Extended High-Frequency Hearing Loss, Not Cochlear Synaptopathy, Predicts Speech Recognition in a Population Cohort

Christopher R. Cederroth,[1, 2, *]

[1]Laboratory of Translational Auditory Neuroscience, Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden
[2]Tübingen Hearing Research Centre, Department of Otolaryngology, Head and Neck Surgery, University of Tübingen, Tübingen, Germany

*Christopher R. Cederroth, Karolinska Institutet, Dpt. of Physiology and Pharmacology, Solnavägen 9, 171 77 Stockholm, Sweden (Tel.: 0046.8.524.87235; Email: christopher.cederroth@ki.se)

Abstract

Age-related hearing loss is the leading sensory deficit in older adults, yet audiometric thresholds at conventional frequencies often poorly predict speech understanding. Two competing hypotheses have emerged: extended high-frequency (eHF) hearing loss beyond 8 kHz may unmask variance in speech performance, while hidden hearing loss from cochlear synaptopathy - detectable via the reduction of auditory brainstem response (ABR) wave I amplitude - may degrade higher order temporal processing independent of audiometry. Here, in 526 ears from 263 tinnitus-free adults in the Swedish Tinnitus Outreach Project (STOP) cohort, we show that eHF pure-tone average (10–16 kHz) is the single most age-sensitive auditory measure, explaining 64% of age-related variance compared to only 16% for conventional audiometry (PTA4, 0.5-4 kHz). EHF strongly predicted speech-in-noise performance, substantially more than any other audiological marker (i.e., PTA4, otoacoustic emissions, or ABR). Whereas Wave I amplitude showed no effect, Wave I latency showed robust predictive value in the decline of word recognition. While in young adults, delayed Wave I latency had little impact on word recognition, in older adults, a 1.8 ms delay lead to ~25% drop in word recognition. These findings challenge the translational relevance of cochlear synaptopathy to age-related speech deficits and identify auditory nerve conduction delays, not synaptic loss, as the peripheral neural mechanism underlying speech comprehension decline in aging. EHF audiometry emerges as the dominant clinical predictor of speech outcomes, with implications for clinical routine and interventional clinical trials in heterogeneous human populations.

Introduction

Hearing loss is defined as a loss of >20 dB hearing levels (HL) at the pure tone average of audiometric thresholds of 0.5, 1, 2, 4 kHz (PTA4). Based on this gold-standard, hearing loss affects over 1.5 billion people globally, with prevalence rising steeply after age 50. Between 1990 and 2019, global moderate-to-complete hearing loss increased 79.1%, from 225.3 million to 403.3 million, with projections suggesting 2.45 billion affected by 2050[1]. Beyond auditory function, hearing loss has emerged as a potentially modifiable risk factor for cognitive decline and dementia[2]. Cross-sectional studies show strong associations of mild cognitive impairement evidencing impaired functional and anatomical connectivity once corrected for β-Amyloid Deposition[3]. Population-based cohort studies reveal that hearing loss associates with incident cognitive impairment and dementia[4]. However, in systematic reviews and meta-analyses that have evidenced a strong relationship between cognitive impairement or dementia, 80% of the studies limited their auditory assessment to 4 kHz[5]. Despite its clinical importance, the PTA4 poorly predicts real-world communication ability and speech understanding in challenging environments. In addition, clinical pure-tone audiometry (up to 8 kHz) explains limited variance in word recognition and speech-in-noise (SPiN) performance, leaving clinicians unable to accurately predict who will struggle most with communication. As a consequence, hearing loss is often detected only once communication is impacted, leading to a delay of near 10 years until interventions (e.g., hearing aids) occur[6], although these could potentially prevent cognitive decline[7].

Difficulties in understanding speech may occur in presence of a normal audiogram. Early detection of SPiN deficits may thus contribute to earlier interventions for hearing and cognitive interventions[8]. Interrestingly, recent work shows that eHF pure-tone averages (10–16 kHz) robustly predict word and phoneme recognition in challenging acoustic environments[9]. Since cochlear damage from aging, noise exposure or ototoxicity preferentially affects the cochlear base, where frequencies above 8 kHz are encoded[10], measurement of extended high-frequency (eHF) thresholds (> 8 kHz) may reveal clinically relevant cochlear pathology that conventional PTA misses. Recent studies have suggested a contribution of cochlear synaptopathy (CS) to "hidden hearing loss" and SPiN deficits[11–13]. CS derives from pre-clinical work showing that moderate noise exposure can selectively destroy high-threshold, low-spontaneous-rate auditory nerve fibers at the inner hair cell synapse without elevating auditory brainstem thresholds, albeit decreasing the auditory brainstem response (ABR) Wave 1 amplitude of the auditory brainstem response[14–16]. This synaptic pathology theoretically persists even after threshold recovery, potentially explaining deficits in speech understanding in noisy environments despite normal or near-normal quiet thresholds. Post-mortem histopathology of human temporal bones found cochlear nerve degeneration to be associated with poorer word recognition performance[17]. Likewise, in young healthy individuals with normal hearing thresholds up to 16 kHz, those with lower ABR Wave I amplitudes had poorer performance on word recognition in noise or with time compression and reverberation[18]. Bridging the two, a reduction in Wave I amplitude has been proposed to serve as a noninvasive proxy of CS in humans.

To clarify which of the two mechanistic pathways, eHF or Wave I amplitude, dominates as a predictor of speech outcomes in humans, I analyzed data from the Swedish Tinnitus Outreach Project, a large general population cohort with comprehensive audiometric and neurophysiological assessment. The analysis focused on a subset of healthy adults without tinnitus, eHF pure-tone audiometry (10–16 kHz), high-reliability ABR recordings with rigorous quality control, and speech recognition measures (word and phoneme recognition in noise). I used multivariate statistical modeling to assess the independent explanatory power of eHF thresholds vs ABR markers after accounting for conventional audiometry and demographic factors.

Results

Cohort characteristics

The STOP cohort comprised 526 ears from 263 participants aged 20 to 84 years (mean ± SD: 49.2 ± 13.6 years), with 68% woman representation (**Table 1**). Participants were measured in a clinical environment with a sound proof booth, with EC marked clinical devices. Overall, the cohort exhibited clinically normal hearing thresholds (cut-off 20 dB HL) in the conventional frequency range (PTA_4 0.5–4 kHz: 5.2 ± 10.7 dB HL), which is the clinical gold-standard used by the WHO for classifying hearing impairment severity. In contrast, substantial variability was found in eHF hearing loss (PTAHF 10–16 kHz: 36.5 ± 22.5 dB HL, range: –12.2 to 78.0 dB HL). Loudness Discomfort Levels (LDLs) plateaued near ceiling values of the medical device (5.2 ± 10.7 dB HL). Distortion Products of Otoacoustic Emissions (DPOAEs), a measure of outer hair cell (OHC) health and function, at 5.273 kHz averaged 19.8 ± 10.1 dB signal to noise ratio (SNR). Speech perception performance (ranging from 0 to 100% accuracy) was tested on 511 ears with hearing abilities in the standard speech weighted noise (+4 dB SNR) was robust, with mean word scores (WS) of 75.1 ± 10.3% and phoneme scores (PS) of 89.2 ± 6.0%. Auditory brainstem response measurements (available for 336 ears; **Fig. 1**) were performed on two different medical devices: the Charter EP (Otometrics) and the Eclipse (Interacoustics). In our former study, test-retest reliability was moderate (ICC>0.5) for all Wave latencies in both devices, but poor (ICC<0.5) for all Wave amplitudes, with the exception of the Eclipse Wave I amplitudes being excellent (>0.9). As a consequence, only Wave 1 amplitudes for the Eclipse are reported (0.4 ± 0.2 μV, n = 140 ears), and the Wave latencies (from both devices) were 1.4 ± 0.2 ms for Wave I, 3.6 ± 0.2 ms for Wave III, and 5.4 ± 0.4 ms for Wave V. The I-V interwave latency was 4.0 ± 0.3 ms (**Table 1**).

Auditory performance declines with age

Age-stratified analyses revealed progressive deterioration across auditory domains (**Fig. 2a, Supplemental Table 2**): eHF thresholds worsened from 13.8 ± 15.2 dB HL in participants <35 years to 69.0 ± 6.2 dB HL in those ≥75 years. LDLs had near ceiling effects and tolerance to loud sounds at 8 kHz increased with age (**Fig. 2b**). Both PTA and LDL's showed device limitations in the eHF with maximal dbHL values achieved for 10 kHz (93 dB HL); 12.5 kHz (85 dB HL); 4 kHz (75 dB HL) and 16 kHz (55.5 dB HL). DPOAEs (hereafter refered to as OAEs), showed also an age-progressive decline, most evident for the 5.273 kHz, with a 24.5 ± 8.4 dB SNR in participants <35 years to 7.9 ± 10.1 dB SNR in those ≥75 years (**Fig. 2c**). WS and PS declined from 78.3 ± 6.5% to 55.6 ± 19.4%, and 90.9 ± 3.7% to 77.8 ± 12.3% respectively, across the same age range (**Fig. 2d,e**). Wave I amplitude (Eclipse only) demonstrated an age-related decline from 0.5 ± 0.2 μV in younger participants (<35 years) to 0.3 ± 0.1 μV in the 65–74 age group (**Fig. 2f**). Wave latency increased with age in Wave I (from 1.3 ± 0.1 ms in <35 years to 1.5 ± 0.2 ms in 65–74 years) and Wave V (from 5.3 ± 0.3 ms to 5.6 ± 0.3 ms; Fig. 2g,h,i), with corresponding prolongation of central conduction time (I–V inter-wave latency: 3.9 ± 0.3 ms to 4.1 ± 0.3 ms). The ABR data collected on two platforms (ChartR n=190, Eclipse n=146) showed systematic offset (0.21 ms latency, p<0.001) but no device×predictor interactions (all p>0.10), indicating findings replicate across independent measurement systems. The mean and standard error of the mean (SEM) latency values obtained from the two devices using equal settings are reported for each age group (**Fig. 2g,h,i**). **Supplemental Table 3,4** report these values stratified by gender and eHF tertiles.

Nonlinear Relationships Between Auditory Function and Aging

Guided by the non-linear, age-stratified patterns in **Fig. 2**, I next modelled each audiometric and physiological variable as a function of age using univariate polynomial regression (linear vs. quadratic vs. cubic fits; **Fig. 3, Supplemental Table 5**). PTA_4 showed a modest but significant quadratic increase with age (R^2 = 0.158, quadratic p = 0.0057), whereas PTAHF exhibited a much steeper age-related elevation with high overall fit (R^2 = 0.641) but no additional improvement from adding a quadratic term (quadratic p = 0.547), indicating an approximately linear worsening of eHF thresholds across the adult lifespan (**Fig. 3a,b**). OAE amplitude declined with age in a curvilinear fashion (R^2 = 0.237, quadratic p = 0.009; **Fig. 3c**), while both WS and PS showed pronounced

accelerating decline captured by strong quadratic components (WS: $R^2 = 0.296$, quadratic $p < 0.0001$; PS: $R^2 = 0.283$, quadratic $p < 0.0001$), suggesting an inverted-U relationship whereby speech perception performance initially improves with age before plateauing or declining (**Fig. 3d,e**). The Age² coefficients were negative for both outcomes (WS: $\beta = -0.013$, $p < 0.001$; PS: $\beta = -0.0077$, $p < 0.001$), indicating that gains diminish progressively across the lifespan (**Supplemental Table 6**). Gender exerted only marginal effects on speech perception; when age was included in the model, women performed modestly better than men (WS: +1.82 points, $p = 0.034$; PS: +1.00 point, $p = 0.05$), though the gender effect on phoneme scores was marginally significant. Critically, the Age × Gender interaction was not significant for either outcome (WS: $p = 0.271$; PS: $p = 0.067$), indicating that age-related trajectories of speech perception do not differ between genders (**Supplemental Table 6**). Among ABR measures, Wave I amplitude (Eclipse only) decreased with age but was adequately described by a linear fit ($R^2 = 0.195$, quadratic $p = 0.893$, **Fig. 3f**), whereas Wave I, III and V latencies all displayed statistically significant quadratic age effects (Wave I: $R^2 = 0.153$, $p = 0.007$; Wave III: $R^2 = 0.102$, $p = 0.048$; Wave V: $R^2 = 0.127$, $p = 0.028$), consistent with subtly increased delay at older ages (**Fig. 3g–l, Supplemental Table 5**). These polynomial fits confirm that several key predictors (PTA_4, OAE, WS/PS, and brainstem latencies) exhibit non-linear age relationships, providing the rationale for the subsequent univariate polynomial analyses of WS and PS and for including higher-order terms only where formally justified in the multivariable models.

Influence of audiometric measures on speech perception.

To optimize model specification, I systematically evaluated whether audiometric and physiological predictors exhibited nonlinear relationships with speech perception outcomes using likelihood ratio testing (LRT) with Bonferroni correction, fitting univariate polynomial models (linear, quadratic, cubic) for each predictor against word score (WS) and phoneme score (PS). EHF pure-tone average (PTAHF, 10–16 kHz) was the single strongest peripheral predictor of speech performance (**Fig. 4, Supplemental Table 7**): cubic models explained substantially more variance in WS than conventional PTA_4 ($R^2 = 0.345$ vs. 0.221; cubic $p = 0.0014$ for PTAHF, $p = 0.047$ for PTA_4; **Fig. 4a,b**), with an even larger advantage for PS (PTAHF $R^2 = 0.361$ vs. PTA_4 $R^2 = 0.247$; PTAHF cubic $p = 0.0001$, PTA_4 cubic $p < 0.0001$; **Fig. 4g,h**). Both audiometric measures showed robust cubic relationships with speech outcomes, consistent with non-linear effects of hearing loss on speech perception rather than a simple monotonic decline. Critically, PTAHF outperformed PTA_4 by 12.4 percentage points for WS (R^2 difference = 0.124) and 11.1 percentage points for PS ($\Delta R^2 = 0.111$), indicating that eHF thresholds capture speech-relevant cochlear damage that is only partly reflected in conventional audiometry. Together with the steeper age-related deterioration of PTAHF compared to PTA_4 (**Fig. 4a,b**), this pattern supports the clinical added value of incorporating eHF measures into routine assessment. In contrast, OAE amplitude at 5.273 kHz exhibited a linear relationship for WS and cubic for PS (WS: $R^2 = 0.202$, cubic $p = 0.059$; PS: $R^2 = 0.219$, cubic $p < 0.0001$; **Supplemental Figure 1, Supplemental Table 7**), indicating a roughly proportional decrement in speech scores as OHC function decreases. When OAE amplitude was added to models already containing PTAHF, it provided at most a small, borderline-significant improvement in fit for WS and PS ($\Delta R^2 < 0.01$ for both outcomes; quadratic $p = 0.0596$ for WS, $p = 0.0376$ for PS), indicating that most speech-relevant variance in OHC damage is already captured by eHF pure-tone thresholds. These findings suggest that, in this cohort and listening condition (+4 dB SNR), eHF is sufficient to index the cochlear pathology that matters for speech, and OHC-specific measures offer limited incremental value for prediction beyond PTAHF. The corresponding relationships between WS and ABR measures are illustrated in **Fig. 4c-f**: Wave I amplitude (Eclipse only) exhibited a shallow linear association with WS and PS (WS: $R^2 = 0.065$, linear $p < 0.001$; quadratic $p = 0.249$; PS: $R^2 = 0.058$, linear $p < 0.001$; quadratic $p = 0.158$; **Fig. 4c**), with no improvement from higher-order terms, indicating that amplitude contributes minimally to explaining speech variance. Wave I latency showed a linear association for WS (WS: $R^2 = 0.115$, linear $p < 0.0001$; **Fig. 4c**) and quadratic improvement for PS (PS: $R^2 = 0.115$, quadratic $p = 0.024$; **Fig. 4i**), showing modest but significant association with both outcomes. In contrast, Wave III latency required cubic terms for optimal fit (WS: $R^2 = 0.088$, cubic $p = 0.0017$; PS: $R^2 = 0.112$, cubic $p <$

0.0001; **Fig. 4d,j**), indicating strong nonlinearity, while Wave V latency showed only weak linear relationships (WS: $R^2 = 0.065$, quadratic p = 0.868; PS: $R^2 = 0.058$, quadratic p = 0.986; **Fig. 4e,k**, **Supplemental Table 7**). Together, these polynomial fits indicate that eHF audiometry (PTAHF) is the dominant predictor of speech perception, explaining 34–36% of variance in univariate models—substantially more than conventional audiometry 19-25%, OAE 19-22%, or any ABR measure 5-11%. The strong cubic relationships for PTAHF and Wave III latency justify the use of higher-order terms for these predictors in subsequent multivariable models, while OAE, Wave I amplitude, Wave I latency, and Wave V latency are treated as linear predictors.

Comparison of the mechanisms underlying WS and PS.

The comparable outcomes for WS and PS prompted me to investigate further the differences between the two. Indeed, the apparent U-shaped age effect in univariate analysis in **Fig. 4** may mask underlying mechanisms. Once controlling for hearing loss and neural timing measures, WS shows genuine age-acceleration (β = -0.0095, 95% CI -0.018 to -0.001, p=0.028; **Fig. 5a, Supplemental Table 9**), indicating that age-related deterioration accelerates in older age. In contrast, PS showed only linear age effects without significant quadratic acceleration (age^2 β = -0.0034, 95% CI -0.009 to 0.002, p=0.206; **Supplemental Table 9**). Specifically, for each additional year of age, WS declined by approximately 0.3–0.4% at age 50 (linear slope) and 0.6–0.7% at age 75 (age-dependent slope), whereas PS declined uniformly at ~0.15% per year regardless of age. This differential trajectory indicates that complex linguistic processing (word recognition) is more age-vulnerable than basic acoustic perception (phoneme identification).

Multivariable polynomial regression models identified distinct predictor patterns for word versus phoneme recognition. Standardized coefficients revealed that, contrary to predictions from the cochlear synaptopathy hypothesis[14–16] whereby lower Wave I amplitude is associated with speech recognition deficits[18,19], Wave 1 amplitude—the putative neural marker of cochlear inner hair cell-auditory nerve synapse loss—showed no significant relationship to either WS (β = 1.73, 95% CI: -4.38 to 7.85, p=0.579) or PS (β = 1.05, 95% CI: -2.645 to 4.748, p=0.578; **Fig. 5b; Supplemental Table 10**), despite restricting the analysis to the Eclipse subset (N=134–140 ears), where Wave I amplitude reliability is excellent (ICC=0.96)[20], to maximize statistical power for detecting synaptopathy effects. Furthermore, the Age × W1.Amplitude interaction was not significant for either outcome (WS: β = 0.169, 95% CI: −0.043 to 0.773, p=0.584; PS: β = 0.186, 95% CI: −0.207 to 0.578, p=0.354; **Fig. 5b**). Collectively, these results demonstrate that ABR Wave I amplitude shows no association with word or phoneme recognition in presence of background noise, even in the highest-quality, most reliable recording subset. In striking contrast, W1.Latency—reflecting auditory nerve conduction velocity rather than neural synchrony—demonstrated robust predictive value for WS but not PS (WS: β = −9.691, 95% CI: −19.121 to −0.261, p=0.045; PS: β = −4.86, 95% CI: − 10.877 to 1.16, p=0.115; **Fig. 5b**). Each additional millisecond of Wave 1 latency was associated with approximately 10-point decline in Word Score and 5-point decline in Phoneme Score, with WS showing approximately twice the sensitivity to neural timing delays. Critically, the Age × W1.Latency interaction was substantially larger in magnitude for WS compared to PS (**Fig. 5b**). Specifically, WS showed a significant Age × W1.Latency interaction (β = −0.762, 95% CI: −1.485 to −0.039, p=0.0397), indicating that older adults experience amplified speech perception decline as W1 latency increases (**Table 2, Supplemental Table 10**). The PS interaction was not significant being substantially weaker in magnitude (β = −0.367, 95% CI: −0.824 to 0.089, p=0.116), with the WS age-interaction effect approximately 2.11 times the magnitude of the PS effect (coefficient ratio; **Table 2, Supplemental Table 10**). This dissociation demonstrates differential age-dependent vulnerability: while both word and phoneme recognition are sensitive to peripheral neural transmission delays, the age-dependent amplification of neural timing vulnerability is approximately twice as pronounced for word recognition, suggesting task-specific mechanisms of aging. These findings collectively suggest that delayed neural transmission (reflected in increased W1 latency), rather than synapse loss (reflected in decreased amplitude), is the primary mechanistic driver of age-related speech perception decline. The selective age-amplification of the W1.Latency effect for word recognition (vs. phoneme recognition) indicates that complex linguistic

processing may depend more heavily on the precision of early neural timing information than basic acoustic-phonetic processing.

Surprinsingly, Wave III latency cubed (W3.Latency³)—representing higher-order (cubic) nonlinearity in central auditory brainstem processing—significantly predicted both outcomes with large, positive effects (**Fig. 5b, Supplemental Table 10**). Specifically, WS was predicted by W3.Latency³ (β = 81.49, 95% CI: 12.36 to 150.62, p=0.022), and PS was similarly predicted (β = 67.39, 95% CI: 9.24 to 125.54, p=0.024; **Supplemental Table 10**). The comparable magnitude of W3.Latency³ effects between outcomes (difference in coefficients = 14.36 points, approximately 18% relative difference) contrasts sharply with the substantially differential Age × W1.Latency effects (coefficient ratio = 1.97-fold difference). This dissociation is mechanistically informative: higher-order (cubic) central auditory processing effects impact word and phoneme recognition similarly, indicating shared reliance on central auditory complexity. In contrast, the age-dependent neural timing vulnerability—specifically the age-amplified impact of peripheral wave initiation delays (Age × W1.Latency)—shows outcome-specific vulnerability that is unique to word-level processing. This pattern suggests that aging selectively compromises mechanisms dependent on early neural precision (W1 timing), whereas cubic nonlinearities in central processing (W3 cubic term) are inherently task-invariant. Wave 5 latency showed no significant associations in the multivariable model (WS: β = −0.668, 95% CI: −3.596 to 2.2593, p=0.655; PS: β = −0.3492, 95% CI: −1.857 to 1.159, p=0.65; **Supplemental Table 10**), indicating that the most distal brainstem response does not provide independent predictive value when upstream measures (W1 latency, W3 latency) are included in the model.

Age-Stratified Trajectories Reveal Selective Word Score Vulnerability

Predicted trajectories across four age strata (45, 55, 65, 75 years) reveal that while young adults (45y) maintain stable performance across W1 latency values (WS predicted 74.8-81.7%, PS predicted 89.5-92.9%), older adults (75y) show marked declines as W1 latency increases, with WS dropping to 57% and PS to 81% at the maximum observed W1 latency (**Fig. 5c, Supplemental Table 11**). This 25-percentage-point differential drop in WS versus 13-percentage-point drop in PS demonstrates that peripheral neural timing delay effects increase with advancing age, with word-level processing showing approximately 1.9 times greater vulnerability compared to phoneme recognition. The age-stratified trajectories visually demonstrate the Age × W1.Latency interaction: while young adults' performance remains relatively insensitive to W1 latency across the observed range, older adults show substantially steeper slopes, indicating that age unmasks or amplifies the neural timing dependency of speech recognition.

Dissociation Between Word and Phoneme Recognition Across Age Groups

I tested whether word and phoneme scores show differential relationships with age and neural timing by constructing multivariable polynomial regression models with interaction terms. Despite high overall correlation (Pearson r=0.934, p<0.001; **Fig. 5d, Supplemental Table 12**), WS and PS showed dissociable aging patterns and neural correlates. WS exhibited accelerating age-related decline (age² β=−0.010, p=0.012) and strong age-amplified vulnerability to neural timing delays (Age × W1.Latency β=−0.788, p=0.002), whereas PS showed linear aging (age² β=−0.004, p=0.033) without age-amplified neural timing vulnerability (Age × W1.Latency β=−0.401, p=0.006 with substantially reduced effect size; **Supplemental Table 12**). Both outcomes were equally driven by central processing measures (W3.Latency³: WS p=0.0015, PS p<0.001; **Supplemental Table 12**). This dual-process profile suggests that word recognition is preferentially vulnerable to age-related neurobiological changes and early neural timing deficits, while phoneme recognition remains more robust to age-dependent mechanisms, despite both being sensitive to central auditory complexity. Age-stratified analysis revealed selective word score vulnerability despite the high bivariate correlation. Young adults (≤median age 38 years, n=153 ears) clustered tightly in the upper-right quadrant (high WS, high PS) with a regression slope (0.48, 95% CI: 0.44–0.52; **Fig. 5d; Supplemental Table 12**). Older adults (>53 years, n=145 ears) showed greater scatter and were concentrated in the lower-left region (lower WS, lower PS) with a significantly steeper slope

(0.58, 95% CI: 0.54–0.61; **Fig. 5d; Supplemental Table 12**). Older adults show steeper (not flatter) slopes (0.58 vs 0.48), indicating stronger WS-PS coupling in older age groups. This pattern indicates differential age-related WS-PS coupling, where older adults show steeper WS-PS slopes, suggesting that in older individuals, phoneme recognition increases more steeply relative to word score—a pattern consistent with differential age-related vulnerability of complex linguistic versus acoustic-phonetic processing.

Discussion

This cross-sectional analysis of 526 ears from 263 tinnitus-free adults reveals three central findings regarding age-related speech perception decline. First, eHF pure-tone average (10–16 kHz) is the dominant predictor of speech recognition, explaining 34–36% of variance in both word and phoneme scores—substantially more than conventional audiometry or any physiological measure. Second, ABR Wave I amplitude—the putative marker of cochlear synaptopathy—contributes no additional explanatory power, directly contradicting the "hidden hearing loss" hypothesis even when tested in speech-weighted noise (+4 dB SNR) where synaptopathy effects should be most evident. Third, Wave I latency reveals age-dependent vulnerability specific to complex linguistic tasks (word recognition), suggesting that neural conduction slowing, rather than synapse loss, may contribute to age-related speech deficits, though this finding requires validation. Collectively, these results establish eHF audiometry as the primary clinical tool for predicting speech outcomes in aging populations and challenge the translational relevance of ABR-based synaptopathy indices.

The linear worsening of eHF hearing with aging is outstanding for a single auditory measure explaining 64% of the variance. Unlike other measures that show accelerating (quadratic) decline, PTAHF declines steadily and predictably throughout the lifespan - making it a reliable biomarker of cochlear aging far better than any other conventional audiometry, all below 30%. Since PTAHF tracks cumulative cochlear damage with such fidelity, and since speech relies on intact cochlear function across the whole frequency spectrum, PTAHF naturally emerges as the dominant speech predictor. Based on the differential age-related trajectories (PTAHF $R^2 = 0.641$ vs. PTA $R^2 = 0.158$), eHF thresholds cross clinical significance thresholds (≥25 dB HL, mild hearing loss) approximately 20 years earlier than conventional audiometry in the average individual, providing a critical window for early intervention and hearing conservation counseling.

Our finding that PTAHF (10–16 kHz) explains 34–36% of speech variance in univariate polynomial models ($R^2 = 0.34–0.36$), and 24–28% when adjusted for age and gender (adjusted $R^2 = 0.24–0.28$)—substantially more than conventional PTA (19–25% univariate, ~16% adjusted)—aligns with emerging evidence that standard audiometry systematically underestimates cochlear damage in aging populations. Speech contains substantial high-frequency acoustic cues extending into the 8–16 kHz range, with eHF energy accounting for approximately 10–20 dB below mid-frequency peaks when analyzed using physiologically-relevant equivalent rectangular bandwidth (ERB) filters[21,22]. Fricative contrasts (/s/ vs. /ʃ/), stop consonant bursts, and voice pitch harmonics all carry linguistically important information in this range, with the maximum audible low-pass cutoff for speech reaching approximately 13 kHz in young normal-hearing listeners[21]. Critically, speech radiation patterns are frequency-dependent: eHF energy radiates primarily toward the front of a talker, whereas low-frequency energy is omnidirectional[23]. In ecological "cocktail party" scenarios where target talkers face the listener but interfering talkers are rotated away, access to eHFs (up to 20 kHz) significantly improves SRTs by 1.4–2.0 dB relative to speech low-pass filtered at 8 kHz[24] . This frequency-dependent directivity makes eHF energy a salient cue for auditory object formation—facilitating both detection and segregation of target talkers from background talkers[23,24]. Traditional laboratory studies with all talkers facing the listener artificially equalize eHF energy across speech sources, obscuring these contributions[22,25]. Our findings demonstrate that eHF loss predicts speech outcomes even in controlled (4 dB SNR) conditions, suggesting that cumulative loss of eHF phonetic information and reduced segregation cues is clinically meaningful across diverse acoustic environments. The high prevalence of

undetected eHF loss among young adults with self-reported listening difficulties[26] supports extending routine audiometry beyond 8 kHz, particularly for individuals presenting with unexplained speech-in-noise complaints, ototoxicity exposure, or noise-related occupations.

Our finding that ABR wave I amplitude did not account for variability in speech-in-noise performance contrasts with predictions that cochlear synaptopathy (CS) would be a primary driver of SPiN deficits[27]. While animal studies convincingly demonstrate that noise exposure causes permanent loss of auditory nerve fiber (ANF) synapses with inner hair cells, leading to reduced ABR wave I without threshold elevation[14,15,28], human translation has proven challenging[29]. Recent human temporal bone studies provide direct anatomical evidence for age-related neural degeneration contributing to poor word-discrimination scores among individuals with similar audiometric thresholds[17,30,31]. However, multiple large-scale human studies have failed to find consistent relationships between putative physiological markers of CS and speech-in-noise performance[32–36]. The impact of CS on hearing may be most pronounced early in the pathological cascade, before substantial hair cell loss. Early-life noise exposure accelerates age-related cochlear synaptopathy in mice, with effects spreading from high-frequency regions initially affected to lower frequencies as animals age[37]. This noise-accelerated synaptopathy precedes threshold elevation by months, but once animals reach an advanced age (~20 months post-exposure), threshold shifts and OHC loss emerge, converging with aging-only pathology. This suggests CS may be an early vulnerability marker, but its relative contribution to hearing impairment diminishes once threshold-elevating pathologies (OHC loss, strial degeneration) accumulate. Despite mixed behavioral evidence[32–36], computational models strongly support CS impairing supra-threshold temporal coding[11,13,38]. Gerbils with noise-induced CS display impared neural discrimination of vowel-consonant-vowels in background noise at high sound levels (75 dB SPL), which was surprisingly improved at moderate levels (60 dB SPL), something reproduced in a phenomenological model.[39] A fundamental challenge is the inability to directly measure CS in living humans—all proposed markers are indirect and influenced by multiple cochlear pathologies. While CS undoubtedly occurs in aging and noise-exposed human cochleas, our study joins growing evidence that CS alone does not account for large inter-individual variability in speech-in-noise performance among individuals with clinically normal audiograms. Multiple peripheral pathologies (synaptopathy, demyelination, strial degeneration, OHC dysfunction) and central factors (cognitive control, attention, neural gain) likely interact to determine real-world listening abilities.

The selective interaction between Wave I latency and age for word recognition (WS, $p=0.040$) but not phoneme recognition (PS, $p=0.116$) reflects differential temporal integration demands: word-level comprehension requires extended temporal windows (400-600 ms) encompassing multiple phonemes, making it vulnerable to conduction delays that accumulate across longer sequences, whereas phoneme recognition operates within brief acoustic windows (50-100 ms) where conduction delays have minimal impact. Indeed, this dissociation is consistent with hierarchical processing models wherein word-level comprehension requires precise temporal integration across multiple phonemes and syllables—a process that becomes increasingly vulnerable when peripheral timing delays compound with age-related central processing decline. In contrast, phoneme recognition relies on shorter temporal integration windows and benefits from high acoustic-phonetic redundancy, rendering it relatively robust to peripheral timing variability even in older age[40–43]. These findings suggest that peripheral neural timing precision, rather than synaptic loss (as measured by Wave I amplitude), represents the rate-limiting step for complex speech understanding in aging. The absence of an interaction between Wave I amplitude and age ($p=0.584$) further challenges the cochlear synaptopathy hypothesis, indicating that age-related speech deficits are driven by the degradation of auditory nerve conduction velocity rather than cochlear synaptic loss.

WS appears to represent higher-level, plasticity-dependent speech comprehension, while PS reflects more fundamental, robust acoustic processing. This dissociation has important

implications for age-based intervention strategies. Our best models explain 32-38% of speech variance, leaving ~65% unexplained. What accounts for this residual? Working memory, attention, processing speed, and executive function all contribute to speech understanding, especially in challenging conditions[44,45]. Age-related cognitive decline may co-occur with hearing loss but follow an independent trajectory[46], and cognitive reserve may compensate for peripheral deficits[47]. Our models did not include cognitive assessments; future work incorporating measures like the Montreal Cognitive Assessment (MoCA) or digit span could partition variance between sensory and cognitive domains. Cortical auditory processing—phonemic decoding, lexical access, predictive coding—is entirely unobserved in click ABR, which reflects only the first 10 ms of the auditory pathway. Cortical evoked potentials (e.g., P1-N1-P2 complex, mismatch negativity) or speech-evoked envelope following responses may capture variance orthogonal to audiometry[33,48]. Aging affects cortical temporal processing, neural noise, and speech-specific encoding[49], potentially explaining inter-individual differences in speech perception despite similar audiograms. Vocabulary size, phonemic familiarity, native language experience, and lifetime auditory enrichment all modulate speech recognition[50,51]. Comprehensive models integrating eHF audiometry, cognitive assessments, cortical evoked potentials, and speech-in-noise testing across diverse listening environments (reverberation, competing talkers, audiovisual) are needed to fully map the determinants of real-world communication ability in aging.

This study is cross-sectional, and therefore I cannot determine whether eHF loss precedes or predicts accelerated speech decline. Being limited to a rather highly educated Swedish population, literacy and cognitive reserve may be higher in this sample than in general clinical populations. Consequently, external validity should be expanded to other languages, cultures, and educational backgrounds. In addition, we tested speech recognition in Swedish speech-weighted noise at +4 dB SNR, a moderately challenging condition. While this SNR is sufficient to unmask synaptopathy effects if present (as theoretically stronger than quiet), it is less demanding than everyday listening environments with multiple talkers, reverberation, or very low SNRs (≤0 dB). Future work should test a range of SNRs to determine whether synaptopathy or latency effects emerge at greater difficulty levels.

In conclusion, the present study establishes eHF pure-tone audiometry as the dominant clinical predictor of speech recognition performance and challenge the hypothesis that cochlear synaptopathy—as reflected in ABR wave I amplitude changes—contributes meaningfully to speech deficits in this large, community-dwelling population. Routine extension of hearing assessment beyond 8 kHz may provide clinically valuable information regarding speech prediction and communication prognosis, particularly in older adults and those with noise exposure history. Future studies examining hearing-cognition relationships should routinely include eHF audiometry to avoid incomplete characterization of hearing status. The weak explanatory power of synaptopathic ABR markers suggests that previous studies invoking hidden hearing loss as a mechanism should reconsider their interpretations in light of more complete audiometric data. Targeting eHF hearing preservation may be more relevant to communication outcomes than strategies focused on synaptic restoration, at least in populations with measurable eHF thresholds.

Materials and Methods

Ethics statement

The study was approved by the Regional Ethics Review Board in Stockholm (Dnr: 2016/1468-31/2); all participants provided written informed consent.

Study Population and design

Adult participants(>18 years old) were drawn from the Swedish Tinnitus Outreach Project (STOP)[52], a population-based cohort recruited 2016–2020 via the LifeGene national health study[53], web-based outreach, and community advertising. Participants did not receive any financial compensation. The participant flow diagram is reported in **Figure 1**. The present analysis included 263 participants (N=526 ears) who (1) reported no tinnitus (neither occasional nor constant) (2) completed audiometry from 0.125–16 kHz, (3) completed speech testing in quiet (N=511 ears), (4) underwent click ABR testing (N=324 ears). All were Swedish residents, and comprised 87% individuals with Swedish origin. Participants were highly educated (75% University level) and none had hearing aid or cochlear implant users. The ESIT screening questionnaire[54,55] was added to the platform in November 2018 and was answered by a subset of participants (n= 32 missing). This allowed to determine that none were diagnosed for the following ear disfunctions: ear barotrauma, Menières's Disease, Acoustic Neuroma, and Otosclerosis. Less than 2% of the participants had acoustic trauma, chronic otitis, or sudden hearing loss). Chemotherapy and head and neck radiotherapy was reported by 5 participants. Prevalence of other diseases is reported in **Supplemental Table 1**.

Audiometry and otoacoustic emissions

Consenting tinntius and non-tinnitus participants located in the Stockholm area or Skåne and who had answered large tinnitus survey[52] were invited for a battery of auditory measurements in a clinical setting. For external validity, auditory measurements were performed on 927 participants between August 2016 and December 2019 at three different sites: two different hospital audiology clinics in Stockholm (Karolinska Sjukhuset or KS Rosenlund, n = 778; and KS Solna, n = 120) and at the Department for Logopedics, Phoniatrics, and Audiology at Lund University (n = 29). For internal validity, the auditory measures and ABRs were collected by four licensed audiologists.

The session started with a short interview, confirming the participants' tinnitus statuses, and otoscopy and tympanometry (Madsen OTOflex 100 tympanometer, Otometrics; #8-04-04900, GN Resound AB, Göteborg) was performed. The DPOAEs were measured using the Madsen Capella 2 (Madsen Capella 2, Otometrics; #8-04-15320, GN Resound AB, Göteborg) with L1 = 65 dB SPL and L2 = 55 dB SPL at ten points per octave between F2 = 1 kHz and F2 = 10 kHz. Pure tone audiometry was performed at standard (.125–8 kHz) and high (8–16 kHz) frequencies using a Madsen Astera 2 clinical audiometer (MADSEN Astera2, Otometrics; #8-04-13100, GN Resound AB, Göteborg) and HDA 200 headphones (Sennheiser) using fixed-frequency Bekesy audiometry with a pulsed pure tone (550 ms, 50% duty cycle). Bone conduction thresholds were verified manually for thresholds >20 dB HL between 0.125 and 4 kHz.

Speech-in-noise (SPiN) testing was performed using the clinical standard method and material in Sweden, presenting 50 CNC-words in +4dB SNR per ear. Word Score (WS) and Phoneme Score (PS) were assessed using the Connected Speech Test (CST) or equivalent validated speech-in-noise paradigm presented at 50-75 dB SPL against a background noise at 40 dB SPL (signal-to-noise ratio +10 dB). Participants repeated back isolated words (WS; 50 key words) and individual phonemes (PS; 3 phonemes per word, accuracy scored by consonant clusters). WS was quantified as the percentage of words in noise repeated correctly, and PS was estimated as percentage of phonemes in noise repeated correctly. Both measures were converted to percent correct (0–100%), where higher scores indicate better performance.

The LDLs were assessed after SIN at all threshold frequencies. A 1.5-second pure tone was presented at 60 dB HL, increased by 5 dB, and presented again. The subject was instructed to inform the tester if the tone presented would have been uncomfortable to listen to for a longer period (1–2 minutes). The session concluded with measurements of the ABR, detailed below.

Auditory Brainstem Response (ABR)

Two setups for the ABR recordings were used in this study, the Madsen EP200 Chartr (ICS Chartr EP 200, Otometrics; #8-04-12701, GN Resound AB, Göteborg) and the Eclipse (Interacoustics, Stockholm) with ER-3A insert earphones and gold-wrapped tip trodes. The settings for both systems were identical, with high and low pass filters of .1 and 3 kHz, respectively, with 100 µs click stimuli of alternating polarity presented at 9.1 clicks/s at 90 dBnHL through insert earphones, with contralateral masking of −40 dB relative to the stimulus ear. Each recording consisted of 2,000 accepted clicks. The participants were relaxed in a reclined position in a dimly lit room during the recording. An audiologist identified relevant ABR features, i.e., wave I, III, and V peaks and troughs, then double-checked and, if necessary, corrected by two independent licensed audiologists blinded to the age of the participant. Amplitude was defined as peak to trough. In the present study, 190 non tinnitus ears were collected using the Chartr EP (Otometrics) and 146 ears with the Eclipse (Interacoustics). ABR Waves (N=331, n= 5 missing) were classified as having good morphology (n = 151, 46%), fair morphology (n = 31, 9%), no clear wave forms for waves I (n = 2, 1%), III (n = 2, 1%), V (n = 4, 1%), and poor morphology (n = 141, 43%). All ABR Waves with poor morphology were excluded (mean age 46.4 +/- SD 10.9; 59.3% women), leaving with with a sample of 190 ears of high quality wave forms. In a former study, test–retest reliability was established for these two devices: ICC for latencies 0.77–0.96 (good-to-excellent), ICC for amplitudes 0.26–0.47 (poor) on Chartr, 0.18-0.48 (poor) for Wave III and V but 0.96 (excellent) for Wave I amplitude on Eclipse[20]. Inter-wave latencies (I–V) were calculated.

Statistical analysis

Age-related trajectories were modelled for all audiometric, OAE, speech, and ABR measures using separate polynomial regressions with age as the predictor and each variable (PTA4, PTAHF, OAE_5273, WS, PS, W1.amp, W1.Plat, W3.Plat, W5.Plat) as the outcome. For each measure, we fitted linear, quadratic, and cubic age models with age mean-centered, and selected the best functional form using hierarchical likelihood ratio tests (LRT) in a sequential testing framework: linear models were compared to quadratic models, and if the quadratic term was significant ($p < 0.05$), quadratic models were compared to cubic models. The polynomial order was selected based on this stepwise LRT approach rather than AIC, to prioritize biological interpretability and avoid overfitting. We report R^2, LRT p-values comparing nested models, and the selected polynomial order for each relationship. Wave I amplitude analyses were restricted to the Eclipse subset, whereas all other variables used the full cohort; the best-fitting age functions from these models were then used to generate the smooth age trajectories.

Univariate linear regression models were constructed with WS or PS as the outcome and audiometric (PTA4, PTAHF), OAE (OAE_5273), or ABR (W1.amp, W1.Plat, W3.Plat, W5.Plat) predictors separately, using centered continuous predictors. Polynomial terms for each predictor were selected using the same hierarchical LRT approach ($\alpha = 0.05$), testing linear vs. quadratic, then quadratic vs. cubic forms sequentially. Models were evaluated using R^2, LRT p-values, and parameter estimates (β, SE, t, p), ensuring consistency in polynomial selection across univariate and multivariable frameworks. For the multivariable analysis predicting WS and PS, we followed a prespecified four-phase hierarchical modeling strategy, with each phase tested against the previous using LRT: **Phase 0** included demographic predictors only (age, age², sex); **Phase 1** added peripheral hearing-loss predictors (PTA4, PTAHF, OAE) with LRT-selected polynomial terms; **Phase 2** added ABR measures (W1.amp, W1.Plat, W3.Plat, W5.Plat) with LRT-selected polynomial terms to capture neural timing and amplitude effects; and **Phase 3** added biologically motivated age interactions (Age × W1.Latency, Age × W1.Amplitude) to test whether

neural predictors show differential age-related vulnerability. Incremental variance explained (ΔR^2) and LRT p-values were computed for each phase transition. For ABR amplitude analyses, we restricted to the Eclipse subset ($N \approx 134–140$) to maximize reliability, and hardware (Chartr vs. Eclipse) was included as a covariate in full-cohort models.

Multicollinearity was assessed via variance inflation factors (VIF < 5 for all models). Robust heteroscedasticity-consistent standard errors (HC3) from the sandwich–car framework were used because Breusch–Pagan tests indicated variance heterogeneity (e.g., $\chi^2(9) = 35.04$, $p = 0.006$ for WS), and HC3-based inferences are reported for all regression coefficients. Model fit was assessed using adjusted R^2, and diagnostic residual analyses in the full sample ($n = 298$), with sensitivity checks confirming that exclusion of extreme outliers did not materially change the pattern of significant predictors. Statistical significance was defined as $\alpha = 0.05$ (two-tailed), and p-values are reported in scientific notation for values < 0.0001 in the main and supplementary tables.

All analyses were performed in R version 4.4.0 (R Foundation for Statistical Computing) using the tidyverse (v2.0.0) for data wrangling, car (v3.1-2) for regression diagnostics and HC3 standard errors, sandwich (v3.0-2) for robust variance estimation, broom for model summaries, and ggplot2 (v4.0.0) with patchwork (v1.2.0) for visualization.

Data availability: Anonymized individual participant data (audiometry, OAE, ABR, speech scores) will be soon made available upon approval by the Swedish ethics regulations.

Code availability: Statistical analysis scripts (R) and figure generation code are available at https://github.com/translational-audiology-lab/STOP_Speech

Funding
This analysis was not funded. The auditory data acquisition was supported by grants from the Swedish Research Council (K2014-99X-22478-01-3) and Tysta Skolan.

Acknowledgments
I am indebted to Jonathon Whitton and Golbarg Mehraei (Decibel Therapeutics Inc.), Hugo Westerlund and Constance Leineweber (SLOSH) and Nancy Pedersen (LG), Inger Uhlen and Andra Lazar (Karolinska Hospital) for their collaboration, support and dedication to STOP. I am very thankful to Otometrics (Clément Sanchez, Aud) and Interacoustics (Stefan Pettersson, Diatec Sweden AB) for sharing their equipment throughout the project duration. I warmly thank Niklas K. Edvall who collected the majority of the STOP data, and also Rickard Redgård, Anders Sjöstrand, and Laura Inan for their contributions in the acquisition of auditory data. I express my gratitude to all other collaborators of TINNET, ESIT, UNITI who guided in the initial steps of STOP, including Rilana Cima, Birgit Mazurek, and Jose Antonio Lopez-Escamez. I would like to thank all the participants who joined the project and supported this study, as well as staff from the Department of Physiology and Pharmacology, Karolinska Institutet.

Author Contributions CRC designed and directed the research, analyzed the data, generated the figures and tables, and wrote the manuscript.

Competing Interest Statement: CRC was supported in 2020 by the UK National Institute for Health Research (NIHR) Biomedical Research Centre but the views expressed herein are his own and do not represent those of NIHR nor the UK Department of Health and Social Care. CRC is also a member of the Scientific Advisory Board of the American Tinnitus Association and member of the Professional Advisory Committee of the Tinnitus UK.

References

1. Haile, L. M. *et al.* Hearing loss prevalence and years lived with disability, 1990–2019: findings from the Global Burden of Disease Study 2019. *The Lancet* **397**, 996–1009 (2021).
2. Livingston, G. *et al.* Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* **396**, 413–446 (2020).
3. An, Y. Y. *et al.* Association of Hearing Loss With Anatomical and Functional Connectivity in Patients With Mild Cognitive Impairment. *JAMA Otolaryngol Head Neck Surg* **149**, 571 (2023).
4. Ishak, E. *et al.* Population Attributable Fraction of Incident Dementia Associated With Hearing Loss. *JAMA Otolaryngol Head Neck Surg* **151**, 568 (2025).
5. Loughrey, D. G., Kelly, M. E., Kelley, G. A., Brennan, S. & Lawlor, B. A. Association of Age-Related Hearing Loss With Cognitive Function, Cognitive Impairment, and Dementia: A Systematic Review and Meta-analysis. *JAMA Otolaryngol Head Neck Surg* **144**, 115 (2018).
6. Davis, A., Smith, P., Ferguson, M., Stephens, D. & Gianopoulos, I. Acceptability, benefit and costs of early screening for hearing disability: a study of potential screening tests and models. *Health Technol Assess* **11**, 1–294 (2007).
7. Cantuaria, M. L. *et al.* Hearing Loss, Hearing Aid Use, and Risk of Dementia in Older Adults. *JAMA Otolaryngol Head Neck Surg* **150**, 157 (2024).
8. Stevenson, J. S., Clifton, L., Kuźma, E. & Littlejohns, T. J. Speech-in-noise hearing impairment is associated with an increased risk of incident dementia in 82,039 UK Biobank participants. *Alzheimer's & Dementia* **18**, 445–456 (2022).
9. Motlagh Zadeh, L. *et al.* Extended high-frequency hearing enhances speech perception in noise. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 23753–23759 (2019).
10. Cederroth, C. R., Dyhrfjeld-Johnsen, J. & Canlon, B. Pharmacological Approaches to Hearing Loss. *Pharmacol Rev* PHARMREV-AR-2024-001195 (2024) doi:10.1124/pharmrev.124.001195.
11. Vasilkov, V., Liberman, M. C. & Maison, S. F. Modeling the impact of cochlear nerve degeneration on speech recognition performance. *PLoS One* **20**, e0336299 (2025).
12. Ponsot, E., Devolder, P., Dhooge, I. & Verhulst, S. Age-related decline in neural phase-locking to envelope and temporal fine structure revealed by frequency following responses: A potential signature of cochlear synaptopathy impairing speech intelligibility. Preprint at https://doi.org/10.1101/2024.12.11.628010 (2024).
13. Grant, K. J. *et al.* Electrophysiological markers of cochlear function correlate with hearing-in-noise performance among audiometrically normal subjects. *Journal of Neurophysiology* **124**, 418–431 (2020).
14. Furman, A. C., Kujawa, S. G. & Liberman, M. C. Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *Journal of neurophysiology* **110**, 577–86 (2013).
15. Kujawa, S. G. & Liberman, M. C. Adding insult to injury: cochlear nerve degeneration after 'temporary' noise-induced hearing loss. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29**, 14077–85 (2009).
16. Liberman, M. C. & Kujawa, S. G. Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. *Hearing research* https://doi.org/10.1016/j.heares.2017.01.003 (2017) doi:10.1016/j.heares.2017.01.003.

17. Wu, P.-Z., O'Malley, J. T., de Gruttola, V. & Liberman, M. C. Primary Neural Degeneration in Noise-Exposed Human Cochleas: Correlations with Outer Hair Cell Loss and Word-Discrimination Scores. *J Neurosci* **41**, 4439–4447 (2021).

18. Liberman, M. C., Epstein, M. J., Cleveland, S. S., Wang, H. & Maison, S. F. Toward a Differential Diagnosis of Hidden Hearing Loss in Humans. *PLoS One* **11**, e0162726 (2016).

19. Grant, K. J. *et al.* Electrophysiological markers of cochlear function correlate with hearing-in-noise performance among audiometrically normal subjects. *J Neurophysiol* **124**, 418–431 (2020).

20. Edvall, N. K. *et al.* Alterations in auditory brain stem response distinguish occasional and constant tinnitus. *Journal of Clinical Investigation* **132**, e155094 (2022).

21. Monson, B. B., Lotto, A. J. & Story, B. H. Analysis of high-frequency energy in long-term average spectra of singing, speech, and voiceless fricatives. *J Acoust Soc Am* **132**, 1754–1764 (2012).

22. Levy, S. C., Freed, D. J., Nilsson, M., Moore, B. C. J. & Puria, S. Extended High-Frequency Bandwidth Improves Speech Reception in the Presence of Spatially Separated Masking Speech. *Ear Hear* **36**, e214-224 (2015).

23. Hunter, L. L. *et al.* Extended high frequency hearing and speech perception implications in adults and children. *Hear Res* **397**, 107922 (2020).

24. Monson, B. B., Rock, J., Schulz, A., Hoffman, E. & Buss, E. Ecological cocktail party listening reveals the utility of extended high-frequency hearing. *Hearing Research* **381**, 107773 (2019).

25. Moore, B. C. J., Füllgrabe, C. & Stone, M. A. Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task. *J Acoust Soc Am* **128**, 360–371 (2010).

26. Aryal, S., Trevino, M., Rodrigo, H. & Mishra, S. Is Noise Exposure Associated With Impaired Extended High Frequency Hearing Despite a Normal Audiogram? A Systematic Review and Meta-Analysis. *Trends in Hearing* **29**, 23312165251343757 (2025).

27. Liberman, M. C. & Kujawa, S. G. Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. *Hear Res* **349**, 138–147 (2017).

28. Mackey, C. A. *et al.* Correlations between cochlear pathophysiology and behavioral measures of temporal and spatial processing in noise exposed macaques. *Hearing Research* **401**, 108156 (2021).

29. Bramhall, N. *et al.* The search for noise-induced cochlear synaptopathy in humans: Mission impossible? *Hearing research* **377**, 88–103 (2019).

30. Wu, P., O'Malley, J. T., de Gruttola, V. & Liberman, M. C. Age-Related Hearing Loss Is Dominated by Damage to Inner Ear Sensory Cells, Not the Cellular Battery That Powers Them. *J. Neurosci.* **40**, 6357–6366 (2020).

31. Wu, P. Z. *et al.* Primary Neural Degeneration in the Human Cochlea: Evidence for Hidden Hearing Loss in the Aging Ear. *Neuroscience* **407**, 8–20 (2019).

32. Coelho De Sousa, S. L., Marrufo Pérez, M. I., Johannesen, P. T., Gómez Álvarez, M. & Lopez-Poveda, E. A. No association between idiopathic hidden hearing loss and behavioral adaptation to noise in humans. *Hearing Research* **464**, 109321 (2025).

33. Ponsot, E., Devolder, P., Dhooge, I. & Verhulst, S. Age-Related Decline in Neural Phase-Locking to Envelope and Temporal Fine Structure Revealed by Frequency Following Responses: A Potential Signature of Cochlear Synaptopathy Impairing Speech Intelligibility. *J Assoc Res Otolaryngol* **26**, 249–270 (2025).

34. Fulbright, A. N. C., Le Prell, C. G., Griffiths, S. K. & Lobarinas, E. Effects of Recreational Noise on Threshold and Suprathreshold Measures of Auditory Function. *Semin Hear* **38**, 298–318 (2017).

35. Grinn, S. K., Wiseman, K. B., Baker, J. A. & Le Prell, C. G. Hidden Hearing Loss? No Effect of Common Recreational Noise Exposure on Cochlear Nerve Response Amplitude in Humans. *Front Neurosci* **11**, 465 (2017).

36. Guest, H., Munro, K. J., Prendergast, G., Millman, R. E. & Plack, C. J. Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. *Hearing research* **364**, 142–151 (2018).

37. Fernandez, K. A., Jeffers, P. W. C., Lall, K., Liberman, M. C. & Kujawa, S. G. Aging after Noise Exposure: Acceleration of Cochlear Synaptopathy in "Recovered" Ears. *J. Neurosci.* **35**, 7509–7520 (2015).

38. Grant, K. J. *et al.* Predicting neural deficits in sensorineural hearing loss from word recognition scores. *Sci Rep* **12**, 8929 (2022).

39. Monaghan, J. J. M., Garcia-Lazaro, J. A., McAlpine, D. & Schaette, R. Hidden Hearing Loss Impacts the Neural Representation of Speech in Background Noise. *Current Biology* **30**, 4710-4721.e4 (2020).

40. Tremblay, K. L., Piskosz, M. & Souza, P. Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical Neurophysiology* **114**, 1332–1343 (2003).

41. Pichora-Fuller, M. K. & Singh, G. Effects of Age on Auditory and Cognitive Processing: Implications for Hearing Aid Fitting and Audiologic Rehabilitation. *Trends in Amplification* **10**, 29–59 (2006).

42. Presacco, A., Simon, J. Z. & Anderson, S. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *Journal of neurophysiology* **116**, 2346–2355 (2016).

43. Anderson, S., Parbery-Clark, A., White-Schwoch, T., Drehobl, S. & Kraus, N. Effects of hearing loss on the subcortical representation of speech cues. *The Journal of the Acoustical Society of America* **133**, 3030–3038 (2013).

44. Parthasarathy, A., Hancock, K. E., Bennett, K., DeGruttola, V. & Polley, D. B. Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. *eLife* **9**, e51419 (2020).

45. Tóth, B. *et al.* Attention and speech-processing related functional brain networks activated in a multi-speaker environment. *PLoS One* **14**, e0212754 (2019).

46. Tran, Y. *et al.* Co-occurring Hearing Loss and Cognitive Decline in Older Adults: A Dual Group-Based Trajectory Modeling Approach. *Front Aging Neurosci* **13**, 794787 (2021).

47. Levett, B. A. *et al.* Hearing impairment and dementia: cause, catalyst or consequence? *J Neurol* **272**, 402 (2025).

48. Garrett, M. *et al.* Deciphering Compromised Speech-in-Noise Intelligibility in Older Listeners: The Role of Cochlear Synaptopathy. *eNeuro* **12**, ENEURO.0182-24.2024 (2025).

49. Karunathilake, I. M. D. *et al.* Effects of aging on cortical representations of continuous speech. *J Neurophysiol* **129**, 1359–1377 (2023).

50. Magnuson, J. S., Mirman, D., Luthra, S., Strauss, T. & Harris, H. D. Interaction in Spoken Word Recognition Models: Feedback Helps. *Front Psychol* **9**, 369 (2018).

51. Blasingame, M. & Bradlow, A. R. Early versus Extended Exposure in Speech Perception Learning: Evidence from Switched-Dominance Bilinguals. *Languages (Basel)* **5**, 39 (2020).

746 52. Edvall, N. K. *et al.* Alterations in auditory brain stem response distinguish occasional and
747 constant tinnitus. *The Journal of clinical investigation* **132**, (2022).

748 53. Almqvist, C. *et al.* LifeGene--a large prospective population-based study of global relevance.
749 *European journal of epidemiology* **26**, 67–77 (2011).

750 54. Genitsaridi, E. *et al.* Standardised profiling for tinnitus research: The European School for
751 Interdisciplinary Tinnitus Research Screening Questionnaire (ESIT-SQ). *Hearing research*
752 **377**, 53–60 (2019).

753 55. Müller, K. *et al.* Validation of Online Versions of Tinnitus Questionnaires Translated into
754 Swedish. *Frontiers in Aging Neuroscience* **8**, 272 (2016).

755

Figure Legends and Tables

Figure 1. Participant flow diagram. Overview of participants excluded and recruited for the electrophysiological study.

Figure 2: Audiometric patterns across age groups in the STOP cohort
Age-stratification of pure-tone audiometry (a), loudness discomfort levels (b), distortion products of otoacoustic emissions (c). Raincloud plots for word (d) and phoneme (e) scores individual points (left), summary statistics as median and IQR (middle) and representation of distribution with one-sided violin plots (right). Auditory brainstem responses are displayed for Wave I amplitiue (f), and latencies of Wave I (g), III (h), V (i). Wave I amplitude are only shown for Eclipse hardware, the Charter ones being unreliable. For Wave latencies, data from the Charter and the Eclipse are disintguishable in the raw data (left) with ligter coloring for the Charter, corresponding to different means and sem displayed on the right of the bar plot as “c” and “e” respectively. Age groups are displayed in different colors: >35 (green), 35-44 (light blue); 45-54 (blue); 55-64 (violet), 65-74 (red), and >75 (orange). Data relates to Tables S2-4.

Figure 3: Age related changes in audiometric and electrophysiological variables
Univariate polynomial regressions (linear vs. quadratic vs. cubic fits) from Fig.2 are displayed for (a) PTA4 (average of 0.5, 1, 2, 4 kHz), (b) PTAHF (average of 10, 12.5, 14, 16 kHz), (c) DPOAEs at 5273 Hz (OAE5273), (d) WS, (e) PS, (f) Wave I amplitude (Eclipse only), and latencies for Wave I (g), III (h) and V (i) in function of age. Best fits are reported for linear (blue), and quadratic (orange) with shaded regions as 95% confidence intervals, and R^2 and p values on top of the respective panels. Data relates to Table S5.

Figure 4: Univariate correlations with data-driven polynomial fits
Word score (WS) versus (a) PTA4, (b) PTAHF, (c) Wave I amplitude (Eclipse only), and latencies of Wave I (d), III (e), and V (f). Phoneme score versus (g) PTA4, (h) PTAHF, (i) Wave I amplitude (Eclipse only), and latencies of Wave I (j), III (k), and V (l). Best fits are reported for linear (blue), and quadratic (orange), and cubic (red) with shaded regions as 95% confidence intervals, and R^2 and p values on top of the respective panels. Data relates to Table S7.

Figure 5. Comparison of the mechanisms underlying WS and PS. (a) WS (red) exhibits accelerating quadratic decline with age (age^2 $\beta=-0.010$, $p=0.012$), while PS (blue) shows linear decline with age (age^2 $p=0.287$). Solid lines represent predicted values from multivariable models controlling for hearing loss, neural timing measures, outer hair cell function, and gender. Shaded regions indicate 95% confidence intervals. Open circles denote raw data points (n=298 ears). (b) Forest plots comparing effect sizes (HC3 robust SE) between WS and PS for key predictors. W1.Amplitude shows no significant relationship to either outcome. Age^2 and Age×W1.Latency effects are significantly stronger for WS, while $W3.Latency^3$ effects are similarly strong for both. Summary statistics are in **Table S10.** (c) Four age groups (45, 55, 65, 75 years) show divergent trajectories as W1.Latency (neural timing) increases (**Table S11**). Older adults (red, 75y) experience steeper Word Score decline (solid line) compared to Phoneme Score decline (dashed line). Young adults (green, 45y) show minimal decline across both measures. (d) High correlation (r=0.934) masks differential aging patterns. Points colored by age reveal that older adults (red) have relatively preserved PS but reduced WS, indicating dissociable age-related mechanisms (**Table S12**). All panels control for age, age^2, sex, low-frequency hearing loss (PTA4, $PTA4^2$), high-frequency hearing loss (PTAHF, $PTAHF^2$, $PTAHF^3$), outer hair cell function (OAE, OAE^2, OAE^3), W1.Latency, Age×W1.Latency, W1.Amplitude, Age×W1.Amplitude, W3.Latency, $W3.Latency^2$, $W3.Latency^3$, and W5.Latency. HC3 heteroscedasticity-consistent standard errors used to account for unequal variances in speech perception measures. N = 298 ears from older

806 adults assessed on speech recognition, auditory brainstem response measures, and audiometric
807 thresholds.
808
809
810
811

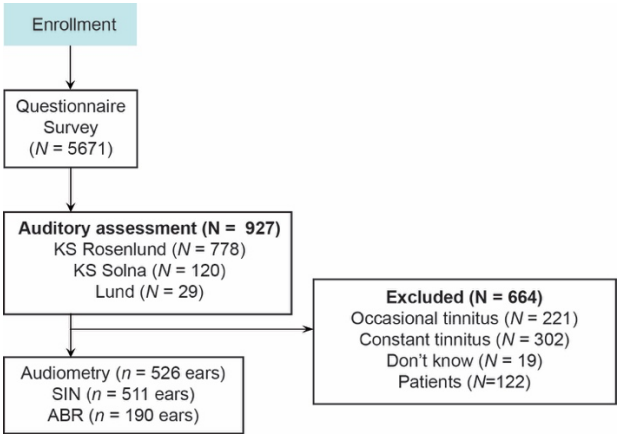
Table 1. Demographics from 263 tinnitus-free STOP participants (526 ears).

Variable	Mean +/- SD or N (%)	Range
Age (years)	49.2 +/- 13.6	20.0 to 84.0
Gender (woman)	356 (68%)	-
PTA4 (0.5-4 kHz) dB HL	5.2 +/- 10.7	-7.4 to 100.0
PTAHF (10-16 kHz) dB HL	36.5 +/- 22.5	-12.2 to 78.0
OAE_5273 (dB SNR)	19.8 +/- 10.1	-31.0 to 44.6
Word score (%)	75.1 +/- 10.3	12.0 to 94.0
Phoneme score (%)	89.2 +/- 6.0	48.0 to 97.3
Wave I amplitude (Eclipse uV)	0.4 +/- 0.2	0.1 to 1.0
Wave I latency (ms)	1.4 +/- 0.2	1.0 to 1.9
Wave V latency (ms)	3.6 +/- 0.2	3.0 to 4.2
I-V inter-wave latency (ms)	5.4 +/- 0.4	4.5 to 6.3

Table 2. Comparison of Age × W1.Amplitude vs Age × W1.Latency Interactions on speech predictions. Significant p values are shown in bold.

Interaction	Outcome	β	SE	95% CI	p-value
Age × W1.Amp	WS	0.17	0.31	-0.43 to 0.77	0.584
Age × W1.Amp	PS	0.19	0.20	-0.21 to 0.58	0.354
Age × W1.Lat	WS	-0.76	0.37	-1.49 to -0.04	**0.040**
Age × W1.Lat	PS	-0.37	0.23	-0.82 to 0.09	0.116

822 **Figure 1**
823



824

825

Figure 2

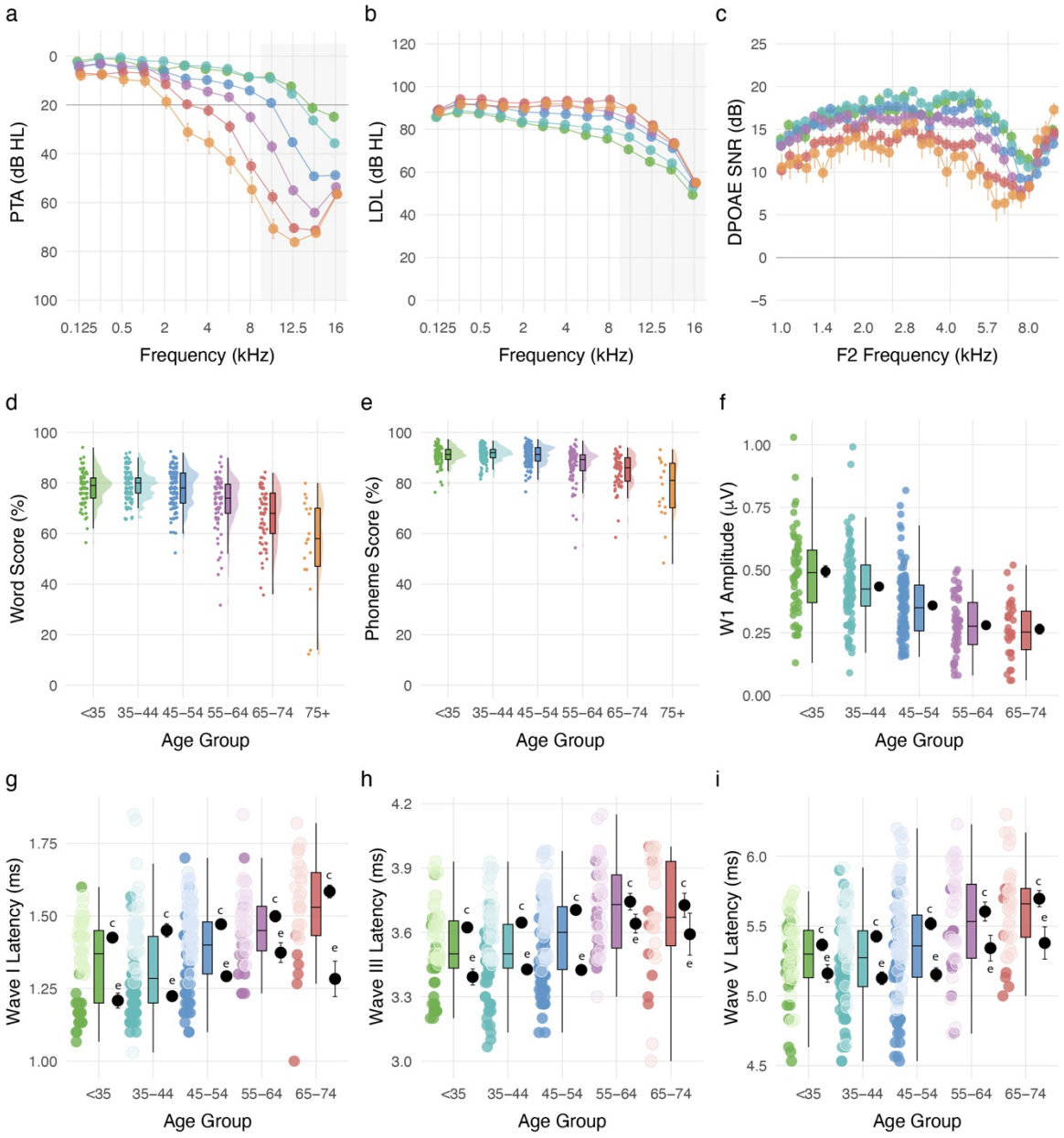

a
PTA (dB HL)
Frequency (kHz)
b
LDL (dB HL)
Frequency (kHz)
c
DPOAE SNR (dB)
F2 Frequency (kHz)
d
Word Score (%)
Age Group
e
Phoneme Score (%)
Age Group
f
W1 Amplitude (µV)
Age Group
g
Wave I Latency (ms)
Age Group
h
Wave III Latency (ms)
Age Group
i
Wave V Latency (ms)
Age Group

Figure 3

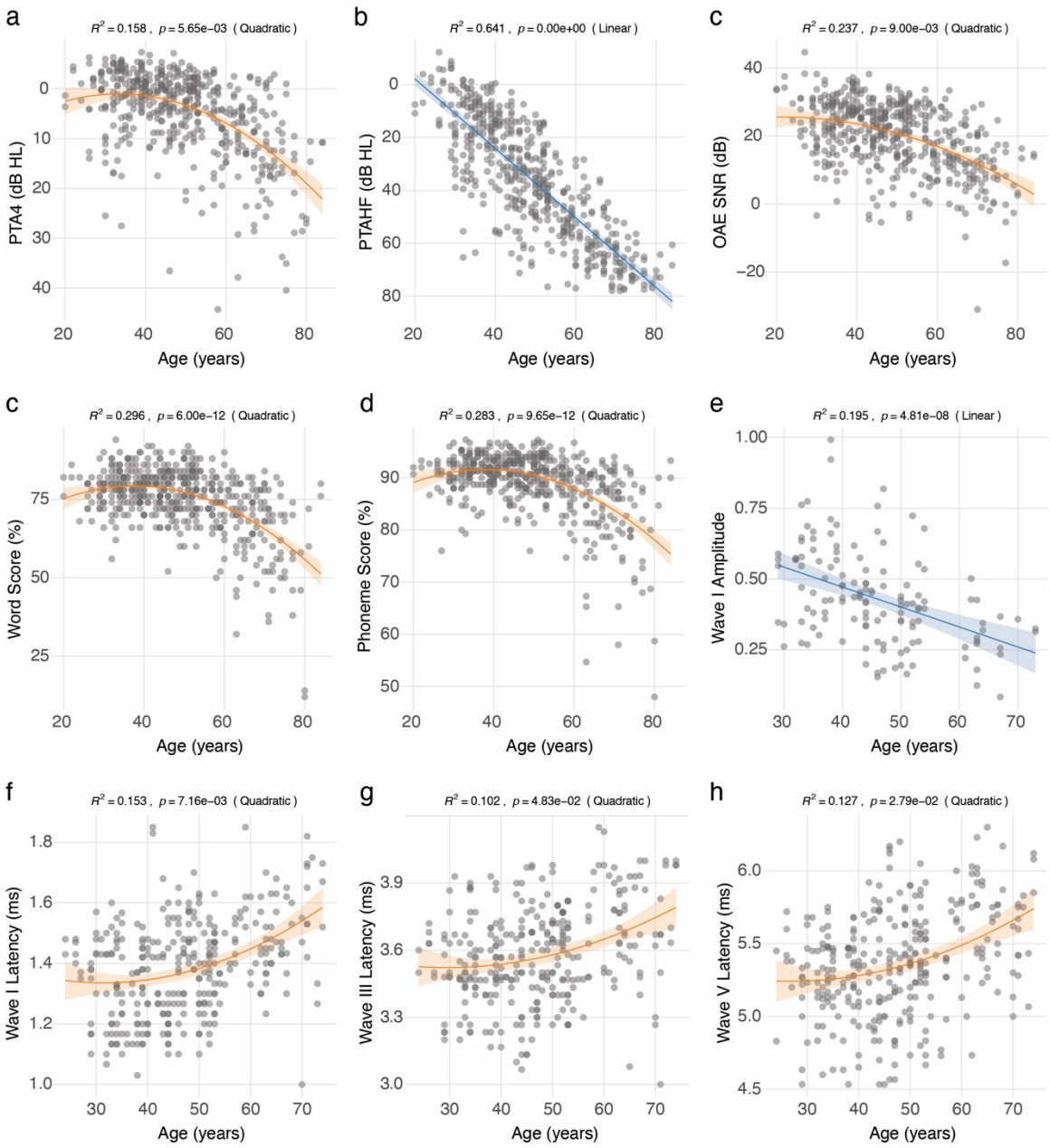
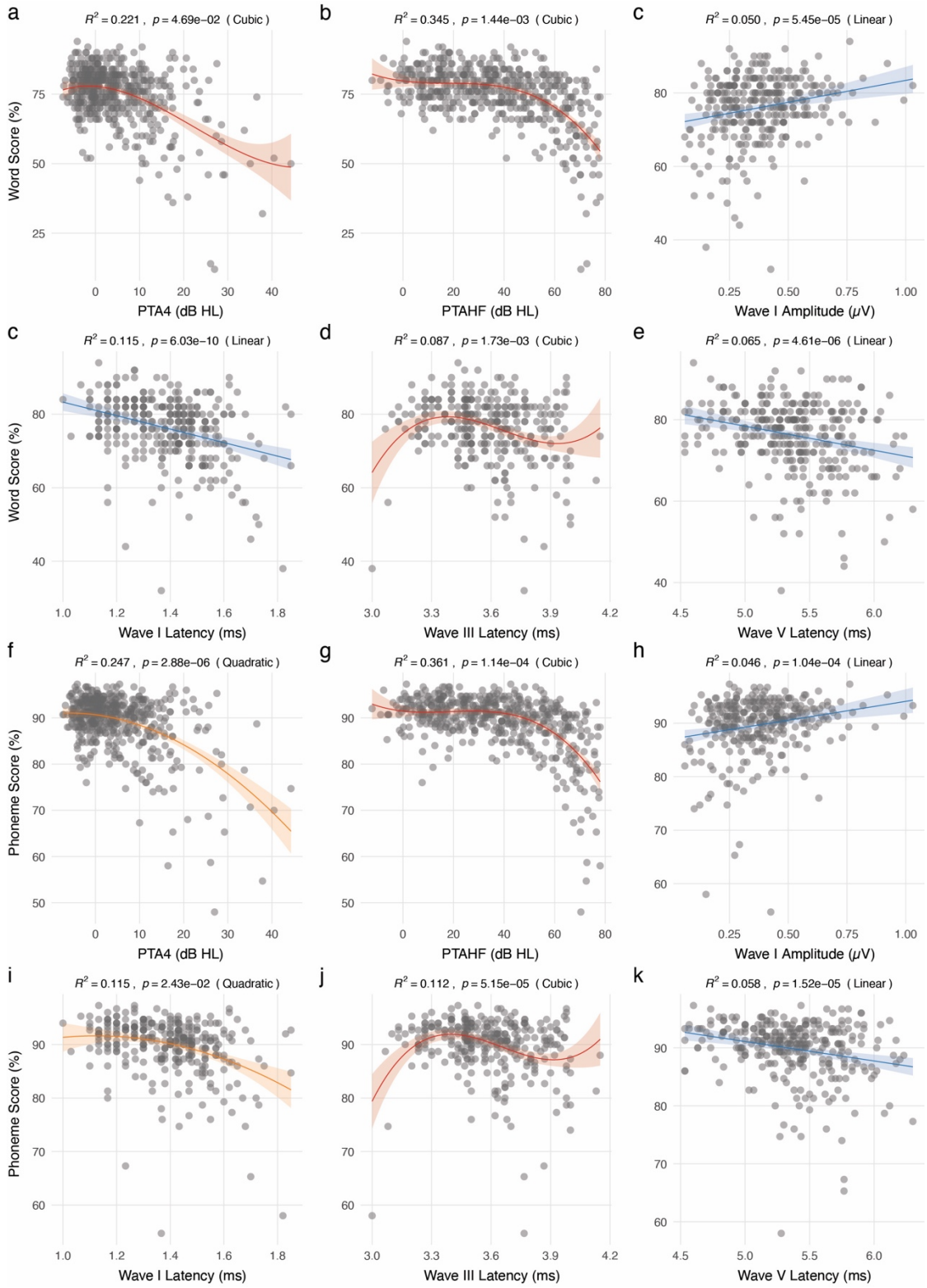
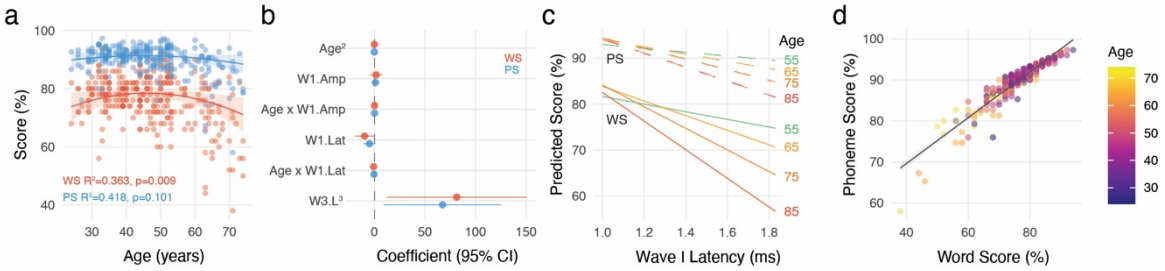

a
$R^2 = 0.158$, $p = 5.65e-03$ (Quadratic)
PTA4 (dB HL)
Age (years)
b
$R^2 = 0.641$, $p = 0.00e+00$ (Linear)
PTAHF (dB HL)
Age (years)
c
$R^2 = 0.237$, $p = 9.00e-03$ (Quadratic)
OAE SNR (dB)
Age (years)
c
$R^2 = 0.296$, $p = 6.00e-12$ (Quadratic)
Word Score (%)
Age (years)
d
$R^2 = 0.283$, $p = 9.65e-12$ (Quadratic)
Phoneme Score (%)
Age (years)
e
$R^2 = 0.195$, $p = 4.81e-08$ (Linear)
Wave I Amplitude
Age (years)
f
$R^2 = 0.153$, $p = 7.16e-03$ (Quadratic)
Wave I Latency (ms)
Age (years)
g
$R^2 = 0.102$, $p = 4.83e-02$ (Quadratic)
Wave III Latency (ms)
Age (years)
h
$R^2 = 0.127$, $p = 2.79e-02$ (Quadratic)
Wave V Latency (ms)
Age (years)

Figure 4


a
$R^2 = 0.221$, $p = 4.69e-02$ (Cubic)
Word Score (%)
PTA4 (dB HL)
b
$R^2 = 0.345$, $p = 1.44e-03$ (Cubic)
PTAHF (dB HL)
c
$R^2 = 0.050$, $p = 5.45e-05$ (Linear)
Wave I Amplitude (μV)
c
$R^2 = 0.115$, $p = 6.03e-10$ (Linear)
Word Score (%)
Wave I Latency (ms)
d
$R^2 = 0.087$, $p = 1.73e-03$ (Cubic)
Wave III Latency (ms)
e
$R^2 = 0.065$, $p = 4.61e-06$ (Linear)
Wave V Latency (ms)
f
$R^2 = 0.247$, $p = 2.88e-06$ (Quadratic)
Phoneme Score (%)
PTA4 (dB HL)
g
$R^2 = 0.361$, $p = 1.14e-04$ (Cubic)
PTAHF (dB HL)
h
$R^2 = 0.046$, $p = 1.04e-04$ (Linear)
Wave I Amplitude (μV)
i
$R^2 = 0.115$, $p = 2.43e-02$ (Quadratic)
Phoneme Score (%)
Wave I Latency (ms)
j
$R^2 = 0.112$, $p = 5.15e-05$ (Cubic)
Wave III Latency (ms)
k
$R^2 = 0.058$, $p = 1.52e-05$ (Linear)
Wave V Latency (ms)

838 **Figure 5**

839



840