

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

Reviewed Preprint

v2 • June 26, 2025

Revised by authors

Reviewed Preprint

v1 • November 14, 2024

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

 Copyright information

eLife Assessment

This study aims to clarify the effects of cochlear neural degeneration on auditory processing in listeners with normal audiograms (sometimes referred to as 'hidden hearing loss'). 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, the findings pose **compelling** evidence that cochlear synaptopathy is associated with a significant part of hearing difficulties in complex environments.

<https://doi.org/10.7554/eLife.102823.2.sa3>

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 experienced in older age. However, current clinical tests fail to capture hearing difficulties for ~10% of middle-aged adults with normal hearing thresholds 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 are associated with 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 considered 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 (13, 24). However, these behavioral readouts may minimize subliminal changes in perception that are reflected in listening effort but *not* in accuracies (25–27). Specifically, two individuals may show similar accuracies on a listening task, but one individual may need to exert substantially more listening effort to achieve the same accuracy as the other. Here, we used a cross-species approach, combined with simultaneous measurements of behavior and listening effort, to show that CND was 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 gerbils. In the human model, we simultaneously measured pupil-indexed listening effort in participants as they performed a speech-in-noise task and showed 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 CN

Middle-aged (40-55 years) and young adult (18-25 years) listeners were recruited to participate in this study (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 ≤ 25 dB HL for octave frequencies between 250 Hz to 8 kHz (Fig. 1B, Table 1), consistent with WHO guidelines for normal hearing (28). Threshold differences were exaggerated in MAs at extended high frequencies (>8 kHz), which are seldom clinically measured but may be a marker for accumulated lifetime noise damage ((17, 29–31), 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; (32)) 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; (34)) was not significantly different between age groups (Fig. 1F, Table 4). Participants also had normal cognitive function indexed by the Montreal Cognitive Assessment (MoCA ≥ 25 ; (35)) and comparable working memory scores assessed using the operation span task (OSPAN) ((36), 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 CN using neural ensemble responses from the auditory periphery phase-locked to the stimulus amplitude envelope via the 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 at rapid amplitude modulation (AM) frequencies above 600 Hz have been shown to relate to underlying CN in animal models (12, 14) and in humans (37). 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 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 the noise floor, with average signal to noise ratios (SNRs) of 8 dB in younger and 4 dB in middle-aged adults (Fig. I, J). Statistically significant age-related decreases in EFR amplitudes were only present for EFRs to the 1024 Hz AM rate, which has putative generators in the auditory nerve (12, 14) but were not present for slower AM rates with putative generators in the midbrain or cortex (Fig. 1K, Table 5).

To confirm that the EFR parameters used here were indeed sensitive to putative CN, we measured EFRs using identical stimuli, acquisition, and analysis parameters in young (22 wk) and middle-aged (80 wk) Mongolian gerbils (Fig. 2A). The hearing range of gerbils largely overlaps with that of humans at speech frequencies (39), 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 humans (Fig. 2C, Table 6). CN in gerbils was assessed using immunohistological analysis of cochlear whole mounts, where the cell bodies, presynaptic ribbon terminals and the post-synaptic glutamate receptor patches

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.

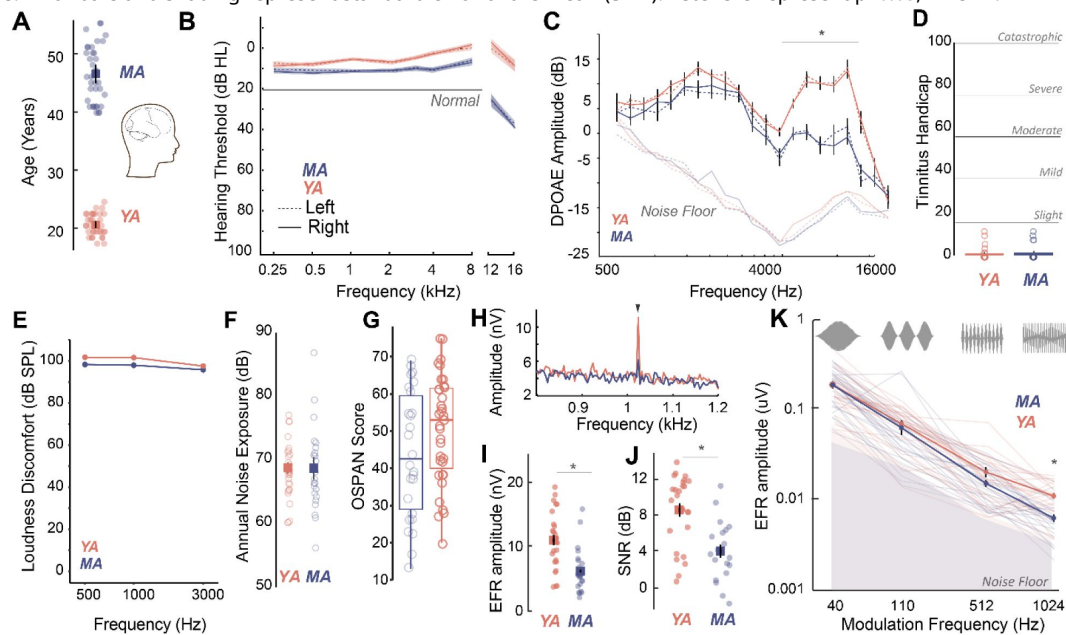


Table 1

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	0.275	.6
AM	1.47	79.49	151.407	< .001 ***
Group:AM	1.47	79.49	0.151	.929
1024:Group	1	55	23.8	< .001 ***
512:Group	1	56	3.171	0.083
110:Group	1	55	0.491	0.487
40:Group	1	54	0.027	.870

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

were immunostained, visualized using confocal 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 abilities with the clinically-used Quick Speech-in-Noise (QuickSIN; (40)) 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 ranging from 25 to 0 dB SNR levels in 5 dB steps (**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). On each trial, participants were required to repeat a target sentence, which contained five key words for identification. Clinically, QuickSIN is scored as dB SNR loss, reflecting the SNR level required to correctly identify key words in noise correctly half the time. No significant age-related differences were observed in clinically scored QuickSIN dB SNR loss (**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 5dB 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 (41–43) while participants performed the QuickSIN task (**Fig. 3A**). Pupillary changes were analyzed using growth curve analysis (GCA, (44)). 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 for individual variability. Two second-order GCAs were fit to different time-windows (**Table 9–10**, see methods). One time window encompassed the onset of the masker through the first 2.8s of the target sentence (listening window). The second window spanned from the end of the target sentence up to the verbal response prompt (integration window). These two time-windows were 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 (45), while the integration window reflects error correction, working memory and comparisons with predictive internal models (46, 47). 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 increase at more challenging SNRs (**Fig. 3D**). Both younger and middle-aged adults showed increases in pupil-indexed effort prior to overt decreases in behavioral performance (**Fig. 3E**). While MAs exhibited larger increases in listening effort compared to YA, 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 (**Fig. 3F**). Pupillary slopes obtained from the GCA increased with task difficulty for both YA and MA. However, middle-aged adults showed a larger increase in listening effort than younger adults with decreasing SNRs, with significant age group listening effort differences at

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 ± 0.86 weeks, $n = 14$) and middle-aged (80 ± 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.

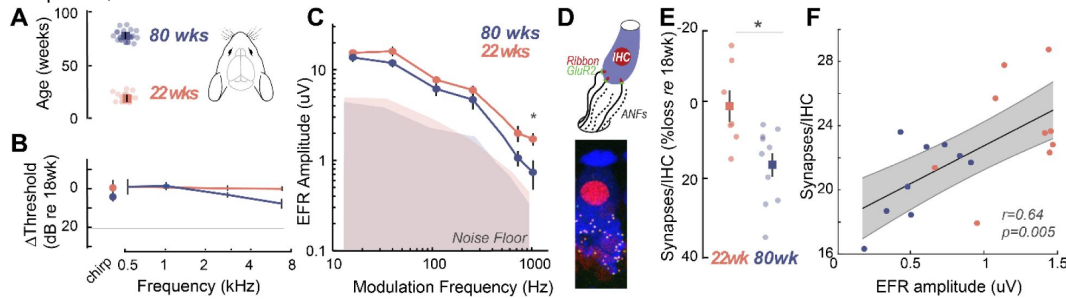


Table 6

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 22 and 80 week-old gerbils using 1-way ANOVA

Measure	22wk (n)	80wk (n)	DFn	DFd	F-Value	p
Synapse Counts	14	12	1	16	4.877	.042 *

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

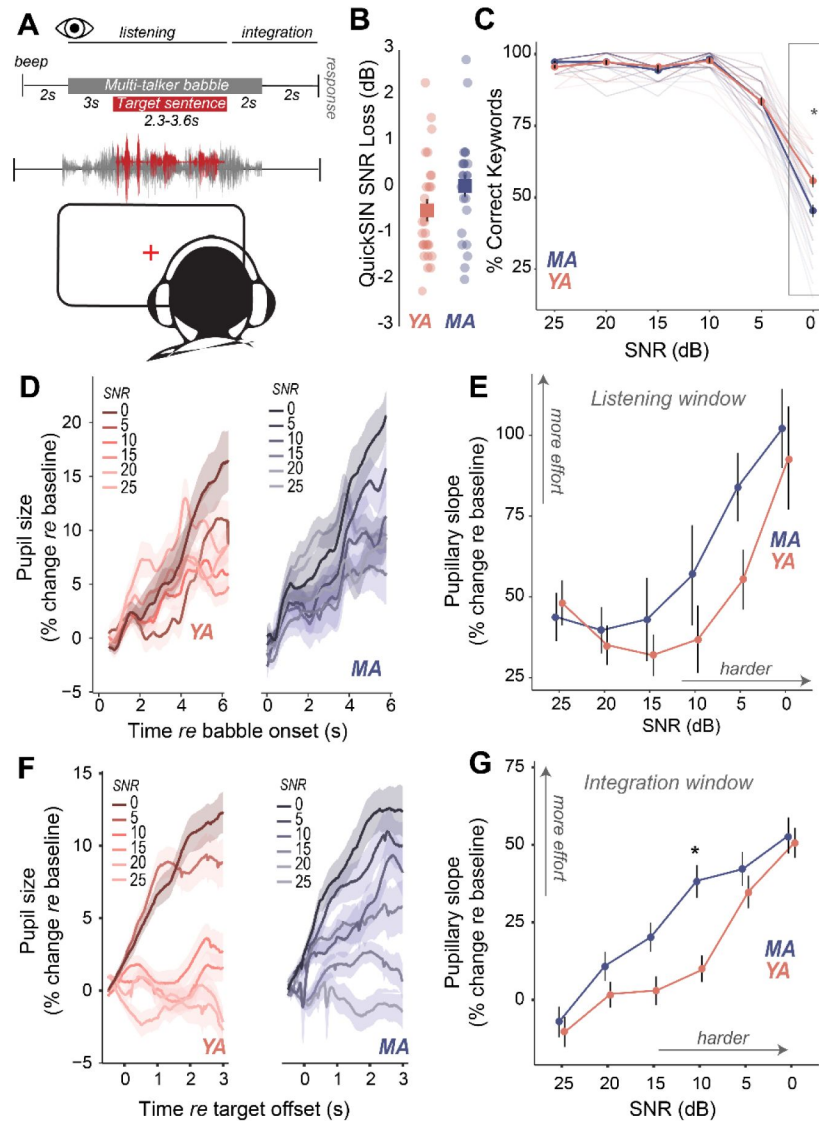


Figure 3.

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 middle-aged (MA) and younger adults (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	p-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)

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)

10dB SNR, 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 putative CND assessed using EFRs at 1024 Hz (**Fig. 4A** [↗](#)). This suggests that peripheral deafferentation may manifest 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 compared to younger adults (**Fig. 3G** [↗](#)), even though behavioral performance was near ceiling for both age groups. 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 growing evidence suggesting that pupil-indexed listening effort to maintain behavioral performance at moderate task difficulties is predictive of behavioral performance at more challenging listening conditions ([48](#) [↗](#)). 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 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 ([49](#) [↗](#)). QuickSIN performance at 0dB SNR (scaled to 0-100) was used as the outcome variable and all other measured variables were inserted as input variables. The elastic net penalized regression framework is a robust method that blends Lasso's ability to perform variable selection and Ridge's ability to handle multicollinearity and grouped covariates. The fitted elastic net regression model showed 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 younger and middle-aged adults ($r = 0.64$ /(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 ([50](#) [↗](#)). Increasing evidence suggests that middle-age is a critical period of rapid changes in brain function ([51](#) [↗](#), [52](#) [↗](#)). The resilience of the brain in keeping with degenerative processes that begin to occur in middle-age predicts further age-related degeneration in later life and presents a critical opportunity for early intervention ([50](#) [↗](#), [53](#) [↗](#)–[55](#) [↗](#)). Hearing loss in middle-age has recently been identified as the largest modifiable risk factor for dementia and Alzheimer's disease later in life ([2](#) [↗](#)). However, the number of middle-aged patients who seek help for hearing difficulties but show no abnormal clinical indicators suggests the need for the development of sensitive biomarkers for hearing challenges experienced by this population ([3](#) [↗](#), [5](#) [↗](#), [6](#) [↗](#), [56](#) [↗](#)).

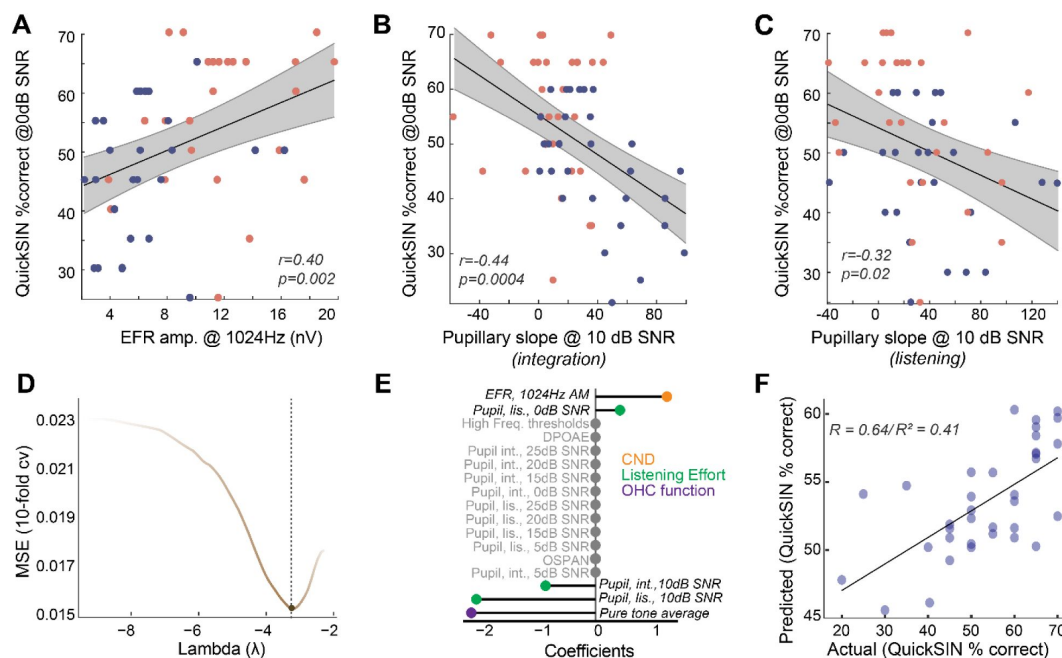


Figure 4.

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 (β) 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 substantial noise exposure history (9–11). Peripheral deafferentation triggers compensatory mechanisms across sensory, language, and attentional systems (57–60). But our understanding of the perceptual consequences of cochlear deafferentation are limited by the lack of consensus on sensitive biomarkers for CND (61). Recent studies have identified multiple promising biomarkers for CND in animal models and human populations (21, 37, 62). Reduced wave I amplitudes in the auditory brainstem response are a reliable marker of CND in animal models (7, 12, 63) but can be challenging to obtain in humans (20, 62). The middle-ear muscle reflex, an acoustic measurement of middle-ear immittance driven by efferent feedback to the middle-ear muscles, has also been identified as a promising marker for CND (19, 21, 64). Here, we used the EFR to identify CND in middle-aged adults with normal audiometric thresholds. As opposed to the middle-ear muscle reflex, EFRs measure peripheral neural coding and central auditory activity by exploiting the divergent phase-locking abilities of the ascending auditory pathway (65, 66). EFRs with modulation rates greater than ~1000 Hz have been associated with CND and are considered to reflect the integrity of the auditory nerve (12, 14), given that midbrain and cortical neurons cannot phase-lock to such high rates (65). We observed decreases in EFRs at modulation rates that were selective to the auditory periphery (i.e., 1024 Hz) in middle-aged adults, while EFRs at slower modulation rates, likely generated from the central auditory structures, were not different from those in younger adults (**Fig. 1K**). The use of a more rapid onset time in the stimulus modulation envelope, such as the rectangular amplitude modulated tones (RAM EFRs), may result in a larger separation of these groups even at slower modulation rates (67, 68), as sharper onset times result in greater EFR amplitudes (37, 69). However, a more intriguing possibility is that middle-aged adults exhibited an increase in relative central auditory activity, or ‘gain’, in the presence of decreased peripheral neural coding (57, 59). The perceptual consequences of this gain are unclear, but our findings align with emerging evidence suggesting that gain is associated with selective deficits in speech-in-noise abilities (59, 70, 71). EFRs at suprathreshold levels presented here also have contributions from higher frequency regions due to a broader excitation at the cochlea (72, 73). Since cochlear synapse loss is also believed to be flat across frequencies with age, EFRs used here likely index cochlear synapse loss equally across a broad range of frequencies (9, 12, 63). This notion is further supported by emerging evidence that suggests that phase-locking measured to lower frequency pure tones also indexes cochlear synaptopathy in ways that are similar to using a faster modulation rate on a higher frequency tone (74, 75).

The Mongolian gerbil provides a robust model for cross-species comparisons with aging humans, due to 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**). Additionally, age-related changes in the EFR were associated with confirmed CND (**Fig. 2F**). CND in gerbils reached ~20% in the middle-aged 80 week group tested here, which is less than what has been observed in middle-aged humans, where CND estimates typically reach 40–50% by the fifth decade of life (9). However, our EFRs were still sensitive to this degree of CND, reiterating that EFRs are a sensitive metric for measuring cochlear deafferentation. Additionally, we confirmed that the gerbils used in this study did not show any changes in hearing thresholds (**Fig. 2B**). Hence, they were unlikely to have strial degenerations that are known to occur in older gerbils that affect auditory thresholds (76). The synapse loss patterns and EFR amplitude changes seen here in gerbils were in agreement with earlier studies using alternate rodent models (12, 14, 69), further confirming that age-related cochlear synapse loss is a pervasive mammalian phenomenon that can be captured using EFRs to rapid modulation frequencies (~1000 Hz).

Strong evidence links CND with altered neural coding of sounds in multiple ascending auditory stations (12, 58, 59). However, the perceptual consequences of CND on speech-in-noise abilities remain unclear (61). Evidence for overt behavioral deficits have been 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 further tease apart these differences (17, 37). In the current study, behavioral deficits began to emerge only at the most challenging SNR levels (Fig. 3). However, perceptual deficits in terms of listening effort began to appear prior to behavioral changes.

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-related changes in pupil diameter (26, 41–43, 77). The mechanisms underlying these pupillary changes are still under study (78, 79) but are hypothesized to involve the Locus Coeruleus – Norepinephrine (LC-NE) system (80, 81). Here, we observed that pupil-indexed listening effort increased in middle-aged adults, even when behavioral performance was 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 (25, 82, 83). Potentially confounding factors impacting pupil measurement such as the decrease of pupil dynamic range with aging (84, 85), participant fatigue, or task habituation (45, 77, 86), can vary between individuals for a multitude of reasons (87). 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 separate approaches – a Pearson's correlation and the elastic net regression model (Fig. 4B–D). We have previously demonstrated similar results in a different test group of young adult participants (48). 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 (78, 79) and thus being a poorer predictor of concomitant behavioral performance. Additionally, our elastic net regression model suggested 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 (Fig. 4D), lending support to studies suggesting that the definition of clinically 'normal' may need revision (3, 88). Our findings demonstrate a need for next-generation diagnostic measures of auditory processing that incorporate both neurophysiological encoding of the temporal elements of sound and cognitive factors associated with listening effort to better capture one's listening abilities. Future studies will directly test the 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 (89, 90).

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 ≥ 25 ; [Nasreddine et al., 2005](#)), normal hearing thresholds (≤ 25 dB HL 250-8000 Hz), no severe tinnitus as self-reported via the Tinnitus Handicap Inventory (THI; [32](#)), and Loudness Discomfort Levels (LDLs) ≥ 80 dB HL at .5, 1, and 3 kHz ([33](#)). Participants were not required to have specific complaints of speech perception in noise difficulties. The Beck's depression Inventory (BDI [91](#)) was administered and participants were excluded if they reported thoughts of self-harm, determined by any response to survey item nine greater than 0. Participants self-reported American English fluency. Thirty-five young (18-25 years old, male = 10) and 37 middle-aged participants (40-55 years old, male = 10) met all eligibility criteria and were tested further using the battery described below.

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 Astera² 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 in alternating polarities. 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; (34 [↗](#))) 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 (34 [↗](#)).

OSPAN

Participants also completed the automated version of the OSPAN task (92 [↗](#)), as a metric of working memory (36 [↗](#)). 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. Following a series of arithmetic-letter presentations, participants were required to recall the letters that were displayed in the order that they appeared. 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.

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; (40 [↗](#))). 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 (93 [↗](#)). Each QuickSIN test list consisted of six sentences masked in four-talker babble at the following SNR levels: 25, 20, 15, 10, 5, and 0 dB. All participants completed four test lists. Participants listened to the sentences through Sennheiser circumaural headphones. The masker was presented at 60 dB SPL, and the sound level of the target sentences were varied to obtain the required SNR level. Participants were instructed 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 (40 [↗](#), 94 [↗](#)). In addition, we calculated the standard clinical QuickSIN score of dB SNR loss, which reflects the lowest SNR level that an individual can accurately identify words 50% of the time. For each participant, the dB SNR loss score was calculated for each test list separately using the following equation: $25.5 - (\text{sum of keywords identified in list})$ (40 [↗](#)). Then, the mean dB SNR loss across all four test lists was calculated and used for analysis.

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. Monocular left-eye pupillary responses were recorded at a 1000 Hz sampling rate using an EyeLink 1000 Plus Desktop Mount camera and chin rest (SR Research). Nine-point eye-tracker calibration was performed prior to the start of the experiment. To start each trial, participants were required to fixate on a cross in the center of the screen for a minimum of 500 ms. This

fixation criterion was applied to control for the effects of saccades, which can alter pupil diameter, and to minimize pupil foreshortening errors (95–97). After meeting the 500 ms fixation criteria, a 100 ms 1000 Hz beep was presented to alert the participant to the start of the trial. There was a 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. After the end of the background masker, there was a two second delay followed by 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

Pupillary data were processed in R (98) using the *eyelinker* package (99) 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 50 Hz. 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 (96). 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.

Growth curve analyses (GCA; Mirman, 2014) were used to obtain a measure of the slope of the pupillary response during QuickSIN. GCA uses orthogonal polynomial time terms to model distinct functional forms of the pupillary response over time. Two GCAs were fit using a second-order orthogonal polynomial to model the interaction of age group with SNR level, separately for the listening window and the integration window. 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 (100) and *p*-values were estimated using the *lmerTest* package (101).

For the listening window, the best-fit GCA model included fixed effects of each time term (ot1, ot2), SNR (reference = 25), Group (reference = younger), 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 = younger), 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 BioSemi ActiveTwo EEG system while participants were seated in a recliner. Stimuli were presented using ER-3C transducers [Etymotic Research Inc., Elk Grove, IL] with gold-foil tiptrodes placed in the ear canals to deliver sound stimuli and record additional channels of evoked potentials. EFRs were recorded to a 250 ms tone with a carrier frequency of 3000Hz, amplitude modulated (AM) at 40, 110, 512, and 1024Hz. Stimuli were presented in alternating polarity, with 500 repetitions each at 85dB SPL to the right ear. Each token was presented at 3.1 repetitions/second, for a period of 322ms.

Preprocessing

EFRs from the Fz to the ipsilateral (right) tiptrode were processed and analyzed using custom written scripts in MATLAB v. 2022a (Mathworks Inc., Natick, Massachusetts). 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 auditory brain stem responses (ABRs) and ending 10ms after stimulus offset. The maximum amplitude of the FFT peak at one of three adjacent bins (~3Hz) around the modulation frequency of the AM rate was 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 was characterized by noise-level data logging and was periodically monitored. Data logging revealed an average noise level of 56 dB, with transients not exceeding 74 dB during regular housing conditions and 88dB once a week during cage changes. All animal procedures were 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. Subdermal electrodes (Ambu) were then 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 (TDT)). The average duration of isoflurane anesthesia during the electrode setup process was approximately 10 min. After placing electrodes, animals were injected with dexmedetomidine (Dexdomitor, 0.3 mg/kg subdermal) and taken off the isoflurane. Dexmedetomidine is an alpha-adrenergic agonist that acts as a sedative and an analgesic and is known to decrease motivation but preserve behavioral and neural responses in rodents ([102](#), [103](#)). This helps to maintain animals in an un-anesthetized state, where they still respond to pain stimuli, such as a foot pinch,

but are otherwise compliant to recordings for a period of about 3 hours. The time window for the effects of isoflurane to wear off was determined empirically as 9 minutes, based on ABRs waveforms and latencies, as well as the response to foot pinch stimuli. Recordings then commenced 15 minutes after cessation of isoflurane.

Stimulus presentation, acquisition, and analysis

Stimuli were presented to the right ear of the animal using insert earphones (ER3C, Etymotic), which matched the stimulus presentation in humans. Stimuli presentation and acquisition were done by a custom program for gerbils in LabView. The output from the insert earphones was calibrated using a Bruel Kjaer microphone and was found to be within ± 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 ABRs 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 300Hz to 30000Hz, 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 AM tones (5ms rise/fall, 250ms duration, 3.1 repetitions/s, alternating polarity) at a 3KHz carrier frequency presented 30dB above auditory thresholds obtained using ABRs at 3kHz. The modulation frequency was systematically varied from 16Hz to 1024Hz 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 μ V were rejected. 250 artifact-free trials of each polarity were averaged to compute the EFR waveform. FFTs were performed on the averaged time-domain waveforms starting 10ms 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 (~ 3 Hz each) around the modulation frequency was recorded as the peak FFT amplitude. The FFT amplitude at the AM frequency was reported as the EFR amplitude. The noise floor of the EFR was calculated as the average of 5 frequency bins (~ 3 Hz each) above and below the central three bins. A response was deemed as significantly above the noise floor if the FFT amplitude was at least 6 dB greater than 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, then 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 ([104](#)) and custom plugins in ImageJ. Cochlear stacks were obtained at the target frequency (3kHz) spanning the cuticular plate to the synaptic pole of ~ 10 hair cells (in 0.25 μ 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 inner hair cells 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 [↗](#), 105 [↗](#)).

Statistical analysis

Analysis of Variance (ANOVA)

Normality of all variables 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 (106 [↗](#)). The function 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

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 were reported.

Elastic Net Regression

We used an linear model with an elastic net penalization/regularization (49 [↗](#)) 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 (107 [↗](#)). The relative strength of selection and shrinkage is controlled by the hyper-parameters λ and α : a higher λ 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. 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 β can be written as $Pen(\beta) = \lambda(\alpha\|\beta\|_1 + (1 - \alpha)/2\|\beta\|^2)$. 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 ($\|\beta\|^2$) encourages grouped variable selection and removes the limitation of number of selected variables unlike LASSO while stabilizing the selection path. To choose the tuning parameters λ and α , we used a 10-fold cross-validation that minimizes the out-of-sample root mean-squared error (RMSE). We used the R packages *glmnet* (108 [↗](#)) and *caret* (109 [↗](#)) for training the elastic net regularizer.

Data Availability

All data reported and analyzed in this study can be found on the Open Science Framework at <http://doi.org/10.17605/OSF.IO/4BGDA> [↗](#)

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. Thanks also to Dr. Simon Warkins, Katherine Helfrich and Mike Calderon at the Center for Biological Imaging at the University of Pittsburgh, supported by NIH grant 1S10RR028478-01 for collaboration on confocal imaging, and the Clinical and Translational Science Institute at the University of Pittsburgh, supported by the NIH Clinical and Translational Science Award (CTSA) program, grant UL1 TR001857 for assistance with participant recruitment.

Additional information

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, JRM

Funding

National Institute on Deafness and Other Communication Disorders (R21DC018882)

National Institute on Deafness and Other Communication Disorders (T32DC011499)

National Institute on Deafness and Other Communication Disorders (F31DC020085)

National Institutes of Health (1S10RR028478)

National Institutes of Health (UL1TR001857)

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 [Google Scholar](#)
2. Livingston G., et al. (2017) **Dementia prevention, intervention, and care** *The Lancet* **390**:2673–2734 [Google Scholar](#)
3. Hind S. E., et al. (2011) **Prevalence of clinical referrals having hearing thresholds within normal limits** *Int. J. Audiol* **50**:708–716 [Google Scholar](#)
4. Tremblay K. L., et al. (2015) **Self-Reported Hearing Difficulties Among Adults With Normal Audiograms: The Beaver Dam Offspring Study** *Ear Hear.* **36**:E290–E299 [Google Scholar](#)
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**:13636 [Google Scholar](#)
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**:e51419 <https://doi.org/10.7554/eLife.51419> | [Google Scholar](#)
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 [Google Scholar](#)
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 [Google Scholar](#)
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> | [Google Scholar](#)
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> | [Google Scholar](#)
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 [Google Scholar](#)
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 [Google Scholar](#)
13. Mepani A. M., et al. (2021) **Envelope following responses predict speech-in-noise performance in normal-hearing listeners** *J. Neurophysiol* **125**:1213–1222 [Google Scholar](#)

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> | [Google Scholar](#)
15. Prendergast G., et al. (2017) **Effects of noise exposure on young adults with normal audiograms I: Electrophysiology** *Hear. Res* **344**:68–81 [Google Scholar](#)
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 [Google Scholar](#)
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 [Google Scholar](#)
18. Grant K. J., et al. (2022) **Predicting neural deficits in sensorineural hearing loss from word recognition scores** *Sci. Rep* **12**:8929 [Google Scholar](#)
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 [Google Scholar](#)
20. Mehraei G., et al. (2016) **Auditory Brainstem Response Latency in Noise as a Marker of Cochlear Synaptopathy** *J. Neurosci* **36**:3755–3764 [Google Scholar](#)
21. Bharadwaj H. M., et al. (2022) **Cross-species experiments reveal widespread cochlear neural damage in normal hearing** *Commun Biol* **5**:1–10 [Google Scholar](#)
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** [Google Scholar](#)
23. Prendergast G., et al. (2017) **Effects of noise exposure on young adults with normal audiograms II: Behavioral measures** *Hear. Res* **356**:74–86 [Google Scholar](#)
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> | [Google Scholar](#)
25. Pichora-Fuller M. K., et al. (2016) **Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL)** *Ear Hear* **37**:5–27 [Google Scholar](#)
26. Peelle J. E. (2018) **Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior** *Ear Hear* **39**:204–214 [Google Scholar](#)
27. 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 [Google Scholar](#)
28. WHO (no date) **International Classification of Functioning, Disability and Health (ICF)** <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health>

29. Škerková M., et al. (2023) **Extended high-frequency audiometry: hearing thresholds in adults** *Eur. Arch. Otorhinolaryngol* **280**:565–572 [Google Scholar](#)
30. 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 [Google Scholar](#)
31. Lough M., Plack C. J. (2022) **Extended high-frequency audiometry in research and clinical practice** *J. Acoust. Soc. Am* **151**:1944 [Google Scholar](#)
32. Newman C. W., Jacobson G. P., Spitzer J. B. (1996) **Development of the Tinnitus Handicap Inventory** *Arch. Otolaryngol. Head Neck Surg* **122**:143–148 [Google Scholar](#)
33. 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 [Google Scholar](#)
34. 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 [Google Scholar](#)
35. Nasreddine Z. S., et al. (2005) **The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment** *J. Am. Geriatr. Soc.* **53**:695–699 [Google Scholar](#)
36. Turner M. L., Engle R. W. (1989) **Is working memory capacity task dependent?** *J. Mem. Lang* **28**:127–154 [Google Scholar](#)
37. Mepani A. M., et al. (2021) **Envelope following responses predict speech-in-noise performance in normal-hearing listeners** *J. Neurophysiol* **125**:1213–1222 [Google Scholar](#)
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 [Google Scholar](#)
39. Ryan A. (1976) **Hearing sensitivity of the mongolian gerbil, *Meriones unguiculatus*** *J. Acoust. Soc. Am* **59**:1222–1226 [Google Scholar](#)
40. 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 [Google Scholar](#)
41. Beatty J. (1982) **Phasic not tonic pupillary responses vary with auditory vigilance performance** *Psychophysiology* **19**:167–172 [Google Scholar](#)
42. 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 [Google Scholar](#)
43. 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 [Google Scholar](#)
44. Mirman D. (2014) **Growth Curve Analysis and Visualization Using R** Chapman and Hall/CRC [Google Scholar](#)
45. 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**:105010 [Google Scholar](#)

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 [Google Scholar](#)
47. Winn M. B. (2023) **Time Scales and Moments of Listening Effort Revealed in Pupillometry** *Semin. Hear* **44**:106–123 [Google Scholar](#)
48. McHaney J. R., Hancock K. E., Polley D. B., Parthasarathy A. (2024) **Sensory representations and pupil-indexed listening effort provide complementary contributions to multi-talker speech intelligibility** *Sci. Rep* **14**:30882 [Google Scholar](#)
49. Zou H., Hastie T. (2005) **Regularization and Variable Selection Via the Elastic Net** *J. R. Stat. Soc. Ser. B Stat. Methodol* **67**:301–320 [Google Scholar](#)
50. Dohm-Hansen S., et al. (2024) **The “middle-aging” brain** *Trends Neurosci* **0** [Google Scholar](#)
51. Schaum N., et al. (2020) **Ageing hallmarks exhibit organ-specific temporal signatures** *Nature* **583**:596–602 [Google Scholar](#)
52. Salthouse T. A. (2019) **Trajectories of normal cognitive aging** *Psychol. Aging* **34**:17–24 [Google Scholar](#)
53. 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 [Google Scholar](#)
54. 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 [Google Scholar](#)
55. 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 [Google Scholar](#)
56. 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 [Google Scholar](#)
57. 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 [Google Scholar](#)
58. Chambers A. R., et al. (2016) **Central Gain Restores Auditory Processing following Near-Complete Cochlear Denervation** *Neuron* **89**:867–879 [Google Scholar](#)
59. Resnik J., Polley D. B. (2021) **Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise** *Neuron* **109**:984–996 [Google Scholar](#)

60. 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 [Google Scholar](#)
61. Bramhall N., et al. (2019) **The search for noise-induced cochlear synaptopathy in humans: Mission impossible?** *Hear. Res* **377**:88–103 [Google Scholar](#)
62. Bharadwaj M H., et al. (2019) **Non-Invasive Assays of Cochlear Synaptopathy - Candidates and Considerations** *Neuroscience* **407**:53–66 [Google Scholar](#)
63. Sergeyenko Y., Lall K., Liberman M. C., Kujawa S. G. (2013) **Age-Related Cochlear Synaptopathy: An Early-Onset Contributor to Auditory Functional Decline** *J. Neurosci* **33**:13686–13694 [Google Scholar](#)
64. Valero M. D., Hancock K. E., Liberman M. C. (2016) **The middle ear muscle reflex in the diagnosis of cochlear neuropathy** *Hear. Res* **332**:29–38 [Google Scholar](#)
65. Joris P. X., Schreiner C. E., Rees A. (2004) **Neural processing of amplitude-modulated sounds** *Physiol. Rev* **84**:541–577 [Google Scholar](#)
66. Parida S., et al. (2024) **Rapid and objective assessment of auditory temporal processing using dynamic amplitude-modulated stimuli** *Commun Biol* **7**:1–10 [Google Scholar](#)
67. Vasilkov V., Garrett M., Mauermann M., Verhulst S. (2021) **Enhancing the sensitivity of the envelope-following response for cochlear synaptopathy screening in humans: The role of stimulus envelope** *Hear. Res* **400**:108132 [Google Scholar](#)
68. Garrett M., et al. (2025) **Deciphering Compromised Speech-in-Noise Intelligibility in Older Listeners: The Role of Cochlear Synaptopathy** *eNeuro* **12** [Google Scholar](#)
69. Parthasarathy A., Bartlett E. L. (2011) **Age-related Auditory Deficits in Temporal Processing in F-344 Rats** *Neuroscience* **192**:619–630 [Google Scholar](#)
70. 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 [Google Scholar](#)
71. Rumschlag J. A., et al. (2022) **Age-Related Central Gain with Degraded Neural Synchrony in the Auditory Brainstem of Mice and Humans** *Neurobiol. Aging* **115**:50–59 [Google Scholar](#)
72. Parthasarathy A., Lai J., Bartlett E. L. (2016) **Age-Related Changes in Processing Simultaneous Amplitude Modulated Sounds Assessed Using Envelope Following Responses** *Jaro-J. Assoc. Res. Otolaryngol* **17**:119–132 [Google Scholar](#)
73. Lai J., Bartlett E. L. (2018) **Masking Differentially Affects Envelope-following Responses in Young and Aged Animals** *Neuroscience* **386**:150–165 [Google Scholar](#)
74. Märcher-Rørsted J., et al. (2022) **Age-related reduction in frequency-following responses as a potential marker of cochlear neural degeneration** *Hear. Res* **414** [Google Scholar](#)
75. Ponsot E., Devolder P., Dhooge I., Verhulst S. (2024) **Age-related decline in neural phase-locking to envelope and temporal fine structure revealed by frequency following**

responses: A potential signature of cochlear synaptopathy impairing speech intelligibility *bioRxiv* <https://doi.org/10.1101/2024.12.11.628010> | [Google Scholar](#)

76. Gratton M. A., Schulte B. A. (1995) **Alterations in microvasculature are associated with atrophy of the stria vascularis in quiet-aged gerbils** *Hear. Res* **82**:44–52 [Google Scholar](#)
77. 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.* **2006** **73**:1431–1443 [Google Scholar](#)
78. 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 [Google Scholar](#)
79. de Gee J. W., et al. (2020) **Pupil-linked phasic arousal predicts a reduction of choice bias across species and decision domains** *eLife* **9**:e54014 <https://doi.org/10.7554/eLife.54014> | [Google Scholar](#)
80. 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 [Google Scholar](#)
81. Reimer J., et al. (2016) **Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex** *Nat. Commun* **7** [Google Scholar](#)
82. 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 [Google Scholar](#)
83. 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 [Google Scholar](#)
84. 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> | [Google Scholar](#)
85. Piquado T., Isaacowitz D., Wingfield A. (2010) **Pupillometry as a measure of cognitive effort in younger and older adults** *Psychophysiology* **47**:560–569 [Google Scholar](#)
86. Tryon W. W. (1975) **Pupillometry: A Survey of Sources of Variation** *Psychophysiology* **12**:90–93 [Google Scholar](#)
87. 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** [Google Scholar](#)
88. Hunter L. L., et al. (2020) **Extended high frequency hearing and speech perception implications in adults and children** *Hear. Res* **397**:107922 [Google Scholar](#)
89. Shinn-Cunningham B. G., Best V. (2008) **Selective attention in normal and impaired hearing** *Trends Amplif* **12**:283–299 [Google Scholar](#)

90. 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 [Google Scholar](#)
91. 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 [Google Scholar](#)
92. 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 [Google Scholar](#)
93. Meyer J., Dentel L., Meunier F. (2013) **Speech recognition in natural background noise** *PLoS One* **8**:e79279 [Google Scholar](#)
94. 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**:321 [Google Scholar](#)
95. 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 [Google Scholar](#)
96. 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** [Google Scholar](#)
97. 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 [Google Scholar](#)
98. R Core Team (2022) **R: a language and environment for statistical computing** Vienna, Austria: R Foundation for Statistical Computing <https://www.r-project.org/>
99. Barthelme S. (2021) **eyelinker: An R package for importing data from EyeLink ASC files** *GitHub* <https://github.com/a-hurst/eyelinker>
100. Bates D., Mächler M., Bolker B., Walker S. (2015) **Fitting Linear Mixed-Effects Models Using lme4** *J. Stat. Softw* **67**:1–48 [Google Scholar](#)
101. Kuznetsova A., Brockhoff P. B., Christensen R. H. B. (2017) **lmerTest Package: Tests in Linear Mixed Effects Models** *J. Stat. Softw* **82**:1–26 [Google Scholar](#)
102. 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 [Google Scholar](#)
103. 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 [Google Scholar](#)
104. Greenwood D. D. (1990) **A Cochlear Frequency-Position function for several species –29 Years Later** *J. Acoust. Soc. Am* **87**:2592–2605 [Google Scholar](#)
105. 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 [Google Scholar](#)

- 106. Girden E. R. (1992) **ANOVA: Repeated measures** Sage Publications, Inc [Google Scholar](#)
- 107. Smith S. S., Jahn K. N., Sugai J. A., Hancock K. E., Polley D. B. (2024) **The human pupil and face encode sound affect and provide objective signatures of tinnitus and auditory hypersensitivity disorders** *bioRxiv* <https://doi.org/10.1101/2023.12.22.571929> | [Google Scholar](#)
- 108. Friedman J., Hastie T., Tibshirani R. (2010) **Regularization Paths for Generalized Linear Models via Coordinate Descent** *J. Stat. Softw* **33**:1–22 [Google Scholar](#)
- 109. Kuhn M. (2008) **Building Predictive Models in R Using the caret Package** *J. Stat. Softw* **28**:1–26 [Google Scholar](#)

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

+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⁺

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

⁺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

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. But 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.?

[Note added after revision: the authors have added new data, references, and discussion that have answered my initial questions].

(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 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.

[Note added after revision: the authors now make it clear in their response that the modeling tells us how much of the current data can be explained but not necessary about generalization to other datasets.]

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 fixings 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.

[Note added after revision: the authors have added to the Discussion to put their work into a broader perspective.]

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.

[Note added after revision: my intention with this comment was not to make a philosophical or nitty-gritty point about statistics. It was more of a follow on to the previous point. Because I don't know what sort of effect size is big enough to matter (for whatever purpose), I don't find the statistical significance (or lack thereof) of the effect size observed to be informative. But I don't think there is anything more that the authors can or should do in this regard.]

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.102823.2.sa2>

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 for 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 suggest 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 - CND- 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 underly 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, 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 that 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 you comment further on why you think you found a significant difference only at the highest mod. frequency of 1024 Hz in your 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. 2fm or 3fm) see e.g. Garrett et al. bioXiv <https://www.biorxiv.org/content/biorxiv/early/2024/04/30/2020.06.09.142950.full.pdf>

Did you 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 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 you 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 your 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 you 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 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 largest CND have also the higher listening effort)

Comments on revised version:

The authors did well to address all the points raised in my review. This paper will make an important contribution to our assessment of the sources of age-related auditory processing deficits beyond the cochlea that impair speech intelligibility.

<https://doi.org/10.7554/eLife.102823.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Public Reviews:

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

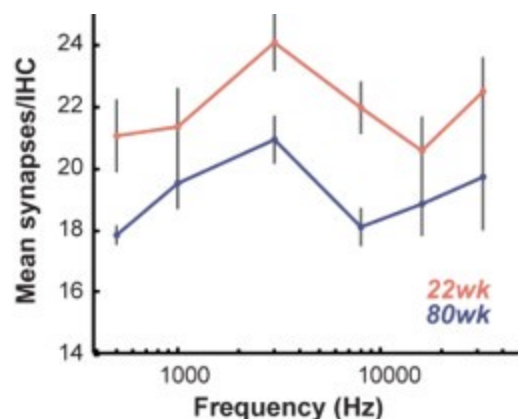
fibers coming from the 3 kHz place? And are we sure that this is the same for gerbils and humans given the differences in cochlear geometry, head size, etc.?

Thank you for raising this important point. The frequency regions that contribute to the generation of EFRs, especially at the suprathreshold sound levels presented here are expected to be broad, with a greater leaning towards higher frequencies and reaching up to one octave above the center frequency. We have investigated this phenomenon in earlier published articles using both low/high pass masking noise and computational models using data from rodent models and humans (Encina-Llamas et al. 2017; Parthasarathy, Lai, and Bartlett 2016). So, the expectation here is that the EFRs reflect a wider frequency region centered at 3 kHz. The difference in cochlear activation regions between humans and gerbils for EFRs have not been systematically studied to our knowledge but given the general agreement between humans and other rodent models stated above, we expect this to be similar to gerbils as well. Additionally, all current evidence points to cochlear synapse loss with age being flat across frequencies, in contrast to cochlear synapse loss with noise which is dependent on the bandwidth of the noise exposure.

Histological evidence for this flat loss across frequencies is found in mice and human temporal bones (Parthasarathy and Kujawa 2018; Sergeyenko et al. 2013; Wu et al. 2018). We find this to be true in our gerbils as well. Author response image 1 shows the patterns of synapse loss as a function of cochlear place. We focused on synapse loss at 3 kHz to keep the analysis focused on the center frequency of the stimulus and minimize compounding errors due to averaging synapse counts across multiple frequency regions. We have now added some explanatory language in the discussion.

Author response image 1.

Cochlear synapse counts per inner hair cell (IHC) in young and middle-aged gerbils as a function of cochlear frequency.



(2) Unless I misunderstood, the predictive power of the final model was not tested on heldout 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.

The goal of the analysis in this current manuscript was inference, rather than prediction, i.e., to find the important/significant variables that contribute to speech intelligibility in noise, rather than predicting the behavioral deficit of speech performance in a yet-unforeseen sample of adults.

Additionally, we used a repeated 10-fold cross-validation approach for our model building exercise as detailed in the Elastic Net Regression section of the methods. This repeated-cross validation calculated the mean square error on a held-out fold and average it repeatedly to reduce the inherent variability of randomly choosing a validation set. The repeated 10-fold CV approach is both more stable and efficient compared to a validation set approach, or splitting the data into two segments: training and test, and provides a better estimate of the test error by utilizing more observations for training (vide Chapter 5, (James et al. 2021). These predictive MSEs along with the R-squared for the final model give us a good idea of the predictive performance, as, for the linear model the R-squared is the correlation between the observed and the predicted response. Future studies with a larger sample size can facilitate having a designated test set and still have enough statistical power to perform predictive analyses.

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.

Studies with CND have so far been largely inferential in humans, since currently we cannot confirm CND in vivo. Hence any measures of putative CND in humans can only be interpreted based on evidence from other animal studies. Our translational approach is partly meant to address this deficit, as mentioned in the Introduction section. By using identical stimuli, recording, acquisition and analysis parameters we hope to reduce some of the variability that may be associated with this inference between human and other animal models. Until direct measurements of CND in humans are possible, the intended goal is to provide diagnostic biomarkers that have face validity – i.e., that explain variance related to speech intelligibility deficits in this population.

We've added more to the discussion to state that our work demonstrates the need for next generation diagnostic measures of auditory processing that incorporate cognitive factors associated with listening effort to better capture speech in noise perceptual abilities.

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.

We agree that pre-determining sample sizes is the optimal approach towards designing a study. The sample sizes here were chosen a priori based on previously published data in young adults with normal hearing thresholds (McHaney et al. 2024; Parthasarathy et al. 2020). With the lack of published literature especially for the EFRs at 1024Hz AM in middle aged adults, there are practical challenges in pre-determining the sample size (given a

prefixed power and an effect size) with limited precursors to supply good estimates of the parameters (e.g., mean, s.d. for each age group for a two-sample test). We hope that this data set now shared will enable us and other researchers to conduct power analyses for successive studies that use similar metrics on this population.

Several authors, including Heinsburg and Weeks (2022) argue that post-hoc power could be “misleading and simply not informative” and encourage using other indicators of poorly powered studies such as the width of the confidence interval. Since the elastic net estimate is a non-linear and non-differentiable function of the response values—even for fixed tuning parameters—it is difficult to obtain an accurate estimate of its standard error (Tibshirani and Taylor 2012). While acknowledging the limitations of post-hoc power analyses, we performed a retrospective power calculation for our linear model with the predictors that we selected (EFR @ 1024Hz, Pupil slope for QuickSIN at selected SNRs and analyses windows, and PTA). The calculated Cohen’s effect size was 0.56, which is considered large (Cohen 2013). With this effect size, a power analysis with our sample size revealed a very high retrospective power of 0.99 with a significance level of 0.05. The minimum number of subjects needed to get 80% power with this effect size was $N = 21$. Hence for the final model, we are confident that our results hold true with adequate statistical power.

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 CNL, but I am unsure how much closer we are to anything that is practically useful.

Thank you for your comments. We hope that this study establishes a framework for the eventual development of the next generation of objective diagnostics tests in the hearing clinic that provide insights into the underlying neurophysiology of the auditory pathway and take into effect top-down contributors such as listening effort.

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. bioXiv: <https://www.biorxiv.org/content/biorxiv/early/2024/04/30/2020.06.09.142950.full.pdf>).

Previously published studies in animal models by us and others suggests that EFRs elicited to AM rates > 700Hz are most sensitive to confirmed CND (Parthasarathy and Kujawa 2018; Shaheen, Valero, and Liberman 2015). This is likely because these AM rates fall well outside of phase-locking limits in the auditory midbrain and cortex (Joris, Schreiner, and Rees 2004), and hence represent a ‘cleaner’ signal from the auditory periphery that may not be modulated by complex excitatory/inhibitory feedback circuits present more centrally (Caspary et al. 2008). We have also demonstrated that we are able to acquire high quality EFRs at 1024Hz AM rates both in a previously published study in young normal hearing adults (McHaney et al. 2024), and in middle aged adults in the present study as seen in Fig. 1 H-J. We posit that the lack of age-related differences at the lower AM rates may be indicative of compensatory plasticity with age (central ‘gain’) that occurs with age in more central regions of the auditory pathway (Auerbach, Radziwon, and Salvi 2019; Parthasarathy and Kujawa 2018). We now expand on this in the discussion. A secondary reason for the lack of change in slower modulation rates may be the difference in stimulus between sinusoidally amplitude modulated tones used here, and the rectangular amplitude modulated tones in other studies, as discussed in response to the comment below.

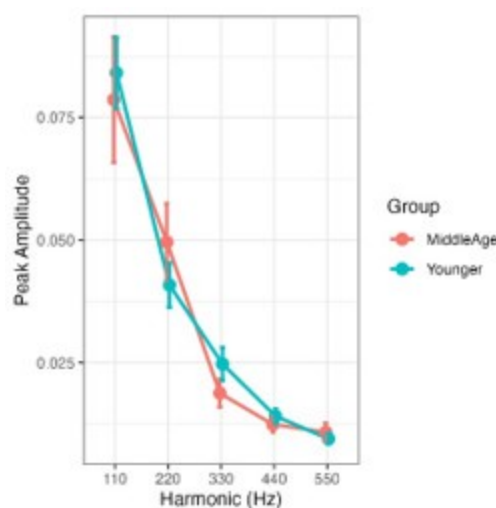
Furthermore, some previous EEG experiments in humans that SAM tones with modulation freqs. of ~100Hz showed that EFRs do not exhibit a single peak, i.e. there are peaks not only at fm but also for the first harmonics (e.g. 2fm or 3fm) see e.g. Garrett et al. bioXiv <https://www.biorxiv.org/content/biorxiv/early/2024/04/30/2020.06.09.142950.full.pdf>. 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).

We examined peak amplitudes for the AM rate and harmonics for the 110 Hz AM condition as shown in Author response image 2. The quantified amplitudes of the first four harmonics did not differ with age ($p > .08$).

Additionally, the harmonic structures obtained were also not as robust as would be expected with rectangular amplitude modulated stimuli. The choice of sinusoidal modulation may explain why. We have previously published studies systematically modulating the rise time of the envelope per cycle in amplitude modulated tones, where the individual period of the envelope is described by $Env(t) = t^x(1-t)$, where t goes from 0 to 1 in one period, and where $x = 0.05$ represents a highly damped envelope akin to the rising envelope of a rectangular modulation, and $x = 1$ representing a symmetric, near-sinusoidal envelope (Parthasarathy and Bartlett 2011). The harmonic structure was much more developed in the damped envelopes compared to the symmetric envelopes and response amplitudes were also higher for the damped envelopes overall, a result also observed in Mepani et. al., 2021. Hence, we believe the rapid rise time may contribute to the harmonic structures evidenced in studies using RAM stimuli, and the absence of this rapid onset may result in reduced harmonic structures in our EFRs. Some language regarding this issue is now added to the discussion.

Author response image 2.

Harmonics analysis for the first four harmonics of envelope following responses elicited to the 110Hz AM stimulus.



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?

We believe the current results are in close agreement with these studies showing deficits in pure tone phase locking with age. These tones are typically at ~300-500Hz or above, and phase locking to these tones likely involves the same or similar peripheral neural generators in the auditory nerve and brainstem. Emerging evidence also seems to suggest that TFS coding measured using pure tone phase locking is closely related to sound with amplitude modulation in the same range (Ponsot et al. 2024). Unpublished observations from our lab support this view as well. In this data set, we begin to see EFR responses at 512 Hz diverge

with age, but this difference does not reach statistical significance. This may be due to specific AM frequencies selected or a lack of statistical power. Using more continuous AM frequency sweeps such as with our recently published dynamic amplitude modulated tones (Parida et al. 2024) may help resolve these AM frequency specific challenges and help us investigate changes over a broader range of AM frequencies. Ongoing studies are currently exploring this hypothesis. Some explanatory language is now presented in the discussion.

(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.

Current evidence seems to suggest that humans have more synaptic loss than gerbils, though exact comparison of lifespan between the two species is challenging due to differences in slopes of growth trajectories between species. Post-mortem temporal bone studies demonstrate a ~40-50% loss of synapses in humans by the fifth decade of life. On the other hand, our gerbils in the current study showed approximately 15-20% loss. Based on our findings and previous studies, it is reasonable to assume that our gerbil data underestimate the temporal processing deficits that would be seen in humans due to CND.

We have added this information and citations to the discussion section.

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.

We chose to record EFRs to stimuli presented at 85 dB SPL in humans, as opposed to 30 dB SL, because 30 dB SL in humans would have corresponded to an intensity that makes EEG recordings unfeasible. The average PTA across younger and middle-aged adults was 7.51 dB HL (~19.51 dB SPL), which would have resulted in an average stimulus intensity of ~50 dB SPL at 30 dB SL. This intensity level would have been far too low to reliably record EFRs without presenting many thousands of trials. In a pilot study, we recorded EFRs at 75 dB SL, which equated to an average of 83.9 dB SPL. Thus, we chose the suprathreshold level of 85 dB SPL for the current study to obtain reliable responses with just 1000 trials.

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)

We agree with the reviewer that phenomenological models would be a useful approach to examining differences between age groups and species. We have previously used the Zilany/Carney model to examine differences in EFRs with age in rats (Parthasarathy, Lai, and Bartlett 2016). It is unclear if such models will directly translate to responses from gerbils. However, this is a subject of ongoing study in our lab.

(3) Synergetic effects of CND and listening effort:

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

We have previously reported that EFRs and listening effort are not linearly related (McHaney et al. 2024). We found the same to be largely true in the current study as well. We ran correlations between EFR amplitudes at 1024 Hz and listening effort at each SNR level in the listening and integrations windows. We did not observe any significant relationships between EFRs at 1024 Hz and listening effort in the listening window (all p s > .05). In the integration window, we did see a significant correlation between listening effort at SNR 5 and EFRs at 1024 Hz, which was significant after correcting for multiple comparisons ($r = -.42$, p -adj = .021). However, we chose to not report these multiple oneto-one correlations in the current study and instead opted for the elastic net regression analysis to better understand the multifactorial contributions to speech-in-noise abilities. These results also do not preclude non-linear relationships between listening effort and EFRs which may be present based on emerging results (Bramhall, Buran, and McMillan 2025), and will be explored in future studies.

Recommendations for the authors:

Reviewer #1 (Recommendations for the authors):

A few more minor comments/questions:

(1) How old were the YA gerbils on average? 18 weeks, or 19 weeks, or 22 weeks?

Young gerbils were on average 22 weeks. We have updated the manuscript accordingly.

(2) "Gerbils share the same hearing frequency range as humans" is misleading; the gerbil hearing range extends to much higher frequencies.

We have revised the statement to say: "The hearing range of gerbils largely overlaps with that of humans, making them an ideal animal model for direct comparison in crossspecies studies."

(3) The writing contains more than a few typos and grammatical errors.

We have completed a thorough revision to correct for grammatical and typographical errors.

(4) Suggesting that correlation and linear modelling are "independent" methods is misleading since they are both measuring linear associations. A better word would be "different".

Thank you for this suggestion. We have rephrased the sentence as "two separate approaches"

(5) The phrase "Our results reveal perceptual deficits ... driven by CND" in the abstract is too strong. Correlation is not causation.

We have revised this phrase to say they "are associated with CND."

Reviewer #2 (Recommendations for the authors):

More general comments:

(1) Recruitment criterion related to hearing-in-noise difficulties:

If I understood correctly, the middle-aged participants recruited for this study do not have specific hearing in noise difficulties, some could, as with 10% in the general population, but they were not recruited using this criterion. If this is correct, this should be stated explicitly, as it constitutes an important methodological choice and a difference with your eLife 2020 study. If you were to use this specific recruitment criterion for both groups here, what differences would you expect?

Our participants were not required to have specific complaints of speech perception in noise challenges to be eligible for this study. We included middle-aged adults here, as opposed to only younger adults as in Parthasarathy et al. (2020), with the assumption that middle-aged adults were likely to have some cochlear synapse loss and individual variability in the degree of synapse loss based on post-mortem data from human temporal bones. We have recently published studies identifying the specific clinical populations of patients with self-perceived hearing loss, including those adults who have received assessments for auditory processing disorders (Cancel et al. 2023). Ongoing studies in the lab are aimed at recruiting from this population.

It is striking here that the QuickSIN test does not exhibit the same variability at low SNRS here as with the digits-in-noise used in your eLife 2020 study. Why would QuickSIN more appropriate than the Digits-in-noise test? Would you expect the same results with the Digits-in-noise test?

Our 2020 eLife study investigated the effects of TFS coding in multi-talker speech intelligibility. TFS coding is specifically hypothesized to be related to multi-talker speech, compared to broadband maskers. The digits test was appropriate in that context as the ‘masker’ there was two competing speakers also speaking digits. In this study, we wanted to test the effects of CND on speech in noise perception using clinically relevant speech in noise tests. The Digits test is devoid of linguistic context and is essentially closed set (participants know that only a digit will be presented). However, QuickSIN consists of open set sentences of moderate context, making it closer to real world listening situations. Additionally, we recently published pupillometry recorded in response to QuickSIN in young adults (McHaney et al. 2024) and identified QuickSIN as a promising screening tool for self-perceived hearing difficulties (Cancel et al. 2023). These factors informed our choice of using QuickSIN in the current study.

(2) Why is the increase in listening effort interpreted as an increase in gain? please clarify (p10, 1st paragraph; [these data suggest a decrease in peripheral neural coding, with a concomitant increase in central auditory activity or ‘gain’])

In the above referenced paragraph, we were discussing the increase in 40 Hz AM rate EFRs in middle-aged adults as an increase in central gain. We have revised parts of this paragraph to better communicate that we were discussing the EFRs and not listening effort: “We observed decreases in EFRs at modulation rates that were selective to the auditory periphery (i.e., 1024 Hz) in middle-aged adults, while EFRs primarily generated from the central auditory structures were not different from those in younger adults (Fig. 1K). These data suggest that middle-aged adults exhibited an increase in central auditory activity, or ‘gain’, in the presence of decreased peripheral neural coding. The perceptual consequences of this gain are unclear, but our findings align with emerging evidence suggesting that gain is associated with selective deficits in speech-in-noise abilities”

(3) Further discussion on the relationship/differences between markers EFR marker of CND (this study) and MEMR marker of CND(Bharadwaj et al., 2022) is needed.

We now make mention of other candidate markers of CND (ABR wave I and MEMRs) in the discussion and expand on why we chose the EFR.

(4) Further analyses and discussion would be needed to be related to extended high-freq thresholds:

Did you test for a potential correlation of your EFR marker of CND with extended high-freq. thresholds ? (could be paralleling the amount of CND in these individuals) Why won't you also consider measuring extended HF in Gerbils?

We acknowledge that there is increasing evidence to suggest extended high frequency thresholds may be an early marker for hidden hearing loss/CND. We have examined an additional correlation for extended high frequency pure tone averages (8k-16k Hz) with EFR amplitudes at 1024 Hz AM rate, which revealed a significant relationship ($r = -.43$, $p < .001$). However, we opted to exclude this analysis from our current study as we wanted to reduce reporting on several one-to-one correlations. Therefore, we chose the elastic net regression model to examine individual contributions to speech in noise abilities. EHF thresholds were included in the elastic net regression models, but were not found to be significant upon accounting for individual differences in PTA.

Additionally, our electrophysiological experimental paradigm was not designed with the consideration of extended high frequencies—we used ER3C transducers which are not optimal for frequencies above ~6kHz. Future studies could use transducers such as the ER2 or free field speakers to examine the influence of extended high frequencies on the EFRs and measure high frequency thresholds in gerbils.

Minor Comments:

(1) Abstract: repetition of 'later in life' in the first two sentences - please reformulate.

We have revised the first two sentences to state: “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 in older age.”

(2) Sentence on page 3 [However, these behavioral readouts may minimize subliminal changes in perception that are reflected in listening effort but not in accuracies (26-28)] is not clear.

We've added a sentence just after that states: “Specifically, two individuals may show similar accuracies on a listening task, but one individual may need to exert substantially more listening effort to achieve the same accuracy as the other.”

(3) The second paragraph of page 11 should go to a methods (model) section, not to the discussion.

We have now moved a portion of this paragraph to the Elastic Net Regression subsection of the Statistical Analysis in the Methods.

(4) Please check references: references 13 and 25 are identical.

Fixed

References

- Auerbach, Benjamin D., Kelly Radziwon, and Richard Salvi. 2019. “Testing the Central Gain Model: Loudness Growth Correlates with Central Auditory Gain Enhancement in a Rodent Model of Hyperacusis.” *Neuroscience* 407:93–107. <https://doi.org/10.1016/j.neuroscience.2018.09.036>.
- Bramhall, Naomi F., Brad N. Buran, and Garnett P. McMillan. 2025. “Associations Between Physiological Indicators of Cochlear Deafferentation and Listening Effort in Military Veterans with Normal Audiograms.” *Hearing Research*, April, 109263. <https://doi.org/10.1016/j.heares.2025.109263>.
- Cancel, Victoria E., Jacie R. McHaney, Virginia Milne, Catherine Palmer, and Aravindakshan Parthasarathy. 2023. “A Data-Driven Approach to Identify a Rapid Screener for Auditory Processing Disorder Testing Referrals in Adults.” *Scientific Reports* 13 (1): 13636. <https://doi.org/10.1038/s41598-023-40645-0>.
- Caspary, D. M., L. Ling, J. G. Turner, and L. F. Hughes. 2008. “Inhibitory Neurotransmission, Plasticity and Aging in the Mammalian Central Auditory System.” *Journal of Experimental Biology* 211 (11): 1781–91. <https://doi.org/10.1242/jeb.013581>.
- Cohen, Jacob. 2013. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. New York: Routledge. <https://doi.org/10.4324/9780203771587>.
- Encina-Llamas, Gerard, Aravindakshan Parthasarathy, James Michael Harte, Torsten Dau, Sharon G. Kujawa, Barbara Shinn-Cunningham, and Bastian Epp. 2017. “Hidden Hearing Loss with Envelope Following Responses (EFRs): The off-Frequency Problem: 40th MidWinter Meeting of the Association for Research in Otolaryngology.” In .
- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2021. *An Introduction to Statistical Learning: With Applications in R*. Springer Texts in Statistics. New York, NY: Springer US. <https://doi.org/10.1007/978-1-0716-1418-1>.
- Joris, P. X., C. E. Schreiner, and A. Rees. 2004. “Neural Processing of Amplitude-Modulated Sounds.” *Physiological Reviews* 84 (2): 541–77. <https://doi.org/10.1152/physrev.00029.2003>.
- McHaney, Jacie R., Kenneth E. Hancock, Daniel B. Polley, and Aravindakshan Parthasarathy. 2024. “Sensory Representations and Pupil-Indexed Listening Effort Provide Complementary Contributions to Multi-Talker Speech Intelligibility.” *Scientific Reports* 14 (1): 30882. <https://doi.org/10.1038/s41598-024-81673-8>.
- Parida, Satyabrata, Kimberly Yurasits, Victoria E. Cancel, Maggie E. Zink, Claire Mitchell, Meredith C. Ziliak, Audrey V. Harrison, Edward L. Bartlett, and Aravindakshan Parthasarathy. 2024. “Rapid and Objective Assessment of Auditory Temporal Processing Using Dynamic Amplitude-Modulated Stimuli.” *Communications Biology* 7 (1): 1–10. <https://doi.org/10.1038/s42003-024-07187-1>.
- Parthasarathy, A., and E. L. Bartlett. 2011. “Age-Related Auditory Deficits in Temporal Processing in F-344 Rats.” *Neuroscience* 192:619–30. <https://doi.org/10.1016/j.neuroscience.2011.06.042>.
- Parthasarathy, A., J. Lai, and E. L. Bartlett. 2016. “Age-Related Changes in Processing Simultaneous Amplitude Modulated Sounds Assessed Using Envelope Following Responses.” *Jaro-Journal of the Association for Research in Otolaryngology* 17 (2): 119–32. <https://doi.org/10.1007/s10162-016-0554-z>.
- Parthasarathy, A., Kenneth E Hancock, Kara Bennett, Victor DeGruttola, and Daniel B Polley. 2020. “Bottom-up and Top-down Neural Signatures of Disordered Multi-Talker Speech Perception in Adults with Normal Hearing.” Edited by Barbara G Shinn-Cunningham, Huan

Luo, Fan-Gang Zeng, and Christian Lorenzi. eLife 9 (January):e51419. <https://doi.org/10.7554/eLife.51419>.

Parthasarathy, Aravindakshan, and Sharon G. Kujawa. 2018. “Synaptopathy in the Aging Cochlea: Characterizing Early-Neural Deficits in Auditory Temporal Envelope Processing.” *The Journal of Neuroscience*. <https://doi.org/10.1523/jneurosci.324017.2018>.

Ponsot, Emmanuel, Pauline Devolder, Ingeborg Dhooge, and Sarah Verhulst. 2024. “AgeRelated Decline in Neural Phase-Locking to Envelope and Temporal Fine Structure Revealed by Frequency Following Responses: A Potential Signature of Cochlear Synaptopathy Impairing Speech Intelligibility.” *bioRxiv*. <https://doi.org/10.1101/2024.12.11.628010>.

Sergeyenko, Yevgeniya, Kumud Lall, M. Charles Liberman, and Sharon G. Kujawa. 2013. “Age-Related Cochlear Synaptopathy: An Early-Onset Contributor to Auditory Functional Decline.” *Journal of Neuroscience* 33 (34): 13686–94. <https://doi.org/10.1523/jneurosci.1783-13.2013>.

Shaheen, L. A., M. D. Valero, and M. C. Liberman. 2015. “Towards a Diagnosis of Cochlear Neuropathy with Envelope Following Responses.” *J Assoc Res Otolaryngol*. <https://doi.org/10.1007/s10162-015-0539-3>.

Tibshirani, Ryan J., and Jonathan Taylor. 2012. “Degrees of Freedom in Lasso Problems.” *The Annals of Statistics* 40 (2): 1198–1232. <https://doi.org/10.1214/12-AOS1003>.

Wu, P. Z., L. D. Liberman, K. Bennett, V. de Gruttola, J. T. O’Malley, and M. C. Liberman. 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>.

<https://doi.org/10.7554/eLife.102823.2.sa0>