

# Increased listening effort and cochlear neural degeneration underlie behavioral deficits in speech perception in noise in normal hearing middle-aged adults

Maggie E Zink, Leslie Zhen, Jacie R McHaney, Jennifer Klara, Kimberly Yurasits, Victoria Cancel, Olivia Flemm, Claire Mitchell, Jyotishka Datta, Bharath Chandrasekaran, Aravindakshan Parthasarathy 

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States • Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, United States • Department of Bioengineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, United States • Department of Otolaryngology, School of Medicine, University of Pittsburgh, Pittsburgh, United States

 [https://en.wikipedia.org/wiki/Open\\_access](https://en.wikipedia.org/wiki/Open_access)

 Copyright information

## eLife Assessment

This study is part of an ongoing effort to clarify the effects of cochlear neural degeneration on auditory processing in listeners with normal audiograms. Here the authors provide **important** new data demonstrating associations between cochlear neural degeneration, non-invasive assays of auditory processing, and speech perception. Based on a cross-species comparison, these findings pose **compelling** evidence that cochlear synaptopathy is associated with a significant part of hearing difficulties in complex environments for some listeners with normal hearing thresholds, such as older individuals.

<https://doi.org/10.7554/eLife.102823.1.sa2>

## Abstract

Middle-age is a critical period of rapid changes in brain function that presents an opportunity for early diagnostics and intervention for neurodegenerative conditions later in life. Hearing loss is one such early indicator linked to many comorbidities later in life. However, current clinical tests fail to capture hearing difficulties for ~10% of middle-aged adults seeking help at hearing clinics. Cochlear neural degeneration (CND) could play a role in these hearing deficits, but our current understanding is limited by the lack of objective diagnostics and uncertainty regarding its perceptual consequences. Here, using a cross-species approach, we measured envelope following responses (EFRs) – neural ensemble responses to sound originating from the peripheral auditory pathway – in young and middle-aged adults with normal audiometric thresholds, and compared these responses to young and middle-aged Mongolian gerbils, where CND was histologically confirmed. We observed near identical

changes in EFRs across species that were associated with CND. Perceptual effects measured as behavioral readouts showed deficits in the most challenging listening conditions and were associated with CND. Additionally, pupil-indexed listening effort increased even at moderate task difficulties where behavioral outcomes were matched. Our results reveal perceptual deficits in middle-aged adults driven by CND and increases in listening effort, which may result in increased listening fatigue and conversational disengagement.

## Introduction

Age-related hearing loss, defined as declines in hearing sensitivity, is exceedingly common; according to some estimates, ~45 million adults in the United States over 50 years of age have age-related hearing loss that is significant enough to interfere with communication (1). Untreated hearing loss decreases quality of life and is thought to be the single-largest modifiable risk factor in middle-age for other age-related comorbidities such as cognitive impairment and dementia (2). However, current measures of hearing sensitivity fail to capture critical aspects of real-world hearing difficulties in this population (3, 4). Hearing difficulties experienced by up to 10% of adults seeking help in the hearing clinic are ‘hidden’ to current diagnostic procedures (3–6). Peripheral deafferentation caused by cochlear neural degeneration (CND) may underlie many of these perceptual difficulties (7, 8). Anatomical evidence for progressive CND with aging is clear – postmortem studies using human temporal bones estimate a 40% deafferentation caused by CND by the fifth decade of life (9–11). CND causes neural coding deficits in the peripheral auditory pathway, affecting the faithful representation of spectrotemporally complex auditory stimuli (12–14). But the evidence linking CND with perceptual deficits is mixed – current assessments of perceptual deficits associated with CND primarily focus on behavioral measures of speech in noise, with mixed evidence of deficits in individuals with putative CND (15–18).

Two challenges impede our understanding of the perceptual consequences of CND. First, while many non-invasive markers of CND have been proposed and validated in animal models (7, 14, 19, 20), non-invasive estimates of putative CND in humans cannot be confirmed with histological assessment of synapses in the same participants. Cross-species comparative studies and computational modeling provide promising avenues for overcoming this gap (21, 22). Secondly, behavioral readouts of perceptual difficulties in humans show mixed results with putative CND depending on the specific test used and degree of spectrotemporal and contextual information provided in that test (17, 23, 24). The most promising tests for CND are ones with no linguistic context and short spectrotemporal processing windows (24, 25). However, these behavioral readouts may minimize subliminal changes in perception that are reflected in listening effort but *not* in accuracies (26–28). Here, we used a cross-species approach, combined with simultaneous measurements of behavior and listening effort, to show that CND is associated with decreased neural coding fidelity and increased listening effort in middle-aged adults with normal audiometric thresholds. We measured putative CND using the envelope following response (EFR) to rapid (~1000Hz) modulation frequencies – a suggested marker for CND (12, 14). Cross-species comparisons with identical recordings in a low-frequency hearing animal model, the Mongolian gerbil, confirmed that decreases in EFRs were selective only for responses with generators in the auditory nerve. These EFRs were also associated with histologically-confirmed CND. In the human model, we simultaneously measured pupil-indexed listening effort in participants as they performed a speech-in-noise task and show that increased listening effort was present despite matched behavioral accuracies. These results point to hitherto underexplored aspects of auditory perceptual difficulties associated with listening effort and CND.

## Results

“Normal” hearing middle-aged adults show evidence of peripheral neural coding deficits that are associated with CND. Middle-aged (MA, 40–55 years) listeners were recruited to participate in this study, and their responses were compared to that of young adult (YA, 18–25 years) listeners (**Fig. 1A**). All participants had clinically normal hearing thresholds and spoke fluent American English. Participants had normal otoscopy by visual examination and air conduction thresholds below 25 dB HL for octave frequencies between 250 Hz to 8 kHz (**Fig. 1B**, **Table 1**), consistent with WHO guidelines for normal hearing (29). Threshold differences were exaggerated in MAs at extended high frequencies (>8 kHz) that are seldom clinically measured but may be a marker for accumulated lifetime noise damage ((17, 30–32), **Fig. 1B**, **Table 2**). Outer hair cell function, assessed using distortion product otoacoustic emissions (DPOAEs), were comparable between young adult and middle-aged listeners up to 4 kHz, the frequency regions that contains most of the spectral information in speech (**Fig. 1C**, **Table 3**). Participants also had no severe symptoms of tinnitus (**Fig. 1D**) assessed using the Tinnitus Handicap Inventory (THI; (33)) and Loudness Discomfort Levels (LDLs; (32)) above 80 dB SPL for frequencies up to 3 kHz (**Fig. 1E**, **Table 4**). Self-reported noise exposure using the Noise

Exposure Questionnaire (NEQ; (35)) was not significantly different between age groups (**Fig. 1F**, **Table 4**). Participants also had normal cognitive function indexed by the Montreal Cognitive Assessment (MoCA  $\geq 25$ ; (36)) and comparable working memory scores assessed using the operation span task (OSPAN) ((37), **Fig. 1G**, **Table 4**). Hence, the middle-aged adults recruited for this study were all “normal” by currently administered behavioral and audiological assessments in the hearing clinic, while exhibiting some sub-clinical outer hair cell dysfunction, especially at frequencies above 4 kHz.

We then measured putative CND using neural ensemble responses from the auditory periphery phase-locked to the stimulus amplitude envelope (Envelope following response, EFR). EFRs can be used to emphasize neural generators in the auditory periphery by exploiting divergent phase-locking abilities along the ascending auditory pathway. EFRs, especially at rapid amplitude modulation (AM) frequencies above 600 Hz, have been shown to relate to underlying CND in animal models (12, 14) and in humans (25). Here, we measured EFRs to AM frequencies that have putative neural generators in the central auditory pathway such as the cortex (40 Hz AM; (12, 38)), as well as faster modulation rates (110 Hz, 512 Hz, and 1024 Hz AM) that emphasize progressively peripheral auditory regions (12). We were able to reliably record EFRs, even to modulation frequencies up to 1024 Hz, by using gold-foil tipped electrodes (“tiptrodes”) placed in the ear canal, closer to the presumptive neural generators in the auditory nerve (**Fig. 1H**). EFR peaks analyzed in the spectral domain were above noise floor, with average signal to noise ratios (SNRs) of 8 dB in YA and 4 dB in MA (**Fig. 1I, J**). Statistically significant decreases in EFR amplitudes were only present for EFRs to the 1024 Hz AM rate, with putative generators in the auditory nerve (12, 14) and were not present for slower AM rates with generators in the midbrain or cortex (**Fig. 1K**, **Table 5**).

To confirm that the EFR parameters used here were indeed sensitive to CND, we measured EFRs using identical stimuli, acquisition, and analysis parameters in young (18 wk) and middle-aged (80 wk) Mongolian gerbils (**Fig. 2A**). Gerbils share the same hearing frequency range as humans, making them an ideal animal model for direct comparison in cross-species studies. Middle-aged gerbils showed no loss of hearing thresholds, similar to middle-aged humans (**Fig. 2B**). Remarkably, gerbils also exhibited a selective decrease in EFR amplitudes for AM rates at 1024 Hz, similar to middle-aged adults (**Fig. 2C**, **Table 6**). CND was assessed using immunohistological analysis of cochlear whole mounts, where the cell bodies, presynaptic ribbon terminals and the post-synaptic glutamate receptor patches were immunostained, visualized using confocal

Figure 1.

**Age-related CND occurs prior to overt changes in hearing thresholds and can be assessed non-invasively by measuring phase-locked neural envelope following responses.**

(A) Thirty middle-aged (MA, 40-55 yrs, mean = 46.1±4.6 yrs) and 36 young adults (YA, 18-25 years, mean = 21.17± 1.8yrs) participated in this study. (B) All participants had clinically normal hearing thresholds with some evidence of threshold losses at extended high frequencies above 8 kHz typically not tested in the clinic. Hearing thresholds in dB HL are shown on the Y axis and frequency in kHz is plotted on the X axis. (C) Outer hair cell function assessed using DPOAEs is comparable between YA and MA up to 4kHz and showed age-related decreases at higher frequencies. Both cohorts show no evidence of self-reported tinnitus (D) or hyperacusis measured as LDLs (E), have comparable self-reported noise exposure levels (F), and comparable working memory scores assessed using OSPAN (G). (H) EFRs to modulation frequencies of 1024Hz can be reliably recorded in young and middle-aged adults using 'tiptrodes'. The panel shows grand-averaged FFT traces for YA and MA. (I) Middle-aged adults showed significant declines in EFR amplitudes at 1024Hz AM, with putative neural generators in the auditory nerve. (J) Signal-to-noise ratios were 8dB on average for YA and 4dB for MA. (K) Statistically significant decreases in EFR amplitudes were selective for 1024Hz AM, the modulation frequency with putative generators in the auditory nerve. All panels: Error bars and shading represent standard error of the mean (SEM). Asterisks represent  $p < 0.05$ , ANOVA.

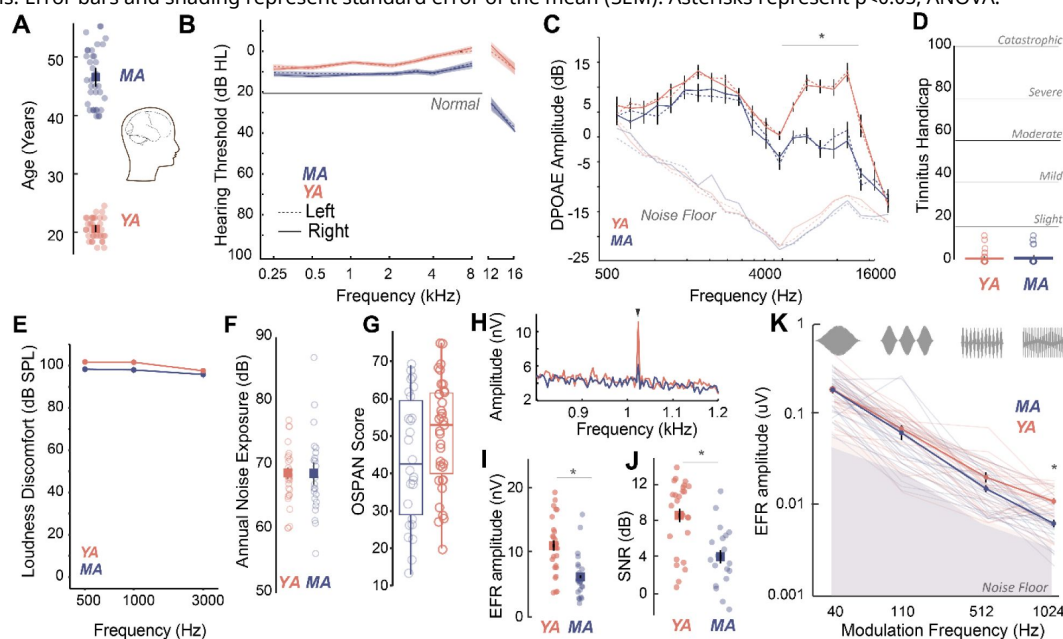


Table 1.

**Comparison of air conduction thresholds using a 3-way ANOVA (MA = 37, YA = 35)**

| Effects        | Dfn | Sum Sq | Mean Sq | F-value | p-value   |
|----------------|-----|--------|---------|---------|-----------|
| Frequency      | 6   | 5067   | 844     | 20.786  | <.001 *** |
| Ear            | 1   | 3      | 3       | 0.068   | .8        |
| Group          | 1   | 5831   | 5813    | 143.521 | <.001 *** |
| Freq:Ear       | 6   | 85     | 14      | 0.349   | .9        |
| Freq:Group     | 6   | 905    | 151     | 3.712   | 0         |
| Ear:Group      | 1   | 7      | 7       | 0.164   | .7        |
| Freq:Ear:Group | 6   | 234    | 39      | 0.961   | .5        |
| Residuals      | 840 | 34125  | 41      |         |           |

\* $p < .05$ ; \*\*\* $p < .001$

Table 2.

Comparison of extended high frequencies using 3-way ANOVA (MA = 37, YA = 35)

| Effects        | DFn | Sum Sq | Mean Sq | F-value | p-value   |
|----------------|-----|--------|---------|---------|-----------|
| Frequency      | 2   | 21209  | 10605   | 74.523  | <.001 *** |
| Ear            | 1   | 6      | 6       | 0.039   | .8        |
| Group          | 1   | 32868  | 32868   | 230.978 | <.001 *** |
| Freq:Ear       | 2   | 142    | 142     | 0.498   | .6        |
| Freq:Group     | 2   | 6016   | 6016    | 21.137  | <.001 *** |
| Ear:Group      | 1   | 152    | 152     | 1.069   | .3        |
| Freq:Ear:Group | 2   | 38     | 19      | 0.134   | .9        |
| Residuals      | 350 | 49805  | 142     |         |           |

\* $p < .05$ ; \*\*\* $p < .001$

Table 3.

Comparison of right ear distortion product otoacoustic emissions using a 2-way ANOVA (MA = 34, YA = 31)

| Effects      | DFn  | DFd    | F-value | p-value    |
|--------------|------|--------|---------|------------|
| Group        | 1    | 63     | 25.85   | < .001 *** |
| Freq         | 9.55 | 601.56 | 58.786  | < .001 *** |
| Group:Freq   | 9.55 | 601.56 | 7.341   | < .001 *** |
| 501:Group    | 1    | 63     | 0.713   | 1.00       |
| 595:Group    | 1    | 63     | 1.939   | 1.00       |
| 707:Group    | 1    | 63     | 0.718   | 1.00       |
| 841:Group    | 1    | 63     | 0.268   | 1.00       |
| 998:Group    | 1    | 63     | 4.38    | .84        |
| 1188:Group   | 1    | 63     | 0.794   | 1.00       |
| 1414:Group   | 1    | 63     | 4.67    | 0.74       |
| 1681:Group   | 1    | 63     | 1.724   | 1.00       |
| 2000:Group   | 1    | 63     | 0.87    | 1.00       |
| 2378:Group   | 1    | 63     | 0.059   | 1.00       |
| 2828:Group   | 1    | 63     | 4.755   | .69        |
| 3365:Group   | 1    | 63     | 2.095   | 1.00       |
| 4001:Group   | 1    | 63     | 10.463  | .04 *      |
| 4757:Group   | 1    | 63     | 18.015  | < .001 *** |
| 5658:Group   | 1    | 63     | 29.947  | < .001 *** |
| 6727:Group   | 1    | 63     | 37.01   | < .001 *** |
| 8000:Group   | 1    | 63     | 28.94   | < .001 *** |
| 9514:Group   | 1    | 63     | 39.235  | < .001 *** |
| 11314:Group  | 1    | 63     | 26.847  | < .001 *** |
| 13454:Group  | 1    | 63     | 7.771   | .147       |
| 160000:Group | 1    | 63     | 0.436   | 1.00       |

\* $p < .05$ ; \*\*\* $p < .001$

**Table 4.**

**Comparisons using 1-way ANOVAs**

| Measure                 | YA (n) | MA (n) | DFn | DFd | F-Value | <i>p</i> |
|-------------------------|--------|--------|-----|-----|---------|----------|
| THI                     | 33     | 37     | 1   | 68  | 0.834   | .364     |
| OSPAN                   | 34     | 34     | 1   | 66  | 3.501   | .066     |
| QuickSIN Clinical Score | 31     | 34     | 1   | 63  | 3.214   | .078     |
| NEQ                     | 32     | 32     | 1   | 66  | 0.8375  | .363     |

Adjusted *p*-values are reported using Bonferroni correction.

**Table 5.**

**Comparison of EFRs using 2-way ANOVAs (MA = 29, YA = 28)**

| Effects    | DFn  | DFd   | F-value | <i>p</i> -value |
|------------|------|-------|---------|-----------------|
| Group      | 1    | 54    | 0.312   | .579            |
| AM         | 1.47 | 79.49 | 164.811 | < .001 ***      |
| Group:AM   | 1.47 | 79.49 | 0.06    | .892            |
| 1024:Group | 1    | 55    | 24.1    | < .001 ***      |
| 512:Group  | 1    | 56    | 0.343   | .561            |
| 110:Group  | 1    | 55    | 0.027   | .87             |
| 40:Group   | 1    | 54    | 3.17    | .081            |

\**p* < .05; \*\*\**p* < .001

microscopy, and quantified from 3D reconstructed images (**Fig. 2D**). Significant decreases in afferent synapse counts were present in middle-aged gerbils, reaching up to 20% losses compared to the young gerbils (**Fig. 2E**, **Table 7**). Further, EFR amplitudes were significantly correlated to the number of remaining cochlear synapses (**Fig. 2F**), thus confirming that our EFRs were a sensitive metric of CND.

## Perceptual deficits manifest as increased listening effort prior to behavioral deficits in middle-aged adults

Do middle-aged adults with putative CND experience challenges with hearing in noise despite having clinically normal hearing thresholds? We measured speech perception in noise using the Quick Speech-in- Noise (QuickSIN; (39)) task, to assess hearing in noise changes that were closer to real-world listening scenarios. QuickSIN tests suprathreshold hearing of medium context sentences presented in varying levels of four-talker background babble (**Fig. 3A**). Further, QuickSIN is a clinically relevant test that we recently identified as being sensitive to detect perceptual deficits in adult populations with normal audiograms (5). Participants are scored on the ability to identify and repeat five key words in each target sentence as the SNR is decreased in 5 dB steps from 25 dB SNR to 0 dB SNR. Clinically, QuickSIN is scored as dB SNR loss, i.e., an estimate of the SNR required to correctly identify key words in noise correctly half the time. No significant age-related decreases were observed in clinically scored QuickSIN measures (**Fig. 3B**, **Table 4**). When analyzing performance at each SNR, accuracy was at near-ceiling from 25 dB SNR to 10 dB SNR, but dropped from 5 dB SNR in both young and middle-aged adults. Statistically significant behavioral deficits with age were observed on QuickSIN only in the most challenging SNR of 0 dB (**Fig. 3C**, **Table 8**).

Are there perceptual deficits experienced by middle-aged adults that are not captured by traditional behavioral readouts? We addressed this question by measuring isoluminous task-related changes in pupil diameter as an index of listening effort (40–42) while participants performed the QuickSIN task (**Fig. 3A**). Pupillary changes were analyzed using growth curve analysis (GCA, (43)). GCAs provide a statistical approach to modeling changes over time in the timing and shape of the pupillary response and has several advantages to analyzing pupillary response over traditional approaches. First, GCA does not require time-binned samples, thus removing the trade-off between temporal resolution and statistical power, and secondly, GCA can account of individual variability. Two second-order GCAs were fit to different time-windows (**Table 9**–10, see methods). One time window from the onset of the masker and covering the first 2.8s of the target sentence (listening window), and second, from the end of the target sentence prior to behavioral response (integration window). These two time-windows are hypothesized to represent effort associated with differing sensory and cognitive processes. The listening window reflects linguistic and semantic processing of ongoing speech stimuli and is a physiological response to auditory processing (44). The integration window reflects error correction, working memory and comparisons with predictive internal models (45). (46). The linear term from the GCA was further analyzed as a marker for the slope of pupillary change over time.

Pupil-indexed listening effort measured during listening was modulated by task difficulty, with pupil diameters showing a larger growth at challenging SNRs (**Fig. 3D**). Both YA and MA showed increases in pupil-indexed effort prior to overt changes in behavioral performance (**Fig. 3E**). While MAs exhibited larger increases in listening effort compared to YAs, this change was not statistically significant (**Fig. 3E**, Supp. **Table 9**). Trends seen in the pupillary responses for the listening window were further amplified in the integration window. Pupillary responses were modulated by task difficulty (**Fig. 3F**). Pupillary slopes obtained from the GCA increased with task difficulty in both YA and MA. However, MA showed a steady increase in listening effort with decreasing SNRs that was higher than YA, reaching a statistically significant increase at 10dB SNR,



Figure 2.

**Cross-species experiments in a rodent model show that EFRs are a sensitive biomarker for histologically confirmed CND.**

(A) Cross-species comparisons were made with young ( $22 \pm 0.86$  weeks,  $n = 14$ ) and middle-aged ( $80 \pm 0.76$  weeks,  $n = 13$ ) Mongolian gerbils, with identical stimuli, recording, and analysis parameters. (B) Middle-aged gerbils did not show any age-related decreases in hearing thresholds. (C) Age-related decreases in EFR amplitudes were isolated to the 1024Hz modulation frequency, similar to middle-aged humans in Fig1K. (D) CND was quantified for a subset of these gerbils ( $n = 10$  young and 10 middle-aged) using immunostained organ of Corti whole mounts, where afferent excitatory synapses were quantified using 3D reconstructed images. (E) Cochlear synapse counts at the 3kHz cochlear region corresponding to the carrier frequency for the EFRs was significantly decreased in middle-aged gerbils, despite matched auditory thresholds. (F) EFR amplitudes at 1024Hz AM were significantly correlated with the number of remaining cochlear synapses, suggesting that these EFRs are a sensitive metric for CND with age. All panels: Error bars and shading represent standard error of the mean (SEM). Asterisks represent  $p < 0.05$ , ANOVA.

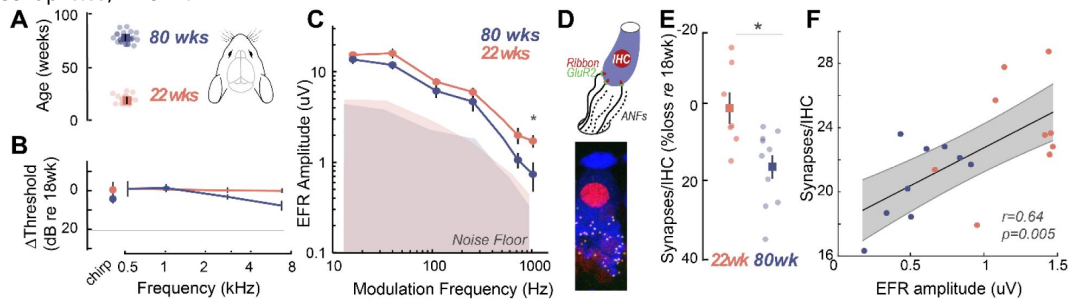


Table 6.

**Comparison of 22 week-old gerbil ( $n = 14$ ) and 80 week-old gerbil ( $n = 12$ ) EFRs using 2-way ANOVAs**

| Effects    | DFn  | DFd   | F-value | p-value    |
|------------|------|-------|---------|------------|
| Group      | 1    | 24    | 4.125   | .053       |
| AM         | 2.68 | 64.28 | 74.636  | < .001 *** |
| Group:AM   | 2.68 | 64.28 | 0.875   | .449       |
| 16:Group   | 1    | 24    | 0.456   | .506       |
| 40:Group   | 1    | 24    | 2.461   | .130       |
| 110:Group  | 1    | 24    | 3.056   | .093       |
| 256:Group  | 1    | 24    | 1.959   | .174       |
| 724:Group  | 1    | 24    | 2.483   | .128       |
| 1024:Group | 1    | 24    | 5.158   | .032 *     |

\* $p < .05$ ; \*\*\* $p < .001$

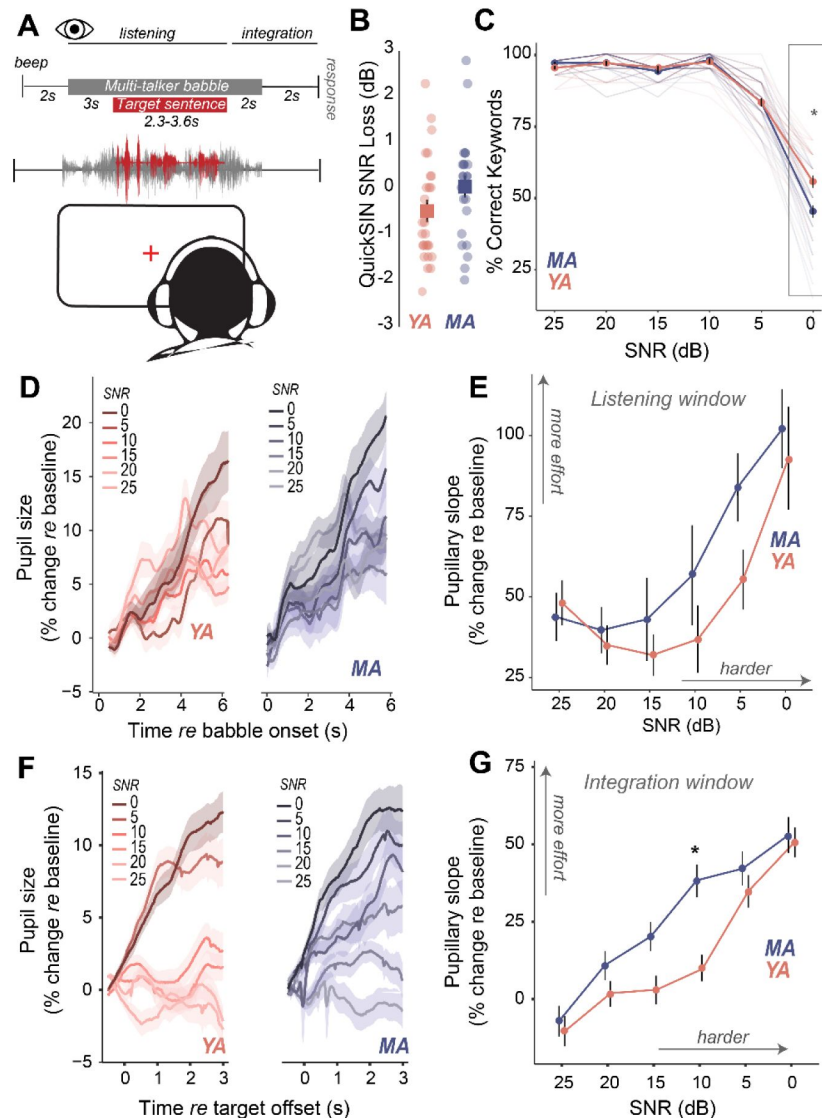
Table 7.

**Comparison of synapse counts at 3000 Hz in 19 and 74 week-old gerbils using 1-way ANOVA**

| Measure        | 19wk (n) | 74wk (n) | DFn | DFd | F-Value | p      |
|----------------|----------|----------|-----|-----|---------|--------|
| Synapse Counts | 14       | 12       | 1   | 16  | 4.877   | .042 * |

\* $p < .05$ ; \*\*\* $p < .001$





**Figure 3.**

### Increased listening effort precedes behavioral deficits in speech in noise perception in middle-aged adults.

(A) Speech perception in noise was assessed using the QuickSIN test, which presents moderate context sentences in varying levels of multi-talker babble. Pupillary measures were analyzed in two time-windows – 1. during stimulus presentation, and 2. after target sentence offset and prior to response initiation (B) No significant age-related differences were observed in clinical QuickSIN scores presented as dB SNR loss. (C) QuickSIN performance is matched between MA and YA until the most difficult noise condition (SNR 0). The x-axis shows the SNR condition that the target sentences were presented in, with 25dB being the easiest noise condition, and 0dB being the most difficult noise condition. The y-axis shows participant accuracy in repeating key words from the target sentences as percent correct. (D) Grand-averaged pupillary responses measured during task listening as an index of effort exhibit modulation with task difficulty, with greater pupillary dilations observed in harder conditions for both groups. (E) Middle-aged adults show consistently higher pupillary responses during performance on the QuickSIN task and at SNR levels prior to when overt behavioral deficits are observed. (F) Grand-averaged pupillary responses measured after target sentence offset as an index of effort exhibit greater modulation with task difficulty, compared to changes in the listening window. (G) Trends seen in the listening window were amplified in this integration window, with middle-aged adults showing even greater effort, especially at moderate SNRs where behavior was matched.

| Effects   | DFn  | DFd    | F-value | <i>p</i> -value |
|-----------|------|--------|---------|-----------------|
| Group     | 1    | 62     | 2.351   | .13             |
| SNR       | 2.46 | 152.32 | 570.184 | < .001 ***      |
| Group:SNR | 2.46 | 152.32 | 7.454   | < .001 ***      |
| 0:Group   | 1    | 62     | 10.5    | .002 *          |
| 5:Group   | 1    | 62     | 0.003   | .957            |
| 10:Group  | 1    | 62     | 1.34    | .252            |
| 15:Group  | 1    | 62     | 1.84    | .181            |
| 20:Group  | 1    | 62     | 0.228   | .635            |
| 25:Group  | 1    | 62     | 6.09    | .016 *          |

$p < .05$ ; \*\*\* $p < .001$

**Table 8.**

**Comparison of QuickSIN performance using a 2-way ANOVA (MA = 34, YA = 31)**

| Fixed Effect         | Estimate  | SE    | 95% CI           | t         | p      |
|----------------------|-----------|-------|------------------|-----------|--------|
| Intercept            | 7.26      | 1.26  | [ 4.79, 9.73]    | 5.76      | < .001 |
| ot1                  | 48.14     | 10.62 | [ 27.33, 68.94]  | 4.53      | < .001 |
| ot2                  | -30.44    | 6.90  | [-43.96, -16.92] | -4.41     | < .001 |
| SNR 0                | -0.43     | 1.80  | [ -3.95, 3.09]   | -0.24     | 0.811  |
| SNR 5                | -3.45     | 1.80  | [ -6.98, 0.07]   | -1.92     | 0.055  |
| SNR 10               | -3.53     | 1.80  | [ -7.06, -0.01]  | -1.96     | 0.049  |
| SNR 15               | -3.72     | 1.77  | [ -7.19, -0.25]  | -2.10     | 0.035  |
| SNR 20               | -2.32     | 1.80  | [ -5.85, 1.20]   | -1.29     | 0.197  |
| ot1 x SNR 0          | 44.77     | 14.58 | [ 16.19, 73.34]  | 3.07      | 0.002  |
| ot1 x SNR 5          | 7.21      | 14.59 | [-21.39, 35.81]  | 0.49      | 0.621  |
| ot1 x SNR 10         | -11.30    | 14.59 | [-39.90, 17.30]  | -0.77     | 0.439  |
| ot1 x SNR 15         | -16.19    | 14.34 | [-44.29, 11.92]  | -1.13     | 0.259  |
| ot1 x SNR 20         | -13.08    | 14.58 | [-41.65, 15.50]  | -0.90     | 0.370  |
| ot2 x SNR 0          | 44.88     | 9.26  | [ 26.73, 63.02]  | 4.85      | < .001 |
| ot2 x SNR 5          | 61.77     | 9.27  | [ 43.60, 79.94]  | 6.66      | < .001 |
| ot2 x SNR 10         | 26.06     | 9.27  | [ 7.88, 44.23]   | 2.81      | 0.005  |
| ot2 x SNR 15         | 14.74     | 9.11  | [ -3.11, 32.59]  | 1.62      | 0.105  |
| ot2 x SNR 20         | 20.07     | 9.26  | [ 1.92, 38.22]   | 2.17      | 0.030  |
| Group (MA vs. YA)    | 1.03      | 1.85  | [ -2.60, 4.65]   | 0.55      | 0.579  |
| Group x SNR 0        | 1.53      | 2.64  | [ -3.65, 6.70]   | 0.58      | 0.564  |
| Group x SNR 5        | 1.87      | 2.63  | [ -3.28, 7.02]   | 0.71      | 0.476  |
| Group x SNR 10       | -4.82e-03 | 2.60  | [ -5.10, 5.10]   | -1.85e-03 | 0.999  |
| Group x SNR 15       | -1.26     | 2.62  | [ -6.39, 3.88]   | -0.48     | 0.632  |
| Group x SNR 20       | -1.26     | 2.64  | [ -6.44, 3.91]   | -0.48     | 0.632  |
| ot1 x Group          | -4.37     | 15.56 | [-34.87, 26.14]  | -0.28     | 0.779  |
| ot1 x Group x SNR 0  | 13.71     | 21.46 | [-28.35, 55.77]  | 0.64      | 0.523  |
| ot1 x Group x SNR 5  | 32.89     | 21.35 | [ -8.94, 74.73]  | 1.54      | 0.123  |
| ot1 x Group x SNR 10 | 24.20     | 21.16 | [-17.26, 65.66]  | 1.14      | 0.253  |
| ot1 x Group x SNR 15 | 15.46     | 21.28 | [-26.25, 57.18]  | 0.73      | 0.468  |
| ot1 x Group x SNR 20 | 8.97      | 21.45 | [-33.07, 51.00]  | 0.42      | 0.676  |
| ot2 x Group          | -4.89     | 10.11 | [-24.70, 14.92]  | -0.48     | 0.628  |
| ot2 x Group x SNR 0  | 3.18      | 13.65 | [-23.57, 29.93]  | 0.23      | 0.816  |
| ot2 x Group x SNR 5  | -11.65    | 13.57 | [-38.26, 14.96]  | -0.86     | 0.391  |
| ot2 x Group x SNR 10 | 16.76     | 13.46 | [ -9.62, 43.13]  | 1.24      | 0.213  |
| ot2 x Group x SNR 15 | 14.82     | 13.53 | [-11.70, 41.35]  | 1.10      | 0.273  |
| ot2 x Group x SNR 20 | 12.19     | 13.63 | [-14.54, 38.91]  | 0.89      | 0.371  |

Growth curve formula:  $\text{Imer}(\text{Pupil} \sim (\text{ot1} + \text{ot2}) * \text{Group} * \text{SNR} + (0 + \text{ot1} + \text{ot2} \mid \text{participant}) + (\text{ot1} + \text{ot2} \mid \text{participant} : \text{SNR}), \text{control} = \text{ImerControl}(\text{optimizer} = 'bobyqa'), \text{REML} = \text{FALSE})$ . Orthogonal polynomial terms: ot1 = linear (slope); ot2 = quadratic (curvature).

**Table 9.**

**Fixed-effect estimates for model of pupillary responses from 0 to 5.8 seconds time-locked to babble masker onset to examine the effect of SNR and age group (observations = 96,612, groups: participant x SNR = 332, participant = 63)**

even though behavioral performance was matched ([Figure 3G](#), [Supp. Table 10](#)). These results suggest that middle-aged adults may maintain comparable performance to younger listeners at moderate task difficulty but at the cost of greater listening effort.

## Pupil-indexed listening effort and CND provide synergistic contributions to speech in noise intelligibility

We sought to understand the relationships between CND, listening effort and speech-in-noise intelligibility in normal-hearing middle-aged adults. Behavioral performance in QuickSIN at 0dB SNR, where there was a group effect of age, was significantly correlated with CND assessed using EFRs at 1024 Hz ([Fig. 4A](#)), suggesting that peripheral deafferentation manifests as overt behavioral deficits under the most challenging listening conditions. Pupil-indexed listening effort was also greater in the integration window in middle-aged adults at 10dB SNR ([Fig. 3G](#)), even though behavioral performance was near ceiling in both young and middle-aged adults. Pupillary slopes at 10dB SNR in the integration window were correlated with behavioral deficits at 0 dB SNR ([Fig. 4B](#)). These results add to the emerging evidence suggesting that pupil-indexed effort to maintain behavioral performance at moderate task difficulties is predictive of behavioral performance at more challenging listening conditions ([47](#)). There were significant correlations between pupillary slopes in the listening window as well, even though there were no group level differences with age ([Fig. 4C](#)). These data suggest that CND and increased listening effort both associated with listening challenges in middle-aged adults.

Is increase in listening effort synergistic with for CND? To understand the multifactorial contributions of sensory and top-down factors that may affect speech perception in noise, we performed a penalized regression with elastic net penalty ([48](#)), with QuickSIN performance at 0dB SNR (scaled to 0-100) as the outcome variable and all other measured variables as the input variables. The elastic net penalized regression framework is a robust method that blends of Lasso's ability to perform variable selection and Ridge's ability to handle multicollinearity and grouped covariates. The fitted elastic net regression model shows an  $R^2$  value of 0.5981, and five significant predictors – hearing thresholds averaged across 500Hz to 4kHz (PTA4k), EFR amplitudes at 1024Hz AM, pupillary slopes at 10dB SNR and 0 dB SNR in the listening window, and pupillary slopes at 10dB SNR in the integration window ([Fig. 4D-E](#)). This model was significantly related to QuickSIN performance and predicted the observed QuickSIN scores across YA and MA ( $r = 0.64$ /(pseudo- $R^2 = 0.41$ , [Fig. 4F](#)). Hence, the output of the elastic net regression suggests that CND and pupil-indexed listening, in addition to subclinical changes in hearing thresholds all provided complementary contributions to speech perception in noise.

## Discussion

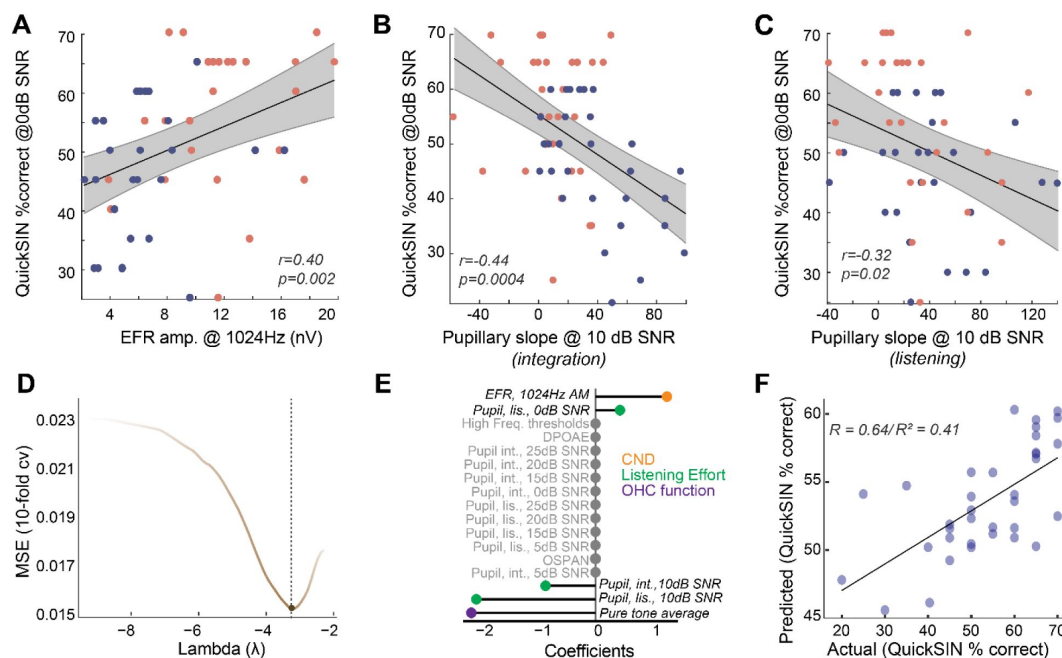
Middle-age, typically defined as the fifth and sixth decade of life, has been historically understudied compared to older age ranges ([49](#)). Yet increasing evidence suggests that middle-age is critical as a period of rapid changes in brain function ([50](#), [51](#)). The resilience of the brain in keeping with degenerative processes that begin to occur in middle-age predicts further age-related degeneration in older ages and presents a critical opportunity for early intervention ([49](#), [52](#)–[54](#)). Hearing loss has been recently identified as the single most modifiable risk factor in middle-age associated with dementia and Alzheimer's disease later in life ([2](#)). However, the number of middle-aged patients who seek help for hearing difficulties but have no abnormal clinical indicators suggests the need for the development of sensitive biomarkers for hearing challenges experienced by this population ([3](#), [5](#), [6](#), [55](#)).

| Fixed Effect         | Estimate | SE    | 95% CI          | t     | p      |
|----------------------|----------|-------|-----------------|-------|--------|
| Intercept            | -0.36    | 0.81  | [-1.95, 1.22]   | -0.45 | .652   |
| ot1                  | -10.33   | 6.06  | [-22.20, 1.54]  | -1.71 | .088   |
| ot2                  | -2.24    | 3.12  | [-8.35, 3.88]   | -0.72 | .474   |
| SNR 0                | 7.40     | 1.00  | [5.45, 9.36]    | 7.43  | < .001 |
| SNR 5                | 6.93     | 1.00  | [4.97, 8.88]    | 6.95  | < .001 |
| SNR 10               | 1.86     | 1.00  | [-0.09, 3.82]   | 1.87  | 0.062  |
| SNR 15               | 0.84     | 1.01  | [-1.13, 2.81]   | 0.83  | .404   |
| SNR 20               | -0.55    | 1.00  | [-2.50, 1.41]   | -0.55 | .583   |
| ot1 x SNR 0          | 60.92    | 7.15  | [46.91, 74.92]  | 8.52  | < .001 |
| ot1 x SNR 5          | 45.16    | 7.15  | [31.15, 59.16]  | 6.32  | < .001 |
| ot1 x SNR 10         | 20.10    | 7.15  | [6.10, 34.11]   | 2.81  | .005   |
| ot1 x SNR 15         | 13.38    | 7.21  | [-0.76, 27.51]  | 1.85  | .064   |
| ot1 x SNR 20         | 12.27    | 7.15  | [-1.74, 26.28]  | 1.72  | .086   |
| ot2 x SNR 0          | -3.41    | 4.19  | [-11.62, 4.81]  | -0.81 | .416   |
| ot2 x SNR 5          | -14.97   | 4.19  | [-23.19, -6.75] | -3.57 | < .001 |
| ot2 x SNR 10         | 6.43     | 4.19  | [-1.78, 14.65]  | 1.53  | .125   |
| ot2 x SNR 15         | 8.83     | 4.23  | [0.54, 17.12]   | 2.09  | .037   |
| ot2 x SNR 20         | 7.83     | 4.19  | [-0.39, 16.05]  | 1.87  | .062   |
| Group (MA vs. YA)    | -0.30    | 1.16  | [-2.57, 1.97]   | -0.26 | .796   |
| Group x SNR 0        | 1.64     | 1.44  | [-1.18, 4.46]   | 1.14  | .254   |
| Group x SNR 5        | 0.37     | 1.43  | [-2.43, 3.16]   | 0.26  | .796   |
| Group x SNR 10       | 3.16     | 1.43  | [0.36, 5.97]    | 2.21  | .027   |
| Group x SNR 15       | 3.79     | 1.45  | [0.95, 6.63]    | 2.62  | .009   |
| Group x SNR 20       | 2.63     | 1.45  | [-0.22, 5.47]   | 1.81  | .071   |
| ot1 x Group          | 3.28     | 8.67  | [-13.72, 20.27] | 0.38  | .706   |
| ot1 x Group x SNR 0  | -0.89    | 10.33 | [-21.13, 19.36] | -0.09 | .932   |
| ot1 x Group x SNR 5  | 4.05     | 10.23 | [-15.99, 24.10] | 0.40  | .692   |
| ot1 x Group x SNR 10 | 25.33    | 10.26 | [5.21, 45.44]   | 2.47  | .014   |
| ot1 x Group x SNR 15 | 14.01    | 10.40 | [-6.37, 34.39]  | 1.35  | .178   |
| ot1 x Group x SNR 20 | 6.24     | 10.43 | [-14.20, 26.67] | 0.60  | .550   |
| ot2 x Group          | 5.50     | 4.48  | [-3.29, 14.29]  | 1.23  | .220   |
| ot2 x Group x SNR 0  | -11.67   | 6.04  | [-23.51, 0.18]  | -1.93 | .053   |
| ot2 x Group x SNR 5  | 3.62     | 5.99  | [-8.11, 15.36]  | 0.61  | .545   |
| ot2 x Group x SNR 10 | -6.72    | 6.01  | [-18.50, 5.06]  | -1.12 | .264   |
| ot2 x Group x SNR 15 | -18.83   | 6.09  | [-30.77, -6.90] | -3.09 | .002   |
| ot2 x Group x SNR 20 | -17.10   | 6.10  | [-29.06, -5.15] | -2.80 | .005   |

Growth curve formula:  $\text{Imer}(\text{Pupil} \sim (\text{ot1} + \text{ot2}) * \text{Group} * \text{SNR} + (\text{ot1} + \text{ot2} \mid \text{participant}) + (\text{ot1} + \text{ot2} \mid \text{participant} : \text{SNR}), \text{control} = \text{ImerControl}(\text{optimizer} = 'bobyqa'), \text{REML} = \text{FALSE})$ . Orthogonal polynomial terms: ot1 = linear (slope); ot2 = quadratic (curvature)

**Table 10.**

**Fixed-effect estimates for model of pupillary responses from 0 to 3 seconds time-locked to QuickSIN target sentence offset to examine the effect of SNR and age group (observations = 63,184, groups: participant x SNR = 359, participant = 63)**



**Figure 4.**

### Listening effort and CND provide complementary contributions to speech in noise intelligibility.

(A) Behavioral performance at the most challenging SNR was significantly correlated with the EFR measures of CND, with lower EFR amplitudes being associated with poorer behavioral performance. (B) Pupillary responses at 10 dB SNR from the integration window were significantly correlated with behavioral performance at 0dB SNR, (B) These correlations between pupillary responses at 10 dB SNR and behavioral performance at 0dB SNR was also found in the listening window, even though there were no group differences in age, further strengthening the link between listening effort at moderate SNRs and behavioral performance at challenging SNRs. (D) an elastic net regression model with 10-fold cross validation (cv) was fit to the QuickSIN scores at 0dB SNR. The tuning parameter Lambda controls the extent to which coefficients contributing least to predictive accuracy are suppressed. (E) A lollipop plot displaying the coefficients ( $\beta$ ) contributing to explaining variance on QuickSIN performance suggests that CND, listening effort and subclinical changes in hearing thresholds all contribute to QuickSIN performance. (F) QuickSIN scores predicted by the elastic net regression are correlated with actual participant QuickSIN scores.



Anatomical evidence from human temporal bones suggests a 40% deafferentation of cochlear synapses in middle-aged adults, even without a substantial noise exposure history (9–11). Peripheral deafferentation triggers compensatory mechanisms across sensory, language, and attentional systems (56–59). But our understanding of the perceptual consequences of cochlear deafferentation are limited by the lack of consensus on sensitive biomarkers for CND (60). Recent studies have identified multiple promising biomarkers for CND in animal models and human populations (21, 25, 61). Here, we used one such marker to identify CND in middle-aged adults with normal audiometric thresholds. EFRs measure peripheral neural coding and central auditory activity by exploiting the divergent phase-locking abilities of the ascending auditory pathway (62). Here, we found decreases in EFRs at modulation rates that are selective for the auditory periphery, while responses from the central auditory structures do not differ with age (Fig. 1K). These data suggest a decrease in peripheral neural coding, with a concomitant increase in central auditory activity or ‘gain’. The perceptual consequences of this gain are unclear, but emerging evidence suggests selective deficits in speech-in-noise abilities (58, 63).

The Mongolian gerbil provides a robust model for cross-species comparisons with aging humans, with their overlapping hearing frequency ranges and experimentally tractable lifespans. Here, using young and middle-aged gerbils, we showed similar EFR decreases as seen in human listeners (Fig. 2C), which are also associated with confirmed CND (Fig. 2F). The gerbils used in this study also do not have any changes in hearing thresholds (Fig. 2B). Hence, they are unlikely to have known strial degenerations that occur in older gerbils and affect auditory thresholds. The synapse loss patterns and EFR amplitude changes seen here in gerbils are in agreement with earlier studies using alternate rodent models (12, 14, 64), further confirming that age-related cochlear synapse loss is a pervasive mammalian phenomenon that can be captured using EFRs to modulation frequencies at 1000Hz AM.

Strong evidence links CND with altered neural coding of sounds in multiple ascending auditory stations (12, 57, 58). However, the perceptual consequences of CND are still unclear (60). Evidence of overt behavioral deficits are mixed and may depend on the specific type of task used for assessment (17, 23). Here, we used QuickSIN, a clinically relevant test that we recently identified as being sensitive to changes in adult normal hearing populations with perceived hearing deficits (5). However, tests that are further challenging in spectrotemporal complexity, such as the addition of time compression or reverberation, may tease apart these differences even more (17, 25). Behavioral deficits here began to emerge only at the most challenging SNRs (Fig. 3). However, perceptual deficits in terms of listening effort began to appear well before these behavioral deficits.

Listening effort is an umbrella term that may assess multiple forms of executive function such as cognitive resource allocation, working memory, and attention, and can be assessed by measuring isoluminous task-linked changes in pupil diameter (27, 40–42, 65). The mechanisms underlying these pupillary changes are still under study (66, 67) but are hypothesized to involve the Locus Coeruleus – Norepinephrine (LC-NE) system (68, 69). Here, we observed that pupil-indexed listening effort increased in middle-aged adults, even when behavioral performance is matched (Fig. 3E, F). This suggests that middle-aged adults expend more effort to maintain behavioral performance, which may lead to more listening fatigue or disengagement from conversations (26, 70, 71). Potentially confounding factors impacting pupil measurement such as the decrease of pupil dynamic range with aging (72, 73), participant fatigue, or task habituation (44, 65, 74) can vary between individuals for a multitude of reasons (75). Here, the effects of these factors were minimized by applying trial-by-trial baseline corrections prior to analysis to match the magnitude of response between young and middle-aged adults.



Interestingly, pupil-indexed listening effort at a moderate SNR was a better predictor of behavioral performance at a more challenging SNR using two independent methods – a Pearson's correlation and the elastic net regression model (**Fig. 4B-D**). We have also previously demonstrated similar results in a different test group of young adult participants. Perhaps akin to predicting a person's ability to run five miles based on assessing their effort required to run one, these results suggest that the amount of effort required to maintain ceiling performance at moderate SNRs are predictive of behavioral performance at harder task difficulties. Pupillary indices at the harder task conditions may be rolling over into hyperexcitability and thus being a poorer predictor of concomitant behavioral performance.

We used a linear model with an elastic net penalization/regularization to simultaneously estimate the underlying contributions of the various predictor variables measured in our studies, and perform model selection. This approach has been previously validated for model selection using multidimensional data related to hearing pathologies like tinnitus and hyperacusis. Elastic net is a regularized regression method that minimizes the negative log-likelihood with a penalty on the parameters that combines the  $l_1$  (LASSO) and  $l_2$  (Ridge) penalty, i.e. the elastic net penalty on the regression parameters  $\beta$  can be written as  $Pen(\beta) = \lambda(\alpha\|\beta\|_1 + (1 - \alpha)/2\|\beta\|_2^2)$ . The relative strength of selection and shrinkage is controlled by the hyper-parameters  $\lambda$  and  $\alpha$ : a higher  $\lambda$  implies more stringent penalization pushing towards the null model and  $0 \leq \alpha \leq 1$  controls the degree of convexity and hence the amount of sparsity with  $\alpha = 0$  implying a Ridge regression with no variable selection. An elastic net regularization has several advantages over both of LASSO or Ridge as well as a simple linear model. The  $l_1$  part of the elastic net ( $\|\beta\|_1$ ) leads to a sparse model where some of the coefficients are shrunk to exact zeroes, thereby performing an automatic model selection without the combinatorial computational complexities of a best-subset selection approach. Further, the quadratic  $l_2$  part ( $\|\beta\|_2^2$ ) encourages grouped variable selection and removes the limitation of number of selected variables unlike LASSO while stabilizing the selection path. Our elastic net regression model suggests that CND and listening effort provided complementary contributions to explaining variance on the QuickSIN task.

Even though both young and middle-aged adults had clinically normal hearing thresholds, subtle changes within this normal range affected speech-in-noise performance, lending support to studies suggesting that the definition of clinical 'normal' may itself need revision. Future studies will directly test this link between cochlear and peripheral neural deficits and listening effort, and explore further contributions of other top-down mechanisms that may influence listening effort such as selective attention or semantic load.

## Methods

### Humans

#### Participants

##### Recruitment

Young ( $n = 38$ ; 18-25 years old, male = 10) and middle-aged ( $n = 45$ ; 40-55 years old, male = 16) adult participants were recruited from the University of Pittsburgh Pitt + Me research participant registry, the University of Pittsburgh Department of Communication Science and Disorders research participant pool, and the broader community under a protocol approved by the University of Pittsburgh Institutional Review Board (IRB#21040125). Participants were compensated for their time, travel, and given an additional monetary incentive for completing all study sessions.

## Eligibility

Participant eligibility was determined during the first session of the study. Eligible participants had normal cognition determined by the Montreal Cognitive Assessment (MoCA  $\geq 25$ ; Nasreddine et al., 2005), normal hearing thresholds ( $\leq 25$  dB HL 250-8000 Hz), no severe tinnitus self-reported via the Tinnitus Handicap Inventory (THI; (33)), and Loudness Discomfort Levels (LDLs)  $\geq 80$  dB HL at .5, 1, and 3 kHz (34). Participants self-reported American English fluency. Thirty-six young (18-25 years old, male = 10) and 30 middle-aged participants (40-55 years old, male = 10) met these eligibility criteria and were tested further using the battery described below. The Beck's depression Inventory (BDI (80)) was administered and participants were excluded if they reported thoughts of self-harm, determined by any response to survey item nine greater than 0.

## Audiological assessment

### Otoscopy

An otoscopic examination was conducted using a Welch Allyn otoscope to examine the patient's external auditory canal, tympanic membrane, and middle ear space for excess cerumen, ear drainage, and other abnormalities. The presence of any such abnormality resulted in exclusion from the study, as these may lead to a conductive hearing loss.

### Audiogram

Hearing thresholds were collected inside a sound attenuating booth using a MADSEN Astera<sup>2</sup> audiometer, Otometrics transducers [Natus Medical, Inc. Middleton, WI], and foam insert eartips sized to the participants' ear canal width. Tones were presented using a pulsed beat and participants were instructed to press a response plunger if they believed that they perceived a tone being played, even if they were unsure. Extended high frequency hearing thresholds (EHFs) were collected at frequencies 8, 12.5, and 16 kHz using Sennheiser circumaural headphones and Sennheiser HDA 300 transducers using the same response instructions.

### Loudness Discomfort Levels (LDLs)

LDLs were collected binaurally using Otometrics transducer [Natus Medical, Inc., Middleton, WI] and foam tip ear inserts. Warble tones were presented, and participants were instructed to rate the loudness on a scale of one to seven, with seven being so loud that they would leave the room.

### Distortion Product Otoacoustic Emissions (DPOAEs)

Outer hair cell function was assessed using DPOAEs. DPOAEs were collected from both the right and left ear individually with a starting frequency of 500 Hz and an ending frequency of 16 kHz. The stimulus had an L1 of 75 dB SPL and an L2 of 65 dB SPL and was presented in 8 blocks of 24 sweeps with alternating polarity. Responses were collected using rubber ear inserts sized to participants' ear canal width and ER-10D DPOAE Probe transducer [Etymotic Research Inc., Elk Grove, IL].

### Noise Exposure History

Participants completed the Noise Exposure Questionnaire (NEQ; (35)) as a self-reported assay of annual noise exposure accounting for both occupational and non-occupational sources. Annual noise exposure was expressed using  $L_{Aeq8760h}$ , representing the annual hourly duration of noise exposure presented in sound pressure level in dB. Calculation of the  $L_{Aeq8760h}$  followed the original article (35).

## OSPAN

Participants also completed the automated version of the OSPAN task<sup>(81)</sup>, which measures working memory<sup>(37)</sup>. Participants were shown simple arithmetic problems and asked to decide whether presented solutions to the problems were correct or incorrect. A letter was displayed on the screen after each problem. Participants were required to recall the letters that were displayed in the order that they appeared following a series of arithmetic problems. The task consisted of 15 letter sequences that spanned three to seven letters (three repetitions of each span). If a participant correctly recalled all letters from a sequence, the span length was added to their score. The maximum possible score on the OSPAN task was 75. Each participant's OSPAN score was used as a measure of working memory.

## Speech perception in noise

### Sentence-level speech perception in noise

Speech perception in noise was indexed using moderate-predictability sentences masked in multitalker babble at six different signal-to-noise ratios (SNR) from the Quick Speech in Noise test (QuickSIN; Killion et al., 2004). QuickSIN is a standardized measure of speech perception in noise that is commonly used in audiology clinics and is representative of a naturalistic listening environment<sup>(82)</sup>. QuickSIN provides a measure of SNR loss. Each QuickSIN test list consisted of six sentences masked in four-talker babble at the following SNR levels: 25, 20, 15, 10, 5, and 0dB. All participants completed four test lists. Participants were instructed to fixate on a point on the screen during listening (to facilitate pupillometry recordings, described below) and to repeat the target sentence to the best of their ability. Each target sentence contained five keywords for identification. The number of key words identified per sentence were recorded. Then, the proportion of keywords correctly identified for each SNR across all four test lists (20 total key words per SNR) was calculated for each participant. In addition to the clinical scoring protocol, participants' performance as the proportion of correctly identified words (i.e., perception accuracy) was also quantified<sup>(39)</sup>,<sup>(83)</sup>.

## Pupillometry

### Acquisition

Pupillary responses were recorded while participants completed the QuickSIN task. Participants were seated in a testing room with consistent, moderate ambient lighting facing a monitor and an EyeLink 1000 Plus Desktop Mount camera (SR Research). During the pupillometry tasks, participants rested their chins on a head-mount and wore Sennheiser circumaural headphones. The masker was presented at 60dB SPL. The sound level of the target sentences was varied to obtain the required signal to noise ratio. The EyeLink 1000 Plus system recorded monocular left eye pupil size in arbitrary units at a 1000 Hz sampling rate. Nine-point eye-tracker calibration was performed prior to the start of the experiment. Participants were required to fixate on a cross on the screen at the start of each trial for a minimum of 500 ms to trigger the start of the QuickSIN stimulus. This fixation criterion was applied to control for the effects of saccades, which can alter pupil diameter and to minimize pupil foreshortening errors<sup>(84–86)</sup>. After meeting the fixation criteria, a 100 ms 1000 Hz beep was presented to alert the participant to the start of the trial. There was two second delay after the beep before the QuickSIN stimulus was presented. The background masker began three seconds before the target sentence and continued for two seconds after the target sentence. Following the end of the stimulus, there was another two second delay and a 100 ms 1000 Hz beep to signal the start of the verbal response period. Manual drift correction was performed at the end of each trial by the experimenter to ensure high quality tracking of the pupil.

## Preprocessing

Raw pupillary data recorded while participants listened to QuickSIN sentences were processed in R (87) using the *eyelinker* package (88) and custom written scripts. Pupillary responses were analyzed in two windows of interest: 1) listening window, from multi-talker babble onset through 5800 ms, and 2) integration window, from target sentence offset to 1000 ms prior to behavioral response period. Separately for each window of interest, data were first processed to remove noise from blinks and saccades. Any trial with more than fifteen percent of the samples detected as saccades or blinks were removed. For the remaining trials, blinks were linearly interpolated from 60 ms before to 160ms after the detected blinks. Saccades were linearly interpolated from 60 ms before to 60 ms after any detected saccade. The de-blinked data were then down sampled to a 50 Hz sampling rate. Pupillary responses were baseline corrected and normalized on a trial-by-trial basis to account for a downward drift in baseline that can occur across a task and for individual differences in pupil dynamic range (85). Baseline pupil size was defined as the average pupil size in the 1000 ms period prior to the start of the window of interest ( $\frac{\text{pupil} - \text{baseline}}{\text{baseline}} \times 100$ ). The pupillary response was then averaged across all four test lists for each SNR per participant in each window of interest. The outcome reported is percent change in pupil size from baseline.

A growth curve analysis (GCA; Mirman, 2014) was used to obtain a measure of the slope of the pupillary response during listening. GCA uses orthogonal polynomial time terms to model distinct functional forms of the pupillary response over time. A GCA was fit using a second-order orthogonal polynomial to model the interaction of age group with SNR level. This second-order model provides three parameters to explain the pupillary response. The first is the intercept, which refers to the overall change in the pupillary response over the time-window of interest. The second is the linear term (ot1), which represents the slope of the pupillary response over time, or the rate of dilation. The third is the quadratic term (ot2), representing curvature of the pupil response, or the change in rate of the pupillary response over time. GCA were conducted in R (R Core Team, 2022) using the *lme4* package (89) and *p*-values were estimated using the *lmerTest* package (90).

For the listening window, the best-fit GCA model included fixed effects of each time term (ot1, ot2), SNR (reference = 25), Group (reference = YA), and all 2- and 3-way interactions between SNR, Group, and time terms. The random effect structure consisted of a random slope of each time term per participant that removed the correlation between random effects, and a random slope of each time term per the interaction of participant and SNR level.

$$\text{Pupil} \sim (\text{ot1} + \text{ot2}) * \text{SNR} * \text{Group} + (0 + \text{ot1} + \text{ot2} | \text{participant}) + (\text{ot1} + \text{ot2} | \text{participant} : \text{SNR})$$

For the integration window, the best-fit GCA model included fixed effects of each time term (ot1, ot2), SNR (reference = 25), Group (reference = YA), and all 2- and 3-way interactions between SNR, Group, and time terms. The random effect structure consisted of a random slope of each time term per participant, and a random slope of each time term per the interaction of participant and SNR level.

$$\text{Pupil} \sim (\text{ot1} + \text{ot2}) * \text{SNR} * \text{Group} + (\text{ot1} + \text{ot2} | \text{participant}) + (\text{ot1} + \text{ot2} | \text{participant} : \text{SNR})$$

## Electrophysiology

### Envelope Following Responses (EFRs)

EFRs were collected in a sound attenuating booth using a 64-channel EEG system (BioSemi ActiveTwo) with stimuli presented using ER-3C transducers [Etymotic Research Inc., Elk Grove, IL]. Gold-foil tippers were positioned in participants' ear canals to deliver sound stimuli and record

additional channels of evoked potentials from the ear canal. EFRs were recorded to 85dB SPL tones with a carrier frequency of 3000Hz, amplitude modulated (AM) at 40, 110, 512, and 1024Hz. Stimuli were presented in alternating polarity, with 500 repetitions in each polarity. Stimulus duration was 250ms, and each AM token was presented at 3.1 repetitions/second, for a period of 322ms. Stimuli were presented to the right ear. During electrophysiology recordings, participants were sat in a comfortable recliner chair in a low-lit room and watched a silent, subtitled streaming show or movie of their choice and were instructed to avoid falling asleep. Researchers checked in periodically between recordings to ensure that participants were awake. Averaged responses were collected and analyzed further using custom-written scripts in MATLAB.

## Preprocessing

EFRs from the Fz to the ipsilateral (right) tiptrode were further analyzed. EFRs were processed using a fourth-order Butterworth filter with a lowpass filter of 3000Hz. The highpass filter cutoffs used were 5Hz, 80Hz, 200Hz, 300Hz for 40Hz, 110Hz, 512Hz, and 1024Hz AM stimuli, respectively. Fast Fourier transforms (FFTs) were performed on the averaged time domain waveforms for each participant at each AM rate starting 10ms after stimulus onset to exclude ABRs and ending 10ms after stimulus offset using MATLAB v. 2022a (MathWorks Inc., Natick, Massachusetts). The maximum amplitude of the FFT peak at one of three adjacent bins (~3Hz) around the modulation frequency of the AM rate is reported as the EFR amplitude.

## Animals

### Subjects

Fourteen young adult Mongolian gerbils aged 18-27 weeks (male = 9) and thirteen middle-aged Mongolian gerbils aged 75-82 weeks (male = 6) were used in this study. All animals are born and raised in our animal care facility from breeders obtained from Charles River. The acoustic environment within the holding facility has been characterized by noise-level data logging and is periodically monitored. Data logging revealed an average noise level of 56 dB, with transients not exceeding 74 dB during regular housing conditions and transients of 88dB once a week during cage changes. All animal procedures are approved by the Institutional Animal Care and Use Committee of the University of Pittsburgh (Protocol #21046600).

### Experimental Setup

Experiments were performed in a double walled acoustic chamber. Animals were placed on a water circulated warming blanket set to 37 °C with the pump placed outside the recording chamber to eliminate audio and electrical interferences. Gerbils were initially anesthetized with isoflurane gas anesthesia (4%) in an induction chamber. The animals were transferred post induction to a manifold and maintained at 1%–1.5% isoflurane. The electrodes were then positioned, and the animals were then injected with dexmedetomidine (Dexdomitor, 0.3 mg/kg subdermal) and taken off the isoflurane. The usual duration of isoflurane anesthesia during this setup process was approximately 10 min. Recordings were commenced 15 min after cessation of isoflurane, with the time window for the effects of isoflurane to wear off determined empirically as 9 min, based on ABRs waveforms and latencies as well as the response to foot pinch stimuli. Dexmedetomidine is an alpha-adrenergic agonist which acts as a sedative and an analgesic, and which is known to decrease motivation but preserve behavioral as well as neural responses in rodents (91 [↗](#), 92 [↗](#)). This helps to maintain animals in an un-anesthetized state, where they still respond to pain stimuli like a foot pinch but are otherwise compliant to recordings for a period of about 3 h. Subdermal electrodes (Ambu) were placed on the animals' scalps for the recordings. A positive electrode was placed along the vertex. The negative electrode was placed under the

ipsilateral ear, along the mastoid, while the ground electrode was placed in the base of the tail. Impedances from the electrodes were always less than 1 kΩ as tested using the head-stage (RA4LI, Tucker Davis technologies, or TDT).

## Stimulus presentation, acquisition, and analysis

The stimulus was presented to the right ear of the animal using insert earphones (ER3C, Etymotic) similar to humans. Signal presentation and acquisition was done by a custom program for gerbils (LabView). The output from the insert earphones was calibrated using a Bruel Kjaer microphone and was found to be within  $\pm 6$  dB for the frequency range tested. Digitized waveforms were recorded with a multichannel recording and stimulation system (RZ-6, TDT) and analyzed with custom written programs in MATLAB (Mathworks).

Hearing thresholds were obtained using auditory brainstem responses presented to tone stimuli that were 5 ms long, with a 2.5 ms on and off ramp, at 27.1 repetitions per second. ABRs were filtered from 300 Hz to 30,000 Hz, and thresholds were determined as the minimum sound level that produced a response as assessed using visual inspection by two blinded, trained observers.

EFRs were elicited to sinusoidally amplitude modulated (sAM) tones (5 ms rise/fall, 250 ms duration, 3.1 repetitions/s, alternating polarity) at a 3 kHz carrier frequency presented 30 dB above auditory thresholds obtained using ABRs at 3 kHz. The modulation frequency was systematically varied from 16 Hz to 1024 Hz AM. Responses were amplified ( $\times 10,000$ ; TDT Medusa 4z amplifier) and filtered (0.1–3 kHz). Trials in which the response amplitude exceeded 200  $\mu$ V were rejected; 250 artifact-free trials of each polarity were averaged to compute the EFR waveform. Fast Fourier transforms were performed on the averaged time-domain waveforms starting 10 ms after stimulus onset to exclude ABRs and ending at stimulus offset using custom-written programs in MATLAB (MathWorks). The maximum amplitude of the FFT peak at 1 of 3 frequency bins ( $\sim 3$  Hz each) around the modulation frequency gave the peak FFT amplitude. This FFT amplitude at the modulation frequency of the AM frequency is reported as the EFR amplitude. The noise floor was calculated as the average of 5 frequency bins ( $\sim 3$  Hz each) above and below the central three bins. A response was deemed as significantly above noise if the FFT amplitude was at least 6 dB above the noise floor.

## Immunohistology

Animals were transcardially perfused using a 4% paraformaldehyde solution (Sigma-Aldrich, 441244) for approximately five minutes before decapitation and isolation of the right and left cochlea. Following intra-labyrinthine perfusion with 4% paraformaldehyde, cochleas were stored in paraformaldehyde for one hour. Cochleae were decalcified in EDTA (Fisher Scientific, BP120500) for 3 to 5 days, followed by cryoprotection with sucrose (Fisher Scientific, D16500) and flash freezing. All chemicals were of reagent grade. Cochlea were thawed prior to dissection and dissected in PBS solution. Immunostaining was accomplished by incubation with the following primary antibodies: 1) mouse anti-CtBP2 (BD Biosciences) at 1:200, 2) mouse anti-GluA2 (Millipore) at 1:2000, 3) rabbit anti-myosin VIIa (Proteus Biosciences) at 1:200; followed by incubation with secondary antibodies coupled to AlexaFluors in the red, green, and blue channels. Piece lengths were measured and converted to cochlear frequency using established cochlear maps ([93](#)) and custom plugins in ImageJ. Cochlear stacks were obtained at the target frequency (3 kHz) spanning the cuticular plate to the synaptic pole of  $\sim 10$  hair cells (in 0.25  $\mu$ m z-steps). Images were collected in a 1024  $\times$  1024 raster using a high-resolution, oil-immersion objective (x60) and 1.59x digital zoom using a Nikon A1 confocal microscope. Images were denoised in NIS elements and loaded into an image-processing software platform (Imaris; Oxford Instruments), where IHCs were quantified based on their Myosin VIIa-stained cell bodies and CtBP2-stained nuclei. Presynaptic ribbons and postsynaptic glutamate receptor patches were counted using 3D representations of



each confocal z-stack. Juxtaposed ribbons and receptor puncta constitute a synapse, and these synaptic associations were determined using IMARIS workflows that calculated and displayed the x-y projection of the voxel space (12 [DOI](#), 94 [DOI](#)).

## Statistical analysis

### Analysis of Variance (ANOVA)

Normality was first checked visually using Q-Q plots and statistically using Shapiro-Wilks test with  $\alpha = 0.05$ . Homogeneity of variance was assessed using Levene's test. N-way ANOVAs were completed using R 2022.07.1 for each measure to determine statistically significant differences between groups (95 [DOI](#)). The package employed, `aov()`, uses treatment contrasts in which the first baseline level is compared to each of the following levels. The number of factors was determined based on the conditions tested in each measure. Bonferroni corrections were used to control familywise error rate due to multiple comparisons.

### Correlations

Any outliers were detected using Tukey's Fence with a boundary distance of  $k = 1.5$  and removed. Correlations were computed using Pearson's correlations. Degrees of freedom,  $r$ , and  $p$  values are reported.

### Elastic Net Regression

We used an linear model with an elastic net penalization/regularization (48 [DOI](#)) to simultaneously estimate the underlying contributions of the various predictor variables measured in our studies, and perform model selection. The relative strength of selection and shrinkage is controlled by the hyper-parameters  $\lambda$  and  $\alpha$ : a higher  $\lambda$  implies more stringent penalization pushing towards the null model and  $0 \leq \alpha \leq 1$  controls the degree of convexity and hence the amount of sparsity with  $\alpha = 0$  implying a Ridge regression with no variable selection. To choose the tuning parameters  $\lambda$  and  $\alpha$ , we used a 10-fold cross-validation that minimizes the out-of-sample root mean-squared error (RMSE). We used the R packages *glmnet* (96 [DOI](#)) and *caret* (97 [DOI](#)) for training the elastic net regularizer.

## Acknowledgements

This work was supported by the National Institute on Deafness and Other Communication Disorders-National Institutes of Health Grants R21DC018882 to A.P, T32DC011499 to K. Kandler and B. yates (Trainee: M.E.Z) and F31DC020085 to J.R.M., and the PNC-Trees Charitable Trust (PNC to B.C. and A.P.). We thank Dr. Carl Snyderman for collaboration on the PNC-Trees grant, and Megan Hallihan, Kathryn Bergstrom, Sarah Anthony, and Shaina Wasileski for their assistance with participant recruitment and data collection. Chandrasekaran). Thanks also to The Center for Biological Imaging at the University of Pittsburgh Dr. Simon Watkins, Katherine Helfrich and Mike Calderon (1S10RR028478 01, PI Watkins) for collaboration on confocal imaging and analysis.

## Additional information

### Conflict of Interest

The authors declare no competing financial interests.



## Author Contributions

Conceptualization: AP; Methodology: AP, BC, JM, JD; Data collection: MEZ, JK, KY, VC, OF, CM; Data analysis: MEZ, LZ, JRM, KY, VC, OF, CM; Statistical analysis: JRM, MEZ, LZ, JD; Writing: MEZ, JRM; Editing: AP, JD, BC; Supervision, Project administration: AP, BC; Funding acquisition: AP, BC

## References

1. Lin F. R., Thorpe R., Gordon-Salant S., Ferrucci L. (2011) **Hearing Loss Prevalence and Risk Factors Among Older Adults in the United States** *J. Gerontol. Ser. -Biol. Sci. Med. Sci* **66**:582–590
2. Livingston G., et al. (2017) **Dementia prevention, intervention, and care** *The Lancet* **390**:2673–2734
3. Hind S. E., et al. (2011) **Prevalence of clinical referrals having hearing thresholds within normal limits** *Int. J. Audiol* **50**:708–716
4. Tremblay K. L., et al., . 36, (2015) **Self-Reported Hearing Difficulties Among Adults With Normal Audiograms: The Beaver Dam Offspring Study** *Ear Hear* :E290–E299
5. Cancel V. E., McHaney J. R., Milne V., Palmer C., Parthasarathy A. (2023) **A data-driven approach to identify a rapid screener for auditory processing disorder testing referrals in adults** *Sci. Rep* **13**
6. Parthasarathy A., Hancock K. E., Bennett K., DeGruttola V., Polley D. B. (2020) **Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing** *eLife* **9**
7. Kujawa S. G., Liberman M. C. (2009) **Adding Insult to Injury: Cochlear Nerve Degeneration after “Temporary” Noise-Induced Hearing Loss** *J. Neurosci* **29**:14077–14085
8. Schaette R., McAlpine D. (2011) **Tinnitus with a Normal Audiogram: Physiological Evidence for Hidden Hearing Loss and Computational Model** *J. Neurosci* **31**:13452–13457
9. Wu P. Z., et al. (2018) **Primary Neural Degeneration in the Human Cochlea: Evidence for Hidden Hearing Loss in the Aging Ear** *Neuroscience* <https://doi.org/10.1016/j.neuroscience.2018.07.053>
10. Wu P., O’Malley J. T., Liberman M. C. (2023) **Neural Degeneration in Normal-Aging Human Cochleas: Machine-Learning Counts and 3D Mapping in Archival Sections** *J. Assoc. Res. Otolaryngol* <https://doi.org/10.1007/s10162-023-00909-y>
11. Wu P.-Z., O’Malley J. T., de Gruttola V., Liberman M. C. (2020) **Age-Related Hearing Loss Is Dominated by Damage to Inner Ear Sensory Cells** *Not the Cellular Battery That Powers Them. J. Neurosci. Off. J. Soc. Neurosci* **40**:6357–6366
12. Parthasarathy A., Kujawa S. G. (2018) **Synaptopathy in the Aging Cochlea: Characterizing Early-Neural Deficits in Auditory Temporal Envelope Processing** *J. Neurosci. Off. J. Soc. Neurosci* **38**:7108–7119
13. Mepani A. M., et al. (2021) **Envelope following responses predict speech-in-noise performance in normal-hearing listeners** *J. Neurophysiol* **125**:1213–1222

14. Shaheen L. A., Valero M. D., Liberman M. C. (2015) **Towards a Diagnosis of Cochlear Neuropathy with Envelope Following Responses** *J Assoc Res Otolaryngol* <https://doi.org/10.1007/s10162-015-0539-3>
15. Prendergast G., et al. (2017) **Effects of noise exposure on young adults with normal audiograms I: Electrophysiology** *Hear. Res* **344**:68–81
16. Guest H., Munro K. J., Plack C. J. (2017) **Tinnitus with a normal audiogram: Role of high-frequency sensitivity and reanalysis of brainstem-response measures to avoid audiometric over-matching** *Hear. Res* **356**:116–117
17. Grant K. J., et al. (2020) **Electrophysiological markers of cochlear function correlate with hearing-in-noise performance among audiometrically normal subjects** *J. Neurophysiol* **124**:418–431
18. Grant K. J., et al. (2022) **Predicting neural deficits in sensorineural hearing loss from word recognition scores** *Sci. Rep* **12**
19. Valero M. D., Hancock K. E., Maison S. F., Liberman M. C. (2018) **Effects of cochlear synaptopathy on middle-ear muscle reflexes in unanesthetized mice** *Hear. Res* **363**:109–118
20. Mehraei G., et al. (2016) **Auditory Brainstem Response Latency in Noise as a Marker of Cochlear Synaptopathy** *J. Neurosci* **36**:3755–3764
21. Bharadwaj H. M., et al. (2022) **Cross-species experiments reveal widespread cochlear neural damage in normal hearing.** *Commun Biol* **5**:1–10
22. Buran B. N., McMillan G. P., Keshishzadeh S., Verhulst S., Bramhall N. F. (2022) **Predicting synapse counts in living humans by combining computational models with auditory physiology** *J. Acoust. Soc. Am* **151**
23. Prendergast G., et al. (2017) **Effects of noise exposure on young adults with normal audiograms II: Behavioral measures** *Hear. Res* **356**:74–86
24. Mepani A. M., et al. (2019) **Middle Ear Muscle Reflex and Word Recognition in “Normal-Hearing” Adults: Evidence for Cochlear Synaptopathy?** *Ear Hear* <https://doi.org/10.1097/AUD.0000000000000804>
25. Mepani A. M., et al. (2021) **Envelope following responses predict speech-in-noise performance in normal-hearing listeners** *J. Neurophysiol* **125**:1213–1222
26. Pichora-Fuller M. K., et al. (2016) **Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL)** *Ear Hear* **37**:5–27
27. Peelle J. E. (2018) **Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior** *Ear Hear* **39**:204–214
28. Zekveld A. A., Kramer S. E., Festen J. M. (2011) **Cognitive Load During Speech Perception in Noise: The Influence of Age, Hearing Loss, and Cognition on the Pupil Response** *Ear Hear* **32**:498–510
29. **International Classification of Functioning, Disability and Health (ICF)**

30. Škerková M., et al. (2023) **Extended high-frequency audiometry: hearing thresholds in adults** *Eur. Arch. Otorhinolaryngol* **280**:565–572
31. Mishra S. K., Saxena U., Rodrigo H. (2022) **Extended High-frequency Hearing Impairment Despite a Normal Audiogram: Relation to Early Aging, Speech-in-noise Perception, Cochlear Function, and Routine Earphone Use** *Ear Hear* **43**:822–835
32. Lough M., Plack C. J. (2022) **Extended high-frequency audiometry in research and clinical practice** *J. Acoust. Soc. Am* **151**
33. Newman C. W., Jacobson G. P., Spitzer J. B. (1996) **Development of the Tinnitus Handicap Inventory** *Arch. Otolaryngol. Head Neck Surg* **122**:143–148
34. Sherlock L. P., Formby C. (2005) **Estimates of loudness, loudness discomfort, and the auditory dynamic range: normative estimates, comparison of procedures, and test-retest reliability** *J. Am. Acad. Audiol* **16**:85–100
35. Johnson T. A., Cooper S., Stamper G. C., Chertoff M. (2017) **Noise Exposure Questionnaire (NEQ): A Tool for Quantifying Annual Noise Exposure** *J. Am. Acad. Audiol* **28**:14–35
36. Nasreddine Z. S., et al. (2005) **The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment** *J. Am. Geriatr. Soc* :695–699
37. Turner M. L., Engle R. W. (1989) **Is working memory capacity task dependent?** *J. Mem. Lang* **28**:127–154
38. Parthasarathy A., Bartlett E. (2012) **Two-channel recording of auditory-evoked potentials to detect age- related deficits in temporal processing** *Hear. Res* **289**:52–62
39. Killion M. C., Niquette P. A., Gudmundsen G. I., Revit L. J., Banerjee S. (2004) **Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners** *J. Acoust. Soc. Am* **116**:2395–2405
40. Beatty J. (1982) **PHASIC NOT TONIC PUPILLARY RESPONSES VARY WITH AUDITORY VIGILANCE PERFORMANCE** *Psychophysiology* **19**:167–172
41. Winn M. B., Edwards J. R., Litovsky R. Y. (2015) **The Impact of Auditory Spectral Resolution on Listening Effort Revealed by Pupil Dilation** *Ear Hear* **36**:e153–e165
42. Kuchinsky S. E., et al. (2013) **Pupil size varies with word listening and response selection difficulty in older adults with hearing loss** *Psychophysiology* **50**:23–34
43. Mirman D. (2014) **Growth Curve Analysis and Visualization Using R** Chapman and Hall/CRC
44. McHaney J. R., Tessmer R., Roark C. L., Chandrasekaran B. (2021) **Working memory relates to individual differences in speech category learning: Insights from computational modeling and pupillometry** *Brain Lang* **222**
45. Winn M. B. (2023) **Time Scales and Moments of Listening Effort Revealed in Pupillometry** *Semin. Hear* **44**:106–123
46. Kuchinsky S. E., et al. (2014) **Speech-perception training for older adults with hearing loss impacts word recognition and effort** *Psychophysiology* **51**:1046–1057

47. McHaney J. R., Hancock K. E., Polley D. B., Parthasarathy A., . [Preprint] (2023) **Sensory representations and pupil-indexed listening effort provide complementary contributions to multi-talker speech intelligibility**
48. Zou H., Hastie T. (2005) **Regularization and Variable Selection Via the Elastic Net** *J. R. Stat. Soc. Ser. B Stat. Methodol* **67**:301–320
49. Dohm-Hansen S., et al. (2024) **The “middle-aging” brain** *Trends Neurosci* 0
50. Schaum N., et al. (2020) **Ageing hallmarks exhibit organ-specific temporal signatures** *Nature* **583**:596–602
51. Salthouse T. A. (2019) **Trajectories of normal cognitive aging** *Psychol. Aging* **34**:17–24
52. Elliott M. L., et al. (2021) **Brain-age in midlife is associated with accelerated biological aging and cognitive decline in a longitudinal birth cohort** *Mol. Psychiatry* **26**:3829–3838
53. Elliott M. L., et al. (2021) **Disparities in the pace of biological aging among midlife adults of the same chronological age have implications for future frailty risk and policy.** *Nat Aging* **1**:295–308
54. Hughes M. L., Agrigoroaei S., Jeon M., Bruzzese M., Lachman M. E. (2018) **Change in Cognitive Performance From Midlife Into Old Age: Findings from the Midlife in the United States (MIDUS) Study** *J. Int. Neuropsychol. Soc* **24**:805–820
55. Spehar B. P., Lichtenhan J. T. (2018) **Patients With Normal Hearing Thresholds but Difficulty Hearing in Noisy Environments: A Study on the Willingness to Try Auditory Training** *Otol. Neurotol. Off. Publ. Am. Otol. Soc. Am. Neurotol. Soc. Eur. Acad. Otol. Neurotol* **39**:950–956
56. Auerbach B. D., Radziwon K., Salvi R. (2019) **Testing the Central Gain Model: Loudness Growth Correlates with Central Auditory Gain Enhancement in a Rodent Model of Hyperacusis** *Neuroscience* **407**:93–107
57. Chambers A. R., et al. (2016) **Central Gain Restores Auditory Processing following Near-Complete Cochlear Denervation** *Neuron* **89**:867–879
58. Resnik J., Polley D. B. (2021) **Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise** *Neuron* **109**:984–996
59. Bharadwaj H. M., Masud S., Mehraei G., Verhulst S., Shinn-Cunningham B. G. (2015) **Individual Differences Reveal Correlates of Hidden Hearing Deficits** *J. Neurosci* **35**:2161–2172
60. Bramhall N., et al. (2019) **The search for noise-induced cochlear synaptopathy in humans: Mission impossible?** *Hear. Res* **377**:88–103
61. Bharadwaj M H., et al. (2019) **Non-Invasive Assays of Cochlear Synaptopathy - Candidates and Considerations** *Neuroscience* **407**:53–66
62. Joris P. X., Schreiner C. E., Rees A. (2004) **Neural processing of amplitude-modulated sounds** *Physiol. Rev* **84**:541–577
63. Dougherty K., Hustedt-Mai A., Hagedorn A., Bharadwaj H. (2021) **Central gain in aging, tinnitus, and temporary hearing loss** *J. Acoust. Soc. Am* **150**:A341–A341

64. Parthasarathy A., Bartlett E. L. (2011) **Age-related Auditory Deficits in Temporal Processing in F-344 Rats** *Neuroscience* **192**:619–630
65. Brown V. A., McLaughlin D. J., Strand J. F., Van Engen K. J. (2020) **Rapid adaptation to fully intelligible nonnative-accented speech reduces listening effort** *Q. J. Exp. Psychol* **73**:1431–1443
66. McGinley M. J., David S. V., McCormick D. A. (2015) **Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection** *Neuron* **87**:179–192
67. de Gee J. W., et al. (2020) **Pupil-linked phasic arousal predicts a reduction of choice bias across species and decision domains** *eLife* **9**
68. Joshi S., Li Y., Kalwani R. M., Gold J. I. (2016) **Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex** *Neuron* **89**:221–234
69. Reimer J., et al. (2016) **Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex** *Nat. Commun* **7**
70. Hornsby B. W. Y., Naylor G., Bess F. H. (2016) **A Taxonomy of Fatigue Concepts and Their Relation to Hearing Loss** *Ear Hear* **37**:136–144
71. Hornsby B. W. Y. (2013) **The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands** *Ear Hear* **34**:523–534
72. Telek H. H. (2018) **The Effects of Age Pupil Diameters at Different Light Amplitudes** *Beyoglu Eye J* <https://doi.org/10.14744/bej.2018.43534>
73. Piquado T., Isaacowitz D., Wingfield A. (2010) **Pupillometry as a measure of cognitive effort in younger and older adults** *Psychophysiology* **47**:560–569
74. Tryon W. W. (1975) **Pupillometry: A Survey of Sources of Variation** *Psychophysiology* **12**:90–93
75. Ansari A. S., Vehof J., Hammond C. J., Bremner F. D., Williams K. M. (2021) **Evidence That Pupil Size and Reactivity Are Determined More by Your Parents Than by Your Environment** *Front. Neurol* **12**
76. Smith S. S., Jahn K. N., Sugai J. A., Hancock K. E., Polley D. B., . [Preprint] (2024) **The human pupil and face encode sound affect and provide objective signatures of tinnitus and auditory hypersensitivity disorders**
77. Hunter L. L., et al. (2020) **Extended high frequency hearing and speech perception implications in adults and children** *Hear. Res* **397**
78. Shinn-Cunningham B. G., Best V. (2008) **Selective attention in normal and impaired hearing** *Trends Amplif* **12**:283–299
79. McLaughlin D. J., et al. (2022) **Pupillometry reveals cognitive demands of lexical competition during spoken word recognition in young and older adults** *Psychon. Bull. Rev* **29**:268–280
80. Beck A. T., Steer R. A., Carbin M. G. (1988) **Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation** *Clin. Psychol. Rev* **8**:77–100

81. Unsworth N., Heitz R. P., Schrock J. C., Engle R. W. (2005) **An automated version of the operation span task** *Behav. Res. Methods* **37**:498–505
82. Meyer J., Dentel L., Meunier F. (2013) **Speech recognition in natural background noise** *PLoS One* **8**
83. Wilson R. H., Abrams H. B., Pillion A. L. (2003) **A word-recognition task in multitalker babble using a descending presentation mode from 24 dB to 0 dB signal to babble** *J. Rehabil. Res. Dev* **40**
84. Zekveld A. A., Festen J. M., Kramer S. E. (2013) **Task Difficulty Differentially Affects Two Measures of Processing Load: The Pupil Response During Sentence Processing and Delayed Cued Recall of the Sentences** *J. Speech Lang. Hear. Res* **56**:1156–1165
85. Winn M. B., Wendt D., Koelewijn T., Kuchinsky S. E. (2018) **Best Practices and Advice for Using Pupillometry to Measure Listening Effort: An Introduction for Those Who Want to Get Started** *Trends Hear* **22**
86. Koelewijn T., Zekveld A. A., Lunner T., Kramer S. E. (2018) **The effect of reward on listening effort as reflected by the pupil dilation response** *Hear. Res* **367**:106–112
87. (2022) **R Core Team, R: a language and environment for statistical computing.** (2022). Available at: <https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing> [Accessed 11 December 2022].
88. Barthelme S. **a-hurst/eyelinker: An R package for importing data from EyeLink ASC files** *GitHub*
89. Bates D., Mächler M., Bolker B., Walker S. (2015) **Fitting Linear Mixed-Effects Models Using lme4** *J. Stat. Softw* **67**:1–48
90. Kuznetsova A., Brockhoff P. B., Christensen R. H. B. (2017) **lmerTest Package: Tests in Linear Mixed Effects Models** *J. Stat. Softw* **82**:1–26
91. Ruotsalainen S., Haapalinna A., Riekkinen P. J., Sirviö J. (1997) **Dexmedetomidine Reduces Response Tendency, but Not Accuracy of Rats in Attention and Short-Term Memory Tasks** *Pharmacol. Biochem. Behav* **56**:31–40
92. Ter-Mikaelian M., Semple M. N., Sanes D. H. (2013) **Effects of spectral and temporal disruption on cortical encoding of gerbil vocalizations** *J. Neurophysiol* **110**:1190–1204
93. Greenwood D. D. (1990) **A Cochlear Frequency-Position function for several species -29 Years Later** *J. Acoust. Soc. Am* **87**:2592–2605
94. Liberman L. D., Wang H. B., Liberman M. C. (2011) **Opposing Gradients of Ribbon Size and AMPA Receptor Expression Underlie Sensitivity Differences among Cochlear-Nerve/Hair-Cell Synapses** *J. Neurosci* **31**:801–808
95. Girden E. R. (1992) **ANOVA: Repeated measures** Sage Publications, Inc
96. Friedman J., Hastie T., Tibshirani R. (2010) **Regularization Paths for Generalized Linear Models via Coordinate Descent** *J. Stat. Softw* **33**:1–22



97. Kuhn M. (2008) **Building Predictive Models in R Using the caret Package** *J. Stat. Softw* **28**:1–26

## Author information

### Maggie E Zink\*

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States

\*Equal contributions

### Leslie Zhen\*

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States

\*Equal contributions

### Jacie R McHaney\*\*

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States

ORCID iD: [0000-0002-1148-4018](https://orcid.org/0000-0002-1148-4018)

\*Equal contributions

<sup>†</sup>Present address: Department of Communication Sciences and Disorders, Northwestern University, Evanston, United States

### Jennifer Klara

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States

### Kimberly Yurasits

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States

### Victoria Cancel

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States

### Olivia Flemm

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States

### Claire Mitchell

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States

**Jyotishka Datta**

Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, United States

**Bharath Chandrasekaran<sup>+</sup>**

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States

<sup>+</sup>Present address: Department of Communication Sciences and Disorders, Northwestern University, Evanston, United States

**Aravindakshan Parthasarathy**

Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, United States, Department of Bioengineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, United States, Department of Otolaryngology, School of Medicine, University of Pittsburgh, Pittsburgh, United States

ORCID iD: [0000-0002-4573-8004](https://orcid.org/0000-0002-4573-8004)

**For correspondence:** [Aravind\\_Partha@pitt.edu](mailto:Aravind_Partha@pitt.edu)

**Editors**

Reviewing Editor

**Jonas Obleser**

University of Lübeck, Lübeck, Germany

Senior Editor

**Andrew King**

University of Oxford, Oxford, United Kingdom

**Reviewer #1 (Public review):**

This study is part of an ongoing effort to clarify the effects of cochlear neural degeneration (CND) on auditory processing in listeners with normal audiograms. This effort is important because ~10% of people who seek help for hearing difficulties have normal audiograms and current hearing healthcare has nothing to offer them.

The authors identify two shortcomings in previous work that they intend to fix. The first is a lack of cross-species studies that make direct comparisons between animal models in which CND can be confirmed and humans for which CND must be inferred indirectly. The second is the low sensitivity of purely perceptual measures to subtle changes in auditory processing. To fix these shortcomings, the authors measure envelope following responses (EFRs) in gerbils and humans using the same sounds, while also performing histological analysis of the gerbil cochleae, and testing speech perception while measuring pupil size in the humans.

The study begins with a comprehensive assessment of the hearing status of the human listeners. The only differences found between the young adult (YA) and middle-aged (MA) groups are in thresholds at frequencies > 10 kHz and DPOAE amplitudes at frequencies > 5 kHz. The authors then present the EFR results, first for the humans and then for the gerbils, showing that amplitudes decrease more rapidly with increasing envelope frequency for MA than for YA in both species. The histological analysis of the gerbil cochleae shows that there

were, on average, 20% fewer IHC-AN synapses at the 3 kHz place in MA relative to YA, and the number of synapses per IHC was correlated with the EFR amplitude at 1024 Hz.

The study then returns to the humans to report the results of the speech perception tests and pupillometry. The correct understanding of keywords decreased more rapidly with decreasing SNR in MA than in YA, with a noticeable difference at 0 dB, while pupillary slope (a proxy for listening effort) increased more rapidly with decreasing SNR for MA than for YA, with the largest differences at SNRs between 5 and 15 dB. Finally, the authors report that a linear combination of audiometric threshold, EFR amplitude at 1024 Hz, and a few measures of pupillary slope is predictive of speech perception at 0 dB SNR.

I only have two questions/concerns about the specific methodologies used:

(1) Synapse counts were made only at the 3 kHz place on the cochlea. However, the EFR sounds were presented at 85 dB SPL, which means that a rather large section of the cochlea will actually be excited. Do we know how much of the EFR actually reflects AN fibers coming from the 3 kHz place? And are we sure that this is the same for gerbils and humans given the differences in cochlear geometry, head size, etc.?

(2) Unless I misunderstood, the predictive power of the final model was not tested on held-out data. The standard way to fit and test such a model would be to split the data into two segments, one for training and hyperparameter optimization, and one for testing. But it seems that the only split was for training and hyperparameter optimization.

While I find the study to be generally well executed, I am left wondering what to make of it all. The purpose of the study with respect to fixing previous methodological shortcomings was clear, but exactly how fixing these shortcomings has allowed us to advance is not. I think we can be more confident than before that EFR amplitude is sensitive to CND, and we now know that measures of listening effort may also be sensitive to CND. But where is this leading us?

I think what this line of work is eventually aiming for is to develop a clinical tool that can be used to infer someone's CND profile. That seems like a worthwhile goal but getting there will require going beyond exploratory association studies. I think we're ready to start being explicit about what properties a CND inference tool would need to be practically useful. I have no idea whether the associations reported in this study are encouraging or not because I have no idea what level of inferential power is ultimately required.

That brings me to my final comment: there is an inappropriate emphasis on statistical significance. The sample size was chosen arbitrarily. What if the sample had been half the size? Then few, if any, of the observed effects would have been significant. What if the sample had been twice the size? Then many more of the observed effects would have been significant (particularly for the pupillometry). I hope that future studies will follow a more principled approach in which relevant effect sizes are pre-specified (ideally as the strength of association that would be practically useful) and sample sizes are determined accordingly.

So, in summary, I think this study is a valuable but limited advance. The results increase my confidence that non-invasive measures can be used to infer underlying CND, but I am unsure how much closer we are to anything that is practically useful.

<https://doi.org/10.7554/eLife.102823.1.sa1>

#### **Reviewer #2 (Public review):**

Summary:

This paper addresses the bottom-up and top-down causes of hearing difficulties in middle-aged adults with clinically-normal audiograms using a cross-species approach (humans vs. gerbils, each with two age groups) mixing behavioral tests and electrophysiology. The study is not only a follow-up of Parthasarathy et al (eLife 2020), since there are several important differences.

Parthasarathy et al. (2020) only considered a group of young normal-hearing individuals with normal audiograms yet with high complaints of hearing in noisy situations. Here, this issue is considered specifically regarding aging, using a between-subject design comparing young NH and older NH individuals recruited from the general population, without additional criterion (i.e. no specifically high problems of hearing in noise). In addition, this is a cross-species approach, with the same physiological EFR measurements with the same stimuli deployed on gerbils.

This article is of very high quality. It is extremely clear, and the results show clearly a decrease of neural phase-locking to high modulation frequencies in both middle-aged humans and gerbils, compared to younger groups/cohorts. In addition, pupillometry measurements conducted during the QuickSIN task suggest increased listening efforts in middle-aged participants, and a statistical model including both EFRs and pupillometry features suggests that both factors contribute to reduced speech-in-noise intelligibility evidenced in middle-aged individuals, beyond their slight differences in audiometric thresholds (although they were clinically normal in both groups).

These provide strong support to the view that normal aging in humans leads to auditory nerve synaptic loss (cochlear neural degeneration - CNR- or, put differently, cochlear synaptopathy) as well as increased listening effort, before any clearly visible audiometric deficits as defined in current clinical standards. This result is very important for the community since we are still missing direct evidence that cochlear synaptopathy might likely underlie a significant part of hearing difficulties in complex environments for listeners with normal thresholds, such as middle-aged and senior listeners. This paper shows that these difficulties can be reasonably well accounted for by this sensory disorder (CND), but also that listening effort, i.e. a top-down factor, further contributes to this problem. The methods are sound and well described and I would like to emphasize that they are presented concisely yet in a very precise manner so that they can be understood very easily - even for a reader who is not familiar with the employed techniques. I believe this study will be of interest to a broad readership.

I have some comments and questions which I think would make the paper even stronger once addressed.

Main comments:

(1) Presentation of EFR analyses / Interpretation of EFR differences found in both gerbils and humans:

a) Could the authors comment further on why they think they found a significant difference only at the highest mod. frequency of 1024 Hz in their study? Indeed, previous studies employing SAM or RAM tones very similar to the ones employed here were able to show age effects already at lower modulation freqs. of ~100Hz; e.g. there are clear age effects reported in human studies of Vasilikov et al. (2021) or Mepani et al. (2021), and also in animals (see Garrett et al. bioRxiv: <https://www.biorxiv.org/content/biorxiv/early/2024/04/30/2020.06.09.142950.full.pdf>).

Furthermore, some previous EEG experiments in humans that SAM tones with modulation freqs. of ~100Hz showed that EFRs do not exhibit a single peak, i.e. there are peaks not only at

fm but also for the first harmonics (e.g.  $2 \cdot fm$  or  $3 \cdot fm$ ) see e.g. Garrett et al. bioXiv <https://www.biorxiv.org/content/biorxiv/early/2024/04/30/2020.06.09.142950.full.pdf>.

Did the authors try to extract EFR strength by looking at the summed amplitude of multiple peaks (Vasilikov Hear Res. 2021), in particular for the lower modulation frequencies? (indeed, there will be no harmonics for the higher mod. freqs).

b) How do the present EFR results relate to FFR results, where effects of age are already at low carrier freqs? (e.g. Märcher-Rørsted et al., Hear. Res., 2022 for pure tones with freq < 500 Hz). Do the authors think it could be explained by the fact that this is not the same cochlear region, and that synapses die earlier in higher compared to lower CFs? This should be discussed. Beyond the main group effect of age, there were no negative correlations of EFRs with age in the data?

(2) Size of the effects / comparing age effects between two species:

Although the size of the age effect on EFRs cannot be directly compared between humans and gerbils - the comparison remains qualitative - could the authors at least provide references regarding the rate of synaptic loss with aging in both humans and gerbils, so that we understand that the yNH/MA difference can be compared between the two age groups used for gerbils; it would have been critical in case of a non-significant age effect in one species.

Equalization/control of stimuli differences across the two species: For measuring EFRs, SAM stimuli were presented at 85 dB SPL for humans vs. 30 dB above the detection threshold (inferred from ABRs) for gerbils - I do not think the results strongly depend on this choice, but it would be good to comment on why you did not choose also to present stimuli 30 dB above thresholds in humans.

Simulations of EFRs using functional models could have been used to understand (at least in humans) how the differences in EFRs obtained between the two groups are \*quantitatively\* compatible with the differences in % of remaining synaptic connections known from histopathological studies for their age range (see the approach in Märcher-Rørsted et al., Hear. Res., 2022)

(3) Synergetic effects of CND and listening effort:

Could you test whether there is an interaction between CNR and listening effort? (e.g. one could hypothesize that MA subjects with the largest CND have also higher listening effort).

<https://doi.org/10.7554/eLife.102823.1.sa0>