearing aids: Worth the effort. AudiologyOnline 2016:18336.
[52] Carlson ML, O'Connell BP, Lohse CM, Driscoll CL, Sweeney AD. Survey of the American Neurotology Society on Cochlear Implantation: part 2, surgical and device-related practice patterns. Otology & Neurotology 2018;39(1):e20-e7.
[53] Wang X, Xu L. Speech perception in noise: Masking and unmasking. Journal of Otology 2021;16(2):109-19.
[54] Plomp R. A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. Journal of Speech, Language, and Hearing Research 1986;29(2):146-54.
[55] Rodenkirch C, Wang Q. Optimization of Temporal Coding of Tactile Information in Rat Thalamus by Locus Coeruleus Activation. Biology 2024;13(2):79.
[56] Lifshitz-Ben-Basat A, Taitelbaum-Swead R, Fostick L. Speech perception following transcranial direct current stimulation (tDCS) over left superior temporal gyrus (STG)(including Wernicke's area) versus inferior frontal gyrus (IFG)(including Broca's area). Neuropsychologia 2024;202:108959.
[57] Kessler SK, Turkeltaub PE, Benson JG, Hamilton RH. Differences in the experience of active and sham transcranial direct current stimulation. Brain Stimulation 2012;5(2):155-62.
[58] Brisson V, Tremblay P. Improving speech perception in noise in young and older adults using transcranial magnetic stimulation. Brain and Language 2021;222:105009.
[59] Brisson V, Tremblay P. Assessing the Impact of Transcranial Magnetic Stimulation on Speech Perception in Noise. Journal of Cognitive Neuroscience 2024;36(10):2184-207.
[60] De Ridder D, Vanneste S, Engineer ND, Kilgard MP. Safety and efficacy of vagus nerve stimulation paired with tones for the treatment of tinnitus: a case series. Neuromodulation: Technology at the Neural Interface 2014;17(2):170-9.
[61] Tyler R, Cacace A, Stocking C, Tarver B, Engineer N, Martin J, et al. Vagus nerve stimulation paired with tones for the treatment of tinnitus: a prospective randomized double-blind controlled pilot study in humans. Scientific Reports 2017;7(1):11960.
[62] Meyers EC, Kasliwal N, Solorzano BR, Lai E, Bendale G, Berry A, et al. Enhancing plasticity in central networks improves motor and sensory recovery after nerve damage. Nature Communications 2019;10(1):5782.
[63] Adcock KS, Chandler C, Buell EP, Solorzano BR, Loerwald KW, Borland MS, et al. Vagus nerve stimulation paired with tones restores auditory processing in a rat model of Rett syndrome. Brain Stimulation 2020;13(6):1494-503.

[64] Darrow MJ, Mian TM, Torres M, Haider Z, Danaphongse T, Rennaker Jr RL, et al. Restoration of somatosensory function by pairing vagus nerve stimulation with tactile rehabilitation. Annals of Neurology 2020;87(2):194-205.
[65] Kilgard MP, Rennaker RL, Alexander J, Dawson J. Vagus nerve stimulation paired with tactile training improved sensory function in a chronic stroke patient. NeuroRehabilitation 2018;42(2):159-65.
[66] Carroll AM, Riley JR, Borland MS, Danaphongse TT, Hays SA, Kilgard MP, et al. Bursts of vagus nerve stimulation paired with auditory rehabilitation fail to improve speech sound perception in rats with hearing loss. iScience 2024;27(4).
[67] Mourdoukoutas AP, Truong DQ, Adair DK, Simon BJ, Bikson M. High-resolution multi-scale computational model for non-invasive cervical vagus nerve stimulation. Neuromodulation: Technology at the Neural Interface 2018;21(3):261-8.
[68] Burger AM, D'Agostini M, Verkuil B, Van Diest I. Moving beyond belief: a narrative review of potential biomarkers for transcutaneous vagus nerve stimulation. Psychophysiology 2020;57(6):e13571.
[69] Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Frontiers in Public Health 2017;5:258.