Age dependent deficits in speech recognition in quiet and noise are reflected in MGB activity and cochlear onset coding

Konrad Dapper, Stephan M Wolpert, Jakob Schirmer, Stefan Fink, Etienne Gaudrain, Deniz Başkent, Wibke Singer, Sarah Verhulst, Christoph Braun, Ernst Dalhoff, et al.

NeuroImage 305 (2025) 120958

Contents lists available at ScienceDirect

NeuroImage

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

Age dependent deficits in speech recognition in quiet and noise are reflected in MGB activity and cochlear onset coding

Konrad Dapper [a,b,1], Stephan M. Wolpert [a,1], Jakob Schirmer [a], Stefan Fink [a], Etienne Gaudrain [c], Deniz Başkent [d], Wibke Singer [a], Sarah Verhulst [e], Christoph Braun [f,g,h], Ernst Dalhoff [a], Lukas Rüttiger [a], Matthias H.J. Munk [a,b], Marlies Knipper [a,*]

[a] Department of Otolaryngology, Head and Neck, University of Tübingen, Tübingen 72076, Germany
[b] Department of Biology, Technical University 64287 Darmstadt, Darmstadt, Germany
[c] Lyon Neuroscience Research Center, Université Claude Bernard Lyon 1, CNRS UMR5292, INSERM U1028, Center Hospitalier Le Vinatier -Bâtiment 462-Neurocampus, 95 boulevard Pinel, Lyon, France
[d] Department of Otorhinolaryngology, University Medical Center Groningen (UMCG), Hanzeplein 1, BB21, Groningen 9700RB, the Netherlands
[e] Department of Information Technology, Ghent University, Zwijnaarde 9052, Belgium
[f] MEG-Center, University of Tübingen, Tübingen 72076, Germany
[g] HIH, Hertie Institute for Clinical Brain Research, Tübingen 72076, Germany
[h] CIMeC, Center for Mind and Brain Research, University of Trento, Rovereto 38068, Italy

ARTICLE INFO

Keywords:
EEG recording
Speech recognition
Cochlear synaptopathy
Aging
Hearing aid

ABSTRACT

The slowing and reduction of auditory responses in the brain are recognized side effects of increased pure tone thresholds, impaired speech recognition, and aging. However, it remains controversial whether central slowing is primarily linked to brain processes as atrophy, or is also associated with the slowing of temporal neural processing from the periphery. Here we analyzed electroencephalogram (EEG) responses that most likely reflect medial geniculate body (MGB) responses to passive listening of phonemes in 80 subjects ranging in age from 18 to 76 years, in whom the peripheral auditory responses had been analyzed in detail (Schirmer et al., 2024). We observed that passive listening to vowels and phonemes, specifically designed to rely on either temporal fine structure (TFS) for frequencies below the phase locking limit (<1500 Hz), or on the temporal envelope (TENV) for frequencies above phase locking limit, entrained lower or higher neural EEG responses. While previous views predict speech content, particular in noise to be encoded through TENV, here a decreasing phoneme-induced EEG amplitude over age in response to phonemes relying on TENV coding could also be linked to poorer speech-recognition thresholds in quiet. In addition, increased phoneme-evoked EEG delay could be correlated with elevated extended high-frequency threshold (EHF) for phoneme changes that relied on TFS and TENV coding. This may suggest a role of pure-tone threshold averages (PTA) of EHF for TENV and TFS beyond sound localization that is reflected in likely MGB delays. When speech recognition thresholds were normalized for pure-tone thresholds, however, the EEG amplitudes remained insignificant, and thereby became independent of age. Under these conditions, poor speech recognition in quiet was found together with a delay in EEG response for phonemes that relied on TFS coding, while poor speech recognition in ipsilateral noise was observed as a trend of shortened EEG delays for phonemes that relied on TENV coding. Based on previous analyses performed in these same subjects, elevated thresholds in extended high-frequency regions were linked to cochlear synaptopathy and auditory brainstem delays. Also, independent of hearing loss, poor speech-performing groups in quiet or with ipsilateral noise during TFS or TENV coding could be linked to lower or better outer hair cell performance and

Abbreviations: ABR, auditory brainstem responses; ANOVA, one-way analysis of variance; ASSR, auditory steady-state responses; CWT, continuous wavelet transform; DPOAE, distortion-product otoacoustic emission; EHF, extended high frequencies; ERF, event-related fields; OLSA, German Oldenburger sentence test (German version of the international matrix test); PCA, principal component analysis; PC, principle component; pDPOAE, short-pulsed DPOAE; PEA, phoneme-evoked amplitude; PED, phoneme-evoked delay; PLL, phase-locking limit; PNOT, PT-normalized OLSA threshold; PT, pure-tone; PTA, pure-tone average; PTA4, pure-tone average (clinical standard); PTA-HF, pure-tone average for high frequencies; PTA-EHF, pure-tone average for extended high frequencies; PTA-LF, pure-tone average for low frequencies; PTT, PT thresholds; SR, spontaneous rate; TENV, temporal envelope; TFS, temporal fine-structure.

* Corresponding author.
E-mail address: marlies.knipper@uni-tuebingen.de (M. Knipper).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.neuroimage.2024.120958
Received 29 May 2024; Received in revised form 25 November 2024; Accepted 26 November 2024
Available online 30 November 2024

delayed or steeper auditory nerve responses at stimulus onset. The amplitude and latency of MGB responses to phonemes requiring TFS or TENV coding, dependent or independent of hearing loss, may thus be a new predictor of poor speech recognition in quiet and ipsilateral noise that links deficits in synchronicity at stimulus onset to neocortical activity. Amplitudes and delays of speech EEG responses to syllables should be reconsidered for future hearing-aid studies.

1. Introduction

Age-related hearing loss (presbycusis) is a growing problem in aging societies (Peelle and Wingfield, 2016), particularly as it is associated with cognitive decline (Golub et al., 2020; Gregory et al., 2020; Livingston et al., 2017). It has devastating effects on the everyday life of subjects, they being handicapped in speech recognition, thus risking social isolation (Keithley, 2020). A reduction of amplitudes and delay of auditory evoked responses with increasing age that correlates with speech recognition difficulties and cognitive decline has been described in numerous studies: Up to the cortex, auditory evoked potentials are found to be reduced and delayed in those components responding within ~200 ms of stimulus onset (Allison et al., 1984; Bertoli et al., 2002; Peelle and Wingfield, 2016; Zhao et al., 2024). Latency shifts in aging subjects have been suggested to influence neural processing, e.g., induce brain atrophy and demyelination (Humes et al., 2013; Peelle and Wingfield, 2016), both independently of peripheral hearing loss (Profant et al., 2014), or without a corresponding increase in latency in the early components (Price et al., 2017). On the other hand, temporal precision deficits and latency shifts during aging have been described in auditory steady state responses **(ASSR)** (Korczak et al., 2012; Lu et al., 2022; Plourde et al., 2008), frequency following responses, and auditory brainstem responses (Anderson et al., 2021; Bidelman and Carter, 2023; Presacco et al., 2016a, 2016b). These all include contributions of synchronized responses phase locked to the stimulus in rather early subcortical generators (Korczak et al., 2012; Lu et al., 2022; Plourde et al., 2008). In addition, speech discrimination deficits occur even in subjects with nearly normal audiograms, as based on pure tone threshold (**PTT**) (Apoux and Bacon, 2004; Bajin et al., 2022; Carcagno and Plack, 2022; Frisina, 2009; Fullgrabe et al., 2014; Guest et al., 2018; Hunter et al., 2020; Le Prell et al., 2013; Levy et al., 2015; Moore et al., 2017; Motlagh Zadeh et al., 2019; Pichora-Fuller et al., 2007; Song et al., 2022; Vitela et al., 2015). Here, particularly speech-intelligibility deficits in noise are suggested to be associated with cochlear synaptopathy, the so-called `hidden hearing loss´, that is a form of primary neural degeneration characterized by damaged synaptic connections between the inner hair cells and auditory-nerve fibers, and is found in animals (Kujawa and Liberman, 2009; Sergeyenko et al., 2013) as well as in humans (Bidelman and Carter, 2023; Harris et al., 2021; Lai and Bidelman, 2022; Wu et al., 2019). Psychophysical experiments with vocoded speech reduced to its temporal envelope demonstrated that speech content, particularly in noise (Carney, 2018; Elhilali et al., 2009; Garrett et al., 2024; Johannesen et al., 2016; Le Prell, 2019; Marrufo-Perez and Lopez-Poveda, 2022; Parthasarathy et al., 2019; Shaheen et al., 2015; Shannon et al., 1995; Snell and Frisina, 2000), is predominantly encoded through temporal envelope modulation (**TENV**), referring to information processing above the human phase-locking limit (PLL). In contrast, it is the role of synchronized firing patterns (phase locking) of neurons in the auditory nerve, that preserve the temporal fine structure (**TFS**) in basilar-membrane vibrations at each characteristic frequency after spectral decomposition in the cochlea, mainly linked to phase locking for sound localization (Lorenzi et al., 2006; Verschooten et al., 2019). Whether phase locking and place information display a role beyond sound localization and is possibly relevant for the discrimination of speech in humans is however under debate (Lorenzi et al., 2006; Verschooten et al., 2019). Here, we examine contributors to speech-discrimination deficits that potentially influence synchronized responses phase locked to the stimuli through EEG recording.

1.1. Previous findings in the same subjects relevant for the study

On the assumption that phase locking is used for binaural processing in humans up to about 1500 Hz (Oxenham, 2018; Verschooten et al., 2019), we previously tested the role of speech stimuli contrasting below and above phase locking in age-dependent speech discrimination using a dual strategy: On the one hand we tested syllable discrimination of phonemes that were designed to specifically rely on either TFS below the human PLL **(1500 Hz)** or TENV above the human PLL (> 1500 Hz) (Schirmer et al., 2024). On the other hand, these psychoacoustic tests of phoneme stimulation were combined with German Oldenburger sentence test **(OLSA)** using three speech-material filtering conditions of either unfiltered broadband speech, low-pass filtered speech (components above 1.5 kHz were deleted from the OLSA power spectrum) or high-pass filtered speech (components below 1.5 kHz were deleted from the OLSA power spectrum) (Schirmer et al., 2024). Young, middle-aged, and older individuals with and without hearing-threshold loss were analyzed for **(i)** PTT that effectively integrate signals over ~500 ms (Smith and Zwislocki, 1975); for **(ii)** an unfiltered, low-pass or high-pass filtered speech-comprehension test in quiet and noise and **(iii)** for methods that enabled the diagnosis of signal transmission at stimulus onset between 10 ms and 100 ms prior to auditory-nerve firing-rate adaptations (Schirmer et al., 2024). The latter methods included short-pulsed distortion-product otoacoustic emissions **(pDPOAE)**, suprathreshold auditory brainstem responses **(ABR),** and the analysis of 116 Hz modulated ASSR that specifically reflects subcortical activity changes to periodic stimuli (Rance, 2008). An unexpected result of these analysis was a noticeable delay and reduction in amplitude of low-frequency induced ABR waves over age that could be linked to elevated PTT, particularly in extended high-frequency regions **(PTA-EHF)**. This finding was interpreted as having the potential of worsened PTA-EHF thresholds over age to negatively influence lower-frequency coding, and thereby speech recognition (Schirmer et al., 2024). In that paper, a principal component analysis and a linear-mixed model to remove the effects of PTT at all measured frequencies (0.25–16 kHz) on the speech-recognition thresholds was also performed, called `pure-tone normalized OLSA threshold´ or **PNOT** (Schirmer et al., 2024). We found in this study that poor speech recognition in PNOT quiet correlated with lower performance in vowel-phoneme discrimination below PLL and thus TFS coding, as well as with reduced amplifier performance of outer hair cells, as detected by lower pDPOAE, and smaller and delayed ABR (Schirmer et al., 2024). On the other hand, listeners with poor speech recognition in noise independent of PTT exhibited poorer performance in vowel-phoneme discrimination above the PLL, and thus TENV coding, increased pDPOAE, and higher uncomfortable loudness levels. These latter response changes were interpreted in the context of steeper slopes of basilar membrane compression that drive a smaller dynamic range (Schirmer et al., 2024).

1.2. Methods applied in subjects who were already included in the previous trial

In the present study, EEG response to phonemes were tested in the same subjects as analyzed in (Schirmer et al., 2024) and using the same stimuli as described above. Specifically, a fundamental frequency of 116

Hz was used for phoneme-induced EEG recordings to **(i)** match the fundamental frequency of OLSA sentences presented to these subjects (Schirmer et al., 2024), and **(ii)** measure subcortical, likely medial geniculate body **(MGB)** activity, rather than cortical components (Kuwada et al., 2002; Lu et al., 2022; Rance, 2008). Also, as used in the previous study (Schirmer et al., 2024), explicitly phonemes below the PLL (/du/, /bu/, /o/) or above PLL (/di/, /bi/, /y/), that were designed to mainly rely on either TFS or TENV coding (Gaudrain et al., 2024) were presented as a passive listening task binaurally in a frozen, pseudo-random order via insert earphones.

1.3. Conclusion

We found strong evidence that differences in amplitudes and delays of likely MGB-based subcortical activity were linked to speech recognition in quiet and in ipsilateral noise, dependent on and independently of age and PTT. MGB amplitudes over age were linked to phonemes that require TENV coding. MGB delays over age were highly correlated with PTA-EHF for both TFS and TENV coding. Moreover, were delays of MGB activity independent of age linked to TFS coding in quiet and shortened MGB activity linked to TENV coding in ipsilateral noise, respectively. The findings are discussed in the context that the changes in the amplitudes and delays of MGB activity have the potential to be a new predictor to link deficits in synchronicity at stimulus onset to neocortical processing deficits.

2. Methods

2.1. Ethics statement

"The study was conducted in the Department of Otolaryngology of the University of Tübingen, and approved by the ethics committee of Tübingen University (Faculty of Medicine; ethical approval number 392/2021BO2). Written, informed consent was given by all participants. All methods followed the Declaration of Helsinki by the World Medical Association (WMA) for human research ethics." (Schirmer et al., 2024).

2.2. Participant recruitment

112 participants (18 to 76 years) were recruited, based on the inclusion and exclusion criteria as described in (Schirmer et al., 2024). Study exclusion included comorbidities, or hearing-related conditions such as tinnitus or previous ear surgery, as well as systemic diseases known to affect hearing (Schirmer et al., 2024). As a result, 80 participants (f/m/d 51/29/0) were finally included that were evenly distributed across three age groups, young (18–29 years, n=25), middle-aged (30–55 years, n=27), and older (56–76 years, n=28). For the individual experimental approaches and correlations, the number of subjects and reasons for specific exclusions are provided in Supplementary Table (Table S1). The age, sex, handedness, middle-ear function as tested by tympanometry of participants included in the present study was listed in detail in the previous study (Schirmer et al., 2024). Given that the statistical analyses involved correlation analyses we estimated the number of subjects using a G-power analysis (Erdfelder et al., 1996). We designed the current study to be sensitive to small effect sizes (0.2). With a standard error probability of 5% and aiming to limit the type II error rate to 15%, a G-power analysis indicated that a minimum sample size of 47 participants was necessary to detect significant correlations.

2.3. Pure-tone audiometry and speech-recognition thresholds

In all 80 subjects, PTTs from 0.125 to 10 kHz [0.125; 0.25; 0.5; 1.0; 1.5; 2.0; 3.0; 4.0; 6.0; 8.0; 10.0 kHz] and the uncomfortable loudness levels from 0.25 to 6 kHz [0.25; 0.5; 1.0; 2.0; 4.0; 6.0 kHz] were measured using Beyerdynamic AT1350A on-ear headphones (Beyerdynamic, Heilbronn, Germany). Additionally, extended high-frequency **(EHF)** thresholds from 11 to 16 kHz [11.2; 12.5; 14.0; 16.0 kHz] using Sennheiser HDA300 (Sennheiser, Wedemark-Wennebostel, Germany) on-ear headphones were measured as described (Schirmer et al., 2024).

As described in Schirmer et al. (2024), four specific pure-tone threshold averages **(PTA)** were calculated for different frequency ranges: (i) low frequencies "**PTA-LF"** [0.125 - 1 kHz], (ii) **"PTA4"** [0.5 - 4 kHz], as mostly used to assess the hearing threshold in clinical otolaryngology (den Besten et al., 2019; von Gablenz and Holube, 2017), (iii) high frequencies **"PTA-HF"** [6 – 10 kHz] to go beyond 6 kHz that is typically deemed sufficient to intelligibly convey speech in communication systems (Monson et al., 2014), and that still cover conventional frequency ranges measured in many clinical studies (Sliwinska-Kowalska, 2015; von Gablenz and Holube, 2017), and (iv) extended high frequencies (EHF) **"PTA-EHF"** [11.2–16 kHz], which have been assumed to play a decisive role in improved spatial hearing (Hunter et al., 2020).

Moreover, as described in Schirmer et al. (2024), the differences in speech-recognition thresholds in quiet, reflecting the state of audibility, and speech-recognition thresholds under ipsilateral noise, challenging the discrimination capacity under background noise, were tested using the standard German Matrix test Oldenburger Satztest **(OLSA)** (Brand and Kollmeier, 2002; Wagener et al., 1999) for either unfiltered "broadband" speech, low-pass filtered speech (components above 1.5 kHz were deleted from the OLSA power spectrum), or high-pass filtered speech (components below 1.5 kHz were deleted from the OLSA power spectrum), as previously described (Garrett et al., 2024; Schirmer et al., 2024). OLSA was performed in quiet and with speech-shaped noise at 70 dB SPL (ipsilateral, contralateral) and with insert earphones. For details: (Schirmer et al., 2024).

2.4. Auditory brain stem responses, ASSR and pDPOAEs recordings

The ABR measurements were performed mono-aurally using three electrodes (Neuroline 720, Ambu, Bad Nauheim, Germany) as described in detail in Schirmer et al. (2024). **In brief:** Acoustic click stimuli were presented at two different stimulus levels (70 dB SPL and 80 dB SPL) with 3000 repetitions. The participants lay on their backs during the measurements to minimize muscle effects. After band-pass filtering between 30 and 2000 Hz, ABR waveform components were averaged (Schirmer et al., 2024). Their latency was determined from the leading positive peak. Click-induced ABR wave peak amplitudes were analyzed using in-ear loudspeakers with a frequency bandwidth limited to approximately 8 kHz that allowed conclusions to be drawn about changes caused by the frequency content within this bandwidth.

The **Auditory Steady-State Response (ASSR)** was measured using the same recording setup as for ABR without changing the position of the participants and as described in Schirmer et al. (2024). **In brief**: The modulation frequency was set to 116 Hz (Vasilkov et al., 2021), and two blocks of 800 epochs each were recorded at carrier frequencies of 4 and 6 kHz at 70 dB SPL rms. ASSR peak amplitudes (μV) were averaged for the first three harmonics (Vasilkov et al., 2021) and analyzed as described (Schirmer et al., 2024).

2.5. Distortion-product otoacoustic emissions

Input-output functions of short-pulsed distortion-product otoacoustic emissions **(pDPOAE)** were measured as decribed in detail in (Schirmer et al., 2024; Zelle et al., 2016): In the present study, pDPOAE input-output functions were measured with an in-ear-probe at 8 frequencies (f_2=0.8; 1.2; 1.5; 2; 3; 4; 6; 8 kHz) using an adaptive algorithm comprising at least four pDPOAE values. For details of recording and method see (Schirmer et al., 2024).

2.6. Phoneme stimuli for EEG recordings

The phoneme stimuli used for the EEG measurements were computer generated using analysis-resynthesis as implemented in the WORLD vocoder (Morise et al., 2016). They had a fundamental frequency (f_0) of 116 Hz, to match the OLSA sentences. As described above, the stimuli were designed assuming an effective PLL of 1.5 kHz (Verschooten et al., 2019; Weiss and Rose, 1988). Some speech stimuli —1 pair of syllables, and 1 pair of vowels— were designed to differ only in the frequency region below the PLL, and could thus rely both on TENV and TFS coding (/du/-/bu/, /u/-/o/). Other stimuli —again 1 pair of syllables, and 1 pair of vowels— were designed to only differ in a frequency region above the PLL, meaning that their discrimination could not rely on TFS, but only on TENV coding (/di/-/bi/, /i/-/y/). The stimuli are available on a public repository (Gaudrain et al., 2024). Time-frequency plotting of the difference between phoneme pairs included a 1/f correction to compensate for the logarithmic frequency scale (Schirmer et al., 2024).

TFS coding stimuli: The vowels /o/ (like in *oder*, "or") and /u/ (like in *Du*, "you" in German) were recorded by a male speaker. Using WORLD, they were then decomposed into an f_0 contour, a time-frequency aperiodicity map, and a spectro-temporal envelope. All manipulations were performed in this decomposed space. The spectral envelope was then smoothed by fitting a Gaussian-mixture representing formants. The center frequencies of the formants for /o/ and /u/ are given in Table 1. The generated envelopes were such that /o/ and /u/ only differed below the 1.5 kHz PLL. After adjusting the f_0, the vowels then resynthesized with a 30-ms raised cosine ramp at the onset and offset, and had a total duration of 414 ms (corresponding to 48 F0 cycles).

The syllables /du/ (like in *Du*, "you" in German) and /bu/ (like in *Budike*, "little shop") were recorded from the same speaker, and also decomposed in WORLD. Differences between stimuli were restricted to a spectro-temporal region comprising the first 100 ms of the stimuli, and all frequencies below the PLL of 1.5 kHz. In this region, the decompositions of the original /du/ and /bu/ were preserved. Outside of this region, the decompositions of the stimuli were averaged, therefore restricting any difference to below 1.5 kHz. Furthermore, the steady-state segment corresponding to the vowel was replaced with the steady-state /u/ vowel described above. Alignment of the stimuli and local corrections were performed to ensure that the syllables remained discriminable. The stimuli were then resynthesized after normalizing the f_0 to 116 Hz, to have a total duration of 471 ms, with a 30-ms offset ramp.

TENV coding stimuli: The same method was used to generate the stimuli differing only above the PLL, hence only allowing TENV coding. The vowels /i/ (like in *die, "the"*) and /y/ (like in *üben, "practice"*) were recorded from the same speaker. The spectral envelope was again represented as Gaussian mixture in such a way that only formants above the PLL would differ between the two stimuli. The center frequencies of the used formants are shown in Table 1. The stimuli were generated to the same duration as the /o/-/u/ pair. For syllables, /di/ (like in *die*, "the") and /bi/ (like in *Bier*, "beer") were, again, recorded from the same speaker and decomposed with WORLD. This time, differences were restricted to a spectro-temporal region comprising the first 100 ms, and restricted to frequencies *above* the PLL of 1.5 kHz. Again, the signal content outside of this region was identical for all stimuli. Furthermore, like for /du/-/bu/, the steady-state vowel segment was replaced with the vowel /i/ as described above.

All the stimuli were spectrally shaped to ensure equal signal-to-noise ratio across frequencies when mixing with the OLSA speech-shaped noise. The stimuli were equalized such that the average level of the stimuli belonging to a given pair was the same for all pairs and provided at 60 dB SPL (L_{eq}). However, to ensure that the level of the formants that remained identical within a pair were not affected, minor overall level fluctuations within a pair were preserved. Calibration was performed with an IEC 60318–4 ear simulator.

Finally, note that in order to keep the experimental time manageable for the participants, the EEG recordings to the vowels already embedded in the syllables were not obtained. Therefore, only /du/, /bu/, /o/ and /di/, /bi/, /y/ were used in the EEG protocol.

Table 1
The frequencies of the first four formants for each vowel contained in the phonemes, in Hz.

Syllable	/o/	/u/	/i/	/y/
F1	380	200	350	350
F2	750	750	1950	1780
F3	3500	3500	2800	2100
F4	10,000	10,000	4000	3000

The vowels /u/ and /i/ are identical to the vowel part in /du/-/bu/ and /di/-/bi/, respectively.

2.7. EEG recording

EEG were recorded to study auditory mid- and long-latency responses to phoneme stimuli (see below). EEG recordings were accomplished using passive stick-on electrodes (Neuroline 720, Ambu, Steinkopfstraße 4, Bad Nauheim, Germany), one mounted on each of the mastoids; the reference electrode was on the center of the forehead below the participant's hairline. The ground electrode was placed between the eyebrows. Signals were preamplified with a gain of 50 (EP50Amp) and digitized by actiCHamp Plus64 (Zeppelinstr.7, Gilching (Munich), Germany). The analog-to-digital converter rate was set to 50,000 samples/s for each channel. To ensure that subjects remained awake during the experiment, the experimenter continuously monitored their brain activity. Here, sudden changes in the roughness of the signal, and strong alpha oscillations served as indicators of sleep. If subjects fell asleep, they were immediately woken up, and any epochs containing sleep were discarded.

All six phonemes (/du/, /bu/, /di/, /bi/, /o/, /y/) were binaurally presented in a frozen pseudo-random order. In total, we played 3456 presentations of alternating polarity per stimulus. In addition, to minimize predictability we ensured that each transition between syllables was equally likely. We presented 4 syllables per second, jittered in time by +/- 5 ms to minimize powerline noise, via insert earphones (ER1–14A) and mu-metal shielded ER2 Transducers (Etymotic Research, Elk Grove Village, Illinois, USA). For the duration of the passive listening task, subjects lay on their back inside a sound-attenuating chamber (Industrial Acoustics, Niederkrüchten, Germany), while watching a silent movie (Metropolis, Fritz Lang, 1929) presented on a Tablet (Samsung a3) placed 60 cm above the subject's head. The horizontal position was chosen to enable the subject to look straight ahead without using their neck muscles. Subjects generally stayed awake during measurements; if subjects fell asleep, they were immediately woken up by the experimenter. Those epochs were excluded from the analysis.

2.8. Data analysis

Data analysis aimed at extracting amplitude and timing information from stable, entrained neural responses to phonemes was carried out using MatLab (2021b).

To reduce the amount of data and computational load, first, all EEG data were downsampled to 5000 samples per second using an anti-aliasing low-pass filter with a cutoff frequency of 2000 Hz and a transition bandwidth of 200 Hz (MATLAB 2021b). Then data were bandpass filtered (60 to 2000 Hz) and segmented into 250 ms-long epochs (-15 to 235 ms), including a 15 ms baseline. The selection of the frequency range of 60 to 2000 Hz was chosen on the lower end to remove motion artifacts, alpha brain activity, and powerline noise. Segments were selected to increase the signal-to-noise ratio by averaging, and chosen

such that they had the maximal length without overlapping the preceding or the subsequent stimulus. The upper limit resulted from down sampling that was implemented to speed up the analysis. We expected signals in the frequency range below 400 Hz (which we confirmed to the end). However, in order to enhance the temporal resolution, we extended the highest limit to the downsized sample rate of 5000 Hz.

We used two stimulus polarities to minimize the influence of cochlear microphonics on the observed responses. For further analyses, we decided to average the brain responses to both stimulus polarities. Averaging over polarities could in principle decrease amplitudes of the neuroelectric response to phoneme stimuli that contained more energy below the human PLL (<1.5 kHz), especially when analyzing responses at frequencies in the formant region. However, since we were looking at responses in frequency bands of the fundamental frequency of the speaker (~116 Hz) rather than at direct responses to the formants, we expected similar responses for both stimulus polarities and thus no extinction when averaging.

Finally, all epoch pairs with variance outside a region of three standard deviations around the mean epoch variance were rejected. Discarding trials with strong activity is a robust way of cleaning the data from artefacts, since head movements, eye movements, eye blinks, and swallowing artifacts were much larger than the brain signals.

Subsequently, the evoked signal's continuous wavelet transform **(CWT),** performed to obtain time-frequency information of the neuroelectric activity, was computed using the MatLab internal CWT function.

2.9. Phase coherence

Two types of neural activity contributed simultaneously to the observable response: stimulus-evoked and stimulus-independent signals. We focus exclusively on stimulus-evoked responses. To maximize the signal-to-noise ratio in the neural response, it was crucial to determine which time-frequency points were sufficiently dominated by the stimulus-evoked response. A data-driven approach was used to identify time-frequency points that had significantly larger phase coherences than could be expected by chance. Delays and amplitudes of the phase-locked brain activity were obtained. For identifying the evoked response components of the EEG signal, we measured the phase coherence compiled for signals across different subjects by averaging the phase vector across all subjects. The averaged phase vector served as the reference to which individual phase delays were compared. A time-frequency point was regarded as coherent if the length of the observed phase coherence was more than 3.5 times larger than the chance level.

2.10. Amplitude and delay of phoneme evoked responses

Based on the phase-coherent points, two measures were extracted from the phoneme-evoked responses of each individual subject: phoneme-evoked delay **(PED)** and phoneme-evoked amplitude **(PEA)**. PEA is the non-phase coherent average of all coherent time-frequency points computed as the average of the absolute values of the CWT coefficients. PED was determined as the average individual delay of all time-frequency points. Each individual time-frequency point delay is the phase difference between a subject's evoked phase and the phase of the averaged recording across all subjects, divided by the angular frequency.

Data for the phoneme induced stimuli were analyzed according to their formant information (Method 2.5) above or below the PLL as follows: Below PLL: (i) /du/-/bu/, or (ii) /du/-/bu/-/o/ (difference in F1, F2; no difference in F3, F4); above PLL: (i) /di/-/bi/, or (ii) /di/-/bi/-/y/ (no difference in F1; difference in F2, F3, F4).

2.11. Pure-tone normalized OLSA threshold (PNOT)

To identify factors that affected the OLSA threshold independent of PTT, OLSA thresholds were corrected by PTT, performing multivariate regression (MatLab 2021b) between OLSA thresholds and the first five principle components **(PCs)** (singular value decomposition, MatLab 2021b) of all PTTs, as previously described (Schirmer et al., 2024). **In brief:** OLSA thresholds were quantitatively normalized for PTT of all available frequencies [0.125 - 16 kHz] by performing multivariate regression between all three OLSA thresholds (broadband, low-pass, and high-pass) and the first five PCs (Matlab Version 2021b) of all audiometric thresholds; the number of components was limited to five to avoid over-fitting. The principle component analyses (PCAs) were performed by employing a singular-value decomposition algorithm that enabled capturing 93% of the variations in audiometric thresholds, as described in Schirmer et al. (2024). OLSA predictors of a single subject were used to evaluate the regression model, resulting in a predicted OLSA value for each subject, as described (Schirmer et al., 2024). To evaluate the residual speech-comprehension performance, three OLSA threshold predictions were performed from the three measured OLSA thresholds that were averaged, and referred to three equally-sized PNOT groups with either "good", "standard", or "poor" speech performance. We verified that this data-driven approach resulted in groups with matched average pure-tone audiometry thresholds within ± 4.5 dB

2.12. Statistical evaluation

The differences between phoneme responses were assessed by z-transformation, with a standard deviation derived by 30,000 permutations of random phoneme pairs, ensuring that all subjects were present in both surrogate phoneme responses. This ensured that the reported difference was caused by the different stimuli, and not by inter-individual differences. To evaluate correlations, an F-statistic against a constant model was used. The difference between the PNOT groups was examined using a Kruskal Wallis test. Only after providing evidence for group differences by these non-parametric tests, regression was performed for parameters with significant group differences. Correlation significance tests were conducted for both PEA and PED, and their relationships to audiometric markers (OLSA in quiet, OLSA ipsi, PNOT in quiet, PNOT ipsi, PTA4, PTA-EHF) and age were examined. To investigate how much variance in speech recognition thresholds could be explained by the speech-processing mechanisms TFS and TENV, we applied the same post-hoc permutation analysis of variance used by Schirmer et al. (2024), using 30,000 permutations and as described in Erdfelder et al. (1996).

2.13. Posthoc analysis of OLSA variance based on a linear mixed model

Analysis of variance for speech perception beyond pure-tone thresholds was performed in the same manner as in the companion paper (Schirmer et al., 2024). The least-square multivariate linear fitting of the five PCs derived from pure-tone thresholds and either PED or PEA was tested for its contribution to total speech-comprehension variance. To ensure uniqueness of the multivariate linear model, we first removed all linear correlations between the five PCs, PEA, and PED. This can be understood as removing the influence of PTT on PED and PEA. An inherent risk of the increase of dimensions of the regression model is the possibility of over-fitting. To eliminate this effect, we compared the observed increase in total explainable variance in the observed 6-dimensional model to the total explainable variance of 10,000 pseudo-models in which we randomly shuffled the additional observable before fitting the model. This gave us a reliable estimate of the gain in the explained variance based on chance. The results are presented as bargraphs showing the percentage of variance that could be attributed to each of the observables. To better illustrate the magnitude of the effect, we computed the standard deviation attributable to the new observable. However, during this computation, we assumed a normal distribution of the five PCs and the additional tested observable, an assumption that is not required for the statistical evaluation based on permutation analysis.

3. Results

We aimed to analyze the central-neural correlates of differences in phoneme processing dependent and independent of age and hearing threshold in quiet and in ipsilateral noise and compare the findings to altered cochlear components previously observed in the same cohort group (Schirmer et al., 2024). Towards this aim, in a total of 80 subjects (see method 2.2 and Tab.S1), speech-EEG was performed by recording the neuroelectric response during binaural passive listening to phonemes that relied either on TFS or on TENV coding. In all of these subjects, pure-tone audiometry spanning the frequency range between 0.5 and 16 kHz, ABR, pDPOAEs, ASSR, as well as speech audiometry in the form of the German matrix language test (OLSA) had previously been carried out (Schirmer et al., 2024). In a first approach phase-coherent evoked neural responses to all six phonemes (two vowels /o/, /y/, and four phonemes with fricatives /du/, /di/, /bu/, /bi/) were extracted from the CWT of continuous EEG recordings (Fig. S1A-E). In a second approach, to evaluate differences among pairs of phonemes, we computed the z-transform of the evoked-response differences using a permutation analysis for estimating the variance at each time-frequency point (Fig. S1F-J).

To illustrate the time courses of the different responses, we averaged z-scores across all frequencies above 60 Hz. By using permutation of phoneme labels within each subject, the L2-norm differences between the time-frequency-points of different phoneme pairs were statistically evaluated against randomness. The differences were largest for the /o/-/y/ contrast (Fig. S1A, $p(/o/-/y/) < 0.001$ (***)), slightly smaller for phonemes with the same fricative, but different vowels (Fig. S1B, $p(/bu/-/bi/) < 0.001$ (***), Fig. S1C, $p(/du/-/di/) < 0.001$ (***)), and even smaller for phoneme pairs with different fricatives, but identical vowels (Fig. S1D, $p(/bu/-/du/) = 0.0174$ (*), Fig. S1E, $p(/bi/-/di/) = 0.1628$ (n.s.)). Notably, the response difference persisted beyond differences in phoneme stimuli (Fig. S1D). Moreover, 1/f corrected time-frequency amplitude plots of phoneme stimulus differences underscored the expected frequency power spectra of phoneme-induced peak amplitudes, being either below PLL (< 1500 Hz, Fig. S1I), above PLL (> 1500 Hz, Fig. S1J), or both (Fig. S1F,G,H).

The time courses of the averaged z-scores of phoneme-evoked responses (Fig. S1A-E) were achieved by averaging z-scores for all frequencies between 60 and 2000 Hz, for reasons explained under method (2.7). Thus, the observations in the findings with CWTs that vowel differences below the PLL (< 1500 Hz), caused more significant neural responses than differences in fricatives could be confirmed (compare Fig. S1A-C for only vowel differences and Fig. S1D, E for differences only in fricatives). When comparing these neural response differences to differences of the phoneme stimuli used, we observed that the response difference for /du/-/bu/ persisted in time beyond the difference between the respective phoneme stimuli (compare Fig. S1D, I).

When averaged evoked EEG responses from both mastoids of all 80 participants during passive phoneme listening were presented in time-frequency space, no difference between neural phoneme responses between left- and right-sided recordings from mastoids were observed (not shown). Neural responses with more information below the PLL showed transient responses with relatively short latency (< 100 ms, Fig. 1A-C). The responses to phonemes above the PLL, however, tended to be more sustained and, in the case of syllables with a fricative, tended to have longer latencies (> 50 ms, Fig. 1E-G). All neuroelectric phoneme responses clustered around the fundamental of the speaker's vocal fold (116 Hz) and, to a lesser degree, around its first harmonic (232 Hz, Fig. 1, arrows on y-axes of A and E). Since in these frequencies the effect

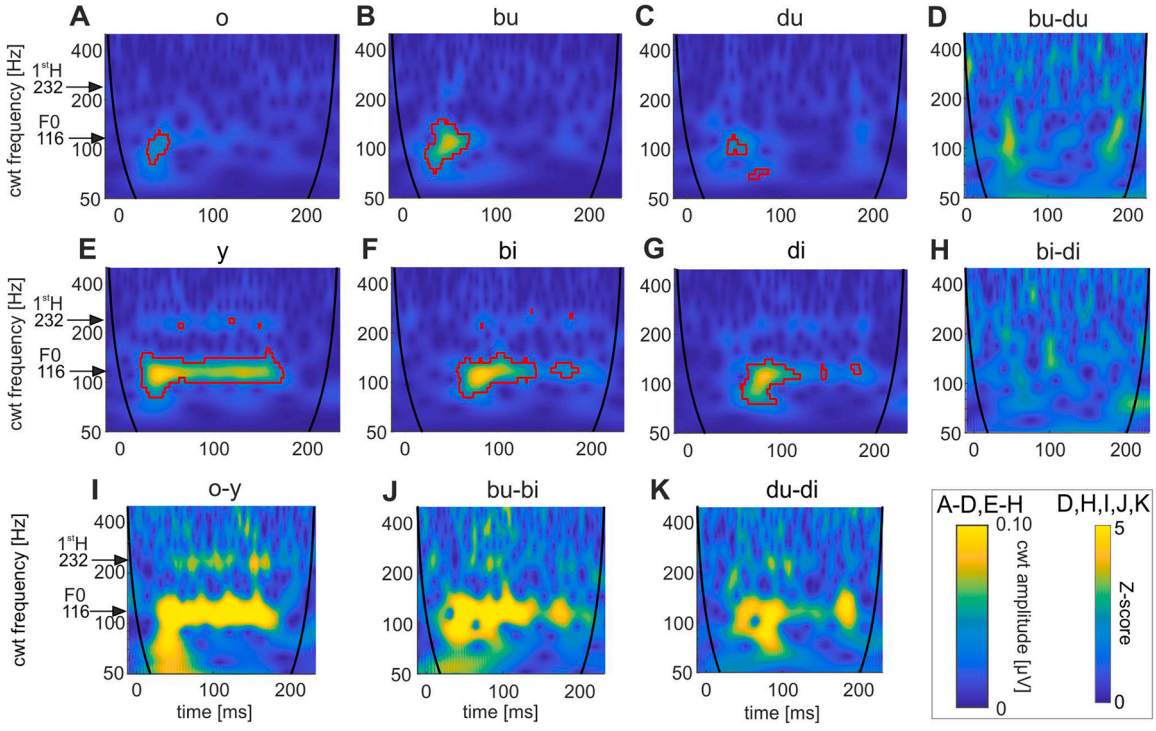

Fig. 1. Evoked EEG response in time-frequency space based on continuous wavelet transform (CWT). Averaged evoked EEG responses of all 80 participants from both mastoids during passive listening, presented in time-frequency space based on CWT. Black lines mark the cone of influence, indicating where edge effects (outside the lines in left and right lower corners) might occur in the CWT. Each plot **A-C** and **E-G** represents the response to a single phoneme, which was presented in pseudo-random order for a total of 3456 repetitions and two different electrical polarities (areas surrounded by red lines). The x-axis represents time for each stimulus epoch in milliseconds; the y-axis shows frequency in Hz on a log scale; the z-axis represents response amplitude in µV as pseudo-color; see left color scale. The red outline highlights all phase-coherent, time-frequency points (> 3.5 chance level). **D, H, I, J, K:** time-frequency plots of response differences for five different phoneme combinations (see labels), x- and y-axis same conventions as in **A-C/E-G**, z-axis: z-values as pseudo-color based on permutation of phoneme identity within each subject, see right color scale.

of stimulus polarity is unimportant, averaging of the brain responses to stimuli with different polarity is justified (see Fig. S2). Note that the fricative /d/ evoked a significant transient response component at slightly lower frequencies (~60–70 Hz, see Fig. 1C,G). Phonemes containing the vowels /i/ and /y/, i. e. that have less information below PLL (Fig. 1E-G), evoked more phase-locked time-frequency points/pixel compared to phonemes containing the vowels /o/ or /u/, i. e. that have more information below PLL (Fig. 1A-C).

As syllable discrimination is a prerequisite for speech recognition (Guilleminot et al., 2023), we examined whether the electrical brain signals carry information about differences in the tested phonemes (Fig. 1D,H, I-K). The differences in neural phoneme responses were larger when the vowels differed (Fig. 1I-K; all p-values < 1/30,000), while they were smaller when only fricatives differed and vowels encoded below PLL (Fig. 1D, p(/bu/-/du/) = 0.0174). They were non-significant when vowels exhibited less information below the PLL (Fig. 1H, p(/bi/-/di/) = 0.1628).

As individual phoneme pairs with a fricative contrast and similar vowels evoked weaker responses than those with vowel contrasts and similar fricatives (as measured by the fewer coherent time-frequency-points; compare, e.g., Fig. 1B,D or Fig. 1B,J), vowel contrasts may either enhance, or fricative contrasts may reduce, entrained neural responses.

To verify that the - on average - lower neural phoneme-response amplitude with more information below the PLL is not merely the result of averaging responses to different stimulus polarities, we averaged the polarities (see methods) independently (Fig. S2 B,E and C,F) and, for the vowels /o/ and /y/, found no difference (compare Fig. S2 B, C and E,F). Furthermore, we observed no polarity effects for any of the syllables (data not shown). This indicates that the increased number of time-frequency-points between phonemes below and above PLL cannot be explained by the averaging of the two different stimuli, and therefore cannot be attributed to contributions of cochlear microphonics, which were suppressed by the alternating-polarity paradigm used. Accordingly, all subsequent analyses were performed using the average responses to both stimulus polarities.

In conclusion, the phoneme-evoked EEG responses during passive listening to /y/, /bi/, /di/ that contain information above PLL, that relies on TENV coding, are stronger than to /o/, /bu/, /du/ with more information below PLL, that relies on TFS coding.

3.1. *Phoneme-evoked EEG amplitude and delay correlate differentially with age, speech recognition, and hearing thresholds*

To investigate whether these two basic response parameters showed significant correlation with either speech recognition thresholds as identified in young, middle-aged, and older subjects by OLSA in quiet or ipsilateral noise, with age, or thresholds within PTA4 or PTA-EHF frequency ranges (see Schirmer et al. 2024 for details), we next decomposed phoneme-evoked EEG responses into distinct electrophysiological components by quantifying PEA (Fig. 2A-E) and PED (Fig. 2F-J).

PEA correlated more significantly with OLSA in quiet (Fig. 2A, p = 0.0142), PTA4 (Fig. 2D, p = 0.004) and age (Fig. 2C, p = 0.009), than with PTA-EHF (Fig. 2E, p = 0.02). There was no correlation with OLSA in ipsilateral noise (Fig. 2B, p = 0.13). In contrast, there was a highly significant correlation between PED and age (Fig. 2H, p < 0.001), as well as between PED and PTA-EHF (Fig. 2J, p < 0.001). PED did not correlate with OLSA in quiet (Fig. 2F, p = 0.067), nor with OLSA in ipsilateral noise (Fig. 2G, p = 0.15) nor PTA4 (Fig. 2I, p = 0.18). See Table 2 for statistical details.

Taken together, these observations show that PEA is most declined over age and with elevated pure-tone thresholds and poor speech recognition in quiet. In contrast, is PED most prolonged in older participants (Fig. 2H, PED green triangles) that also exhibited the highest PTA-EHF thresholds (Fig. 2J, green triangles). In line had younger listeners shorter PED (Fig. 2H, gray circles) and lower PTA4 (Fig. 2E, gray

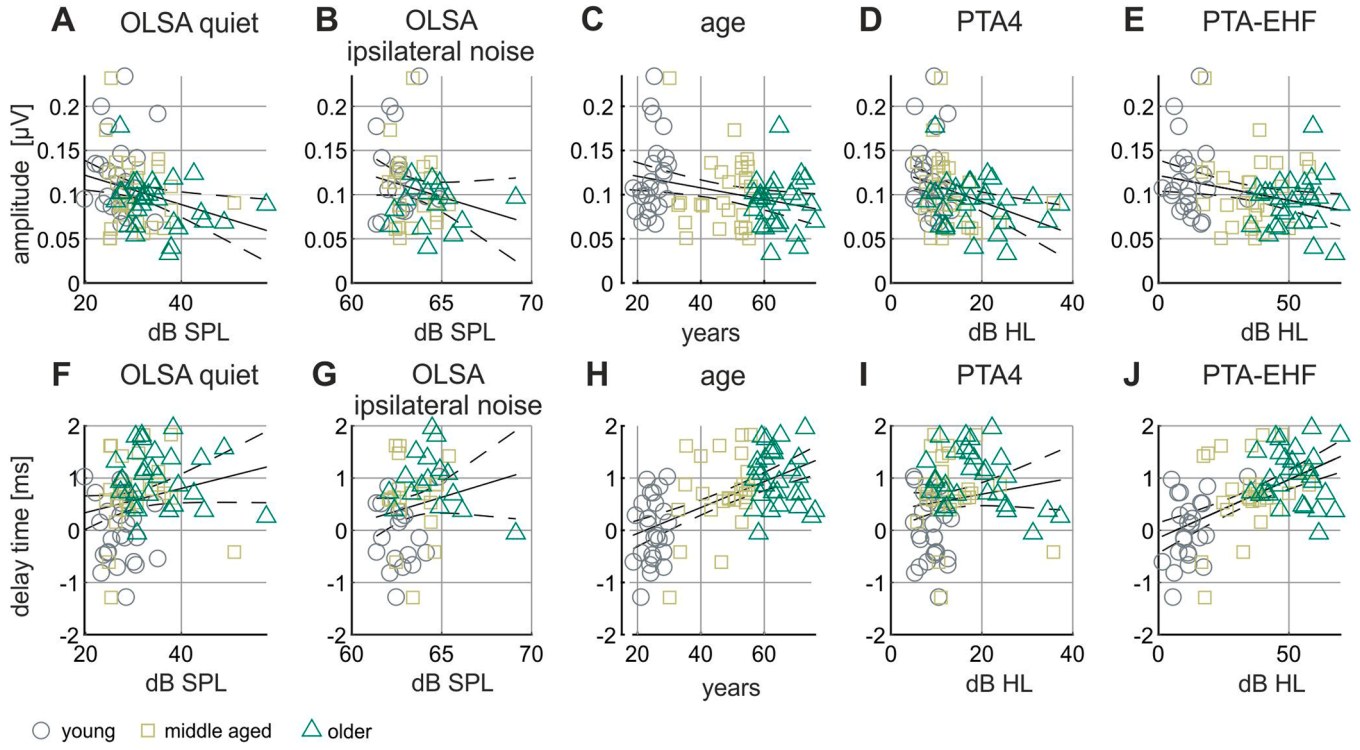

Fig. 2. Correlation of decomposed phoneme-evoked EEG responses to phoneme-induced amplitude (µV) (**A-E**) and phoneme-induced delay (ms) (**F-J**). Correlations were calculated between speech-recognition thresholds identified by OLSA in quiet (**A,F**), OLSA in ipsilateral noise (**B,G**), Age (**C,H**), pure-tone average PTA4 as clinical standard (**D,I**) and pure-tone average for extended high frequencies (PTA-EHF). Subjects are coded for their respective age groups: young (18–30y, open circle), middle-aged (30–56y, open square), older (56–76y, open triangle).

Table 2
Statistical test values for PEA and PED in the indicated categories.

PEA	OLSA quiet (Fig. 2A)	OLSA ipsilateral noise (Fig. 2B)	Age (Fig. 2C, H)	PTA4 (Fig. 2D, I)	PTA-EHF (Fig. 2E, J)
PEA - p	***0.0142***	0.13 (n.s.)	***0.009***	***0.004***	***0.02***
PEA - r	-0.2063	-0.2773	-0.2870	-0.3363	-0.2718
PEA - n	80	55	80	80	80
PED	(Fig. 2A)	(Fig. 2B)	(Fig. 2C, H)	(Fig. 2D, I)	(Fig. 2E, J)
PED - p	0.067 (n. s.)	0.15 (n.s)	***<0.001***	0.18 (n. s.)	***<0.001***
PED - r	0.1954	0.2052	0.5755	0.1495	0.5623
PED - n	80	55	80	80	80
Test type	F-test against constant value				

Table 2: p: probability value for the used statistical test; r: correlation coefficient; n: number of subjects included in the test categories. Age groups: young (18–30y, open circle), middle-aged (30–56y, open square), older (56–76y, open triangle), statistically significant values ($p < 0.05$) in bold italics.

circles), and exhibited lower PTA-EHF thresholds (Fig. 2J, gray circles). PED and PEA were linearly correlated, but are non-redundant, as the linear regression model explained only 18% of each other's variance.

3.2. Phoneme evoked EEG amplitude over age correlate with TENV and EEG delay over PTA-EHF with TFS and TENV coding

The dependency of PEA (Fig. 3A,B) and PED (Fig. 3C,D) was next broken down into the frequency content of the phoneme stimuli (Fig. 3, Table 3). PEA correlated with a negative slope restrictively for phoneme contrast-containing vowels requiring TENV coding (Fig. 3A, /di/, /bi/; Fig. 3B, F1). The correlation with PTA4 and PTA-EHF (Fig. 3 A,B, $p < 0.05$) as well as for OLSA in quiet (Fig. 3B, F1, $p < 0.05$) and OLSA in ipsilateral noise (Fig. 3A, /di/-/bi/, $p < 0.05$) was significant. See for statistic details Table 2.

In contrast, a highly significant correlation for PED was achieved with PTA-EHF for phoneme contrasts and vowels requiring TFS coding (Fig. 3C,D, F1/F2, /du/-/bu/, $p < 0.0001$), as well as with phoneme contrasts and vowels requiring TENV (Fig. 3C,D, /di/-/bi/, F1, $p < 0.0001$). The slope for PED regression was only positively correlated with a trend of significance ($p < 0.1$) with OLSA in quiet and ipsilateral noise (Fig. 3C,D) or with PTA4 for phonemes relying on TFS (Fig. 3D). See for statistic details Table 3.

Taken together, correlation of PEA decline over and with elevated pure-tone thresholds is driven by phoneme pairs having vowels requiring TENV coding, whereas the prolonged PED in older participants that also exhibited the highest PTA-EHF thresholds is driven mainly by phoneme pairs having vowels requiring TFS and TENV coding.

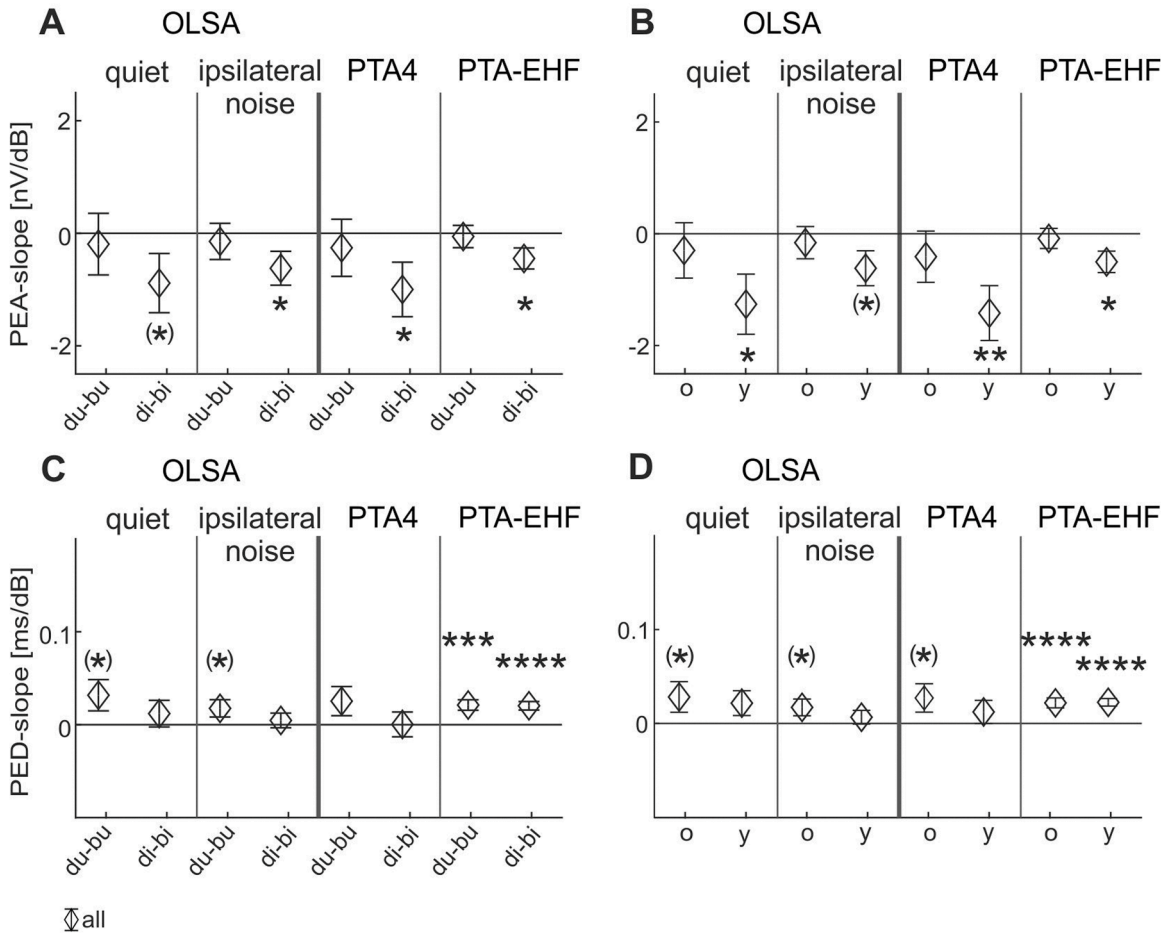

Fig. 3. Slope of regression lines of EEG phoneme evoked amplitude (PEA, A,B**) and phoneme-evoked delay (PED,** C,D**)** with speech recognition thresholds identified by OLSA in quiet, by OLSA in ipsilateral noise, pure-tone average PTA4, and pure-tone average for extended high frequencies (PTA-EHF). Correlations were performed for phonemes exposed and designed as described in method 2.5, 2.9: Below PLL: (i) /du/-/bu/, or (ii) or /du/-/bu/-/o/ (F1/F2); Above PLL: (i) /di/-/bi/ or (ii) /di/-/bi/-/y/ (F1). Error bars label the distribution of fit errors of regression lines, for all subjects (n=80, open diamonds). The OLSA ipsilateral noise slope was scaled down by a factor of 10 for better comparison. Significant correlations are indicated by * for $p < 0.05$ and (*) for $p < 0.1$. F-statistics against constant model.

Table 3
Statistics for PEA and PED depicted in Fig. 3.

	OLSA quiet (Fig.3A)		OLSA quiet (Fig.3B)		OLSA ipsi noise (Fig.3A)		OLSA ipsi noise (Fig.3B)	
	TFS	TENV	TFS	TENV	TFS	TENV	TFS	TENV
PEA	du-bu	di-bi	F1/F2	F1	du-bu	di-bi	F1/F2	F1
PEA - p	0.726	0.096	0.549	0.022	0.656	0.043	0.581	0.053
PEA - F	0.124	2.839	0.362	5.462	0.362	5.462	0.308	3.917
PEA - n	80		80		55		55	
PED	(Fig.3C)		(Fig.3D)		(Fig.3C)		(Fig.3D)	
PED - p	0.063	0.407	0.088	0.106	0.064	0.564	0.062	0.367
PED - F	3.557	0.695	2.985	2.674	3.579	0.337	3.635	0.828
PED - n	80		80		55		55	
	PTA4 (Fig.3A)		PTA4 (Fig.3B)		PTA-EHF (Fig.3A)		PTA-EHF (Fig.3B)	
	TFS	TENV	TFS	TENV	TFS	TENV	TFS	TENV
	du-bu	di-bi	F1/F2	F1	du-bu	di-bi	F1/F2	F1
PEA - p	0.612	0.043	0.374	0.005	0.771	0.019	0.643	0.010
PEA - F	0.259	4.232	0.799	8.347	0.085	5.738	0.217	6.971
PEA - n	80		80		80		80	
	(Fig.3C)		(Fig.3D)		(Fig.3C)		(Fig.3D)	
PED - p	0.106	0.981	0.077	0.327	<0.001	<0.0001	<0.0001	<0.0001
PED - F	3.112	0.273	3.212	0.973	>10	>10	>10	>10
PED - n	80		80		80		80	
Test type	F-test against constant model							

Table 3: F-test statistics values for PEA and PED in the indicated categories p: probability value; F: F-test value; n: number of subjects included in the test categories. statistically significant values ($p < 0.05$) in bold italics.

3.3. Poor speech comprehension in quiet or ipsilateral noise is linked to delayed or advanced MGB activity and TFS or TENV coding, respectively

We analyzed the linear regression of amplitude (PEA) and delay (PED) with pure-tone normalized OLSA thresholds (PNOT) for all participants, as described in methods 2.10 and in previous studies (Schirmer et al., 2024). To distinguish phoneme pairs below or above PLL, linear regressions between PNOT and PEA (Fig. 4) or PNOT and PED (Fig. 5) were calculated for grouped phonemes for PNOT in quiet (Figs. 4A, 5A) and PNOT in ipsilateral noise (Figs. 4B, 5B) as a function of the slope of PEA and PED, (Figs. 4C, 5C). The variance (%) to which PEA or PED for all phonemes or distinct phoneme formants (below and above PLL) contributed to the three different bandpass filtered OLSA in quiet and ipsilateral noise conditions is shown in Figs. 4D, G and 5D, G (red bar). To investigate the contribution of TFS and TENV coding for speech recognition in PNOT, we analyzed the variance (%) of deficits in broadband , lowpass and, highpass filtered OLSA performance that PEA and PED may explain (Figs. 4E,F,H,I; 5E,F,H,I; red bar). We observed that the phoneme-evoked EEG amplitude (PEA) did not correlate significantly with PNOT for OLSA in quiet (Fig. 4A), ipsilateral noise (Fig. 4B), or for distinct phoneme pairs requiring TENV or TFS coding (Fig. 4C,E,F,H,I). Also, the variance of deficits observed upon OLSA performance did not explain PEA for any phoneme, not for all averaged phonemes (Fig. 4D,G) neither for phonemes relying on TENV or TFS coding (Fig. 4E,F, H,I).

In contrast, PED was significantly positively correlated to PNOT in quiet (Fig. 5A), and to phonemes with vowels containing one format below PLL (Fig. 5C, F1, $p < 0.1$, /di/-/bi/, $p < 0.05$, against constant model). The positive correlation of PED with speech recognition indicates that speech recognition in quiet is linked to the delay in EEG responses. For PNOT in ipsilateral noise, the slope of the regression line falls (Fig. 5B,C), although non-significantly ($p = 0.18$, against constant model). The slope of the regression for PED with PNOT in ipsilateral noise was stronger for phonemes requiring TENV than TFS (Fig. 5C, F1, /di/-/bi/), though not reaching significance. The variance explained by PED for the three OLSA filter conditions encompasses (i) for OLSA in quiet the unfiltered "broadband" speech condition (2% of variance or 1.5 dB in speech-recognition thresholds) and low-pass filtered speech condition (3.6% of variance or 1.3 dB in speech-recognition thresholds), when driven by phonemes below PLL accounting TFS coding (Fig. 5D,E, F1/F2, low-pass). (ii) OLSA in ipsilateral noise, for which PED was predictive for high-pass filtered speech performance (Fig. 5G, 5.8%; 1.6 dB), particular by phonemes above PLL accounting TENV coding (Fig. 5I, 6.4%; 1.7 dB).

In conclusion: We found that EEG (MGB) amplitudes over age varied only with phonemes that rely on TENV coding in correlation with worse speech-recognition thresholds in quiet and elevated pure-tone thresholds (Graphical abstract A). In contrast, EEG (MGB delays) varied for both, TFS and TENV coding, particular with thresholds in extended high frequency regions (PTA-EHF) (Graphical abstract B). Finally, was MGB activity delayed independent of age and hearing loss for OLSA (low-pass filtered) in quiet and phonemes requiring TFS coding, while MGB activity was shortened (with a trend) for OLSA (high-pass filtered) in ipsilateral noise for phonemes requiring TENV coding (Graphical abstract C).

4. Discussion

We here demonstrate that the analysis of the entrained amplitudes (PEA) and delays (PED) of EEG responses measured in young, middle-aged, and older subjects to passive listening of phonemes, revealed a differential contribution of phoneme-induced amplitudes and delays to either TENV and TFS coding, respectively. Until now, speech deficits have mainly been described for challenging listening conditions under background noise that linked deficits to reduced TENV coding. We find that TFS and TENV coding deficits for listening conditions in quiet and ipsilateral noise dependent and independent of age can be reflected in reduced amplitude or delays of subcortical activity that based on studies performed in the same subject group allow an association to cochlear synaptopathy. We propose that subcortical EEG activity is a predictor of poor speech recognition, whether dependent or independent of age and hearing loss that can link synchronicity changes at stimulus onset to

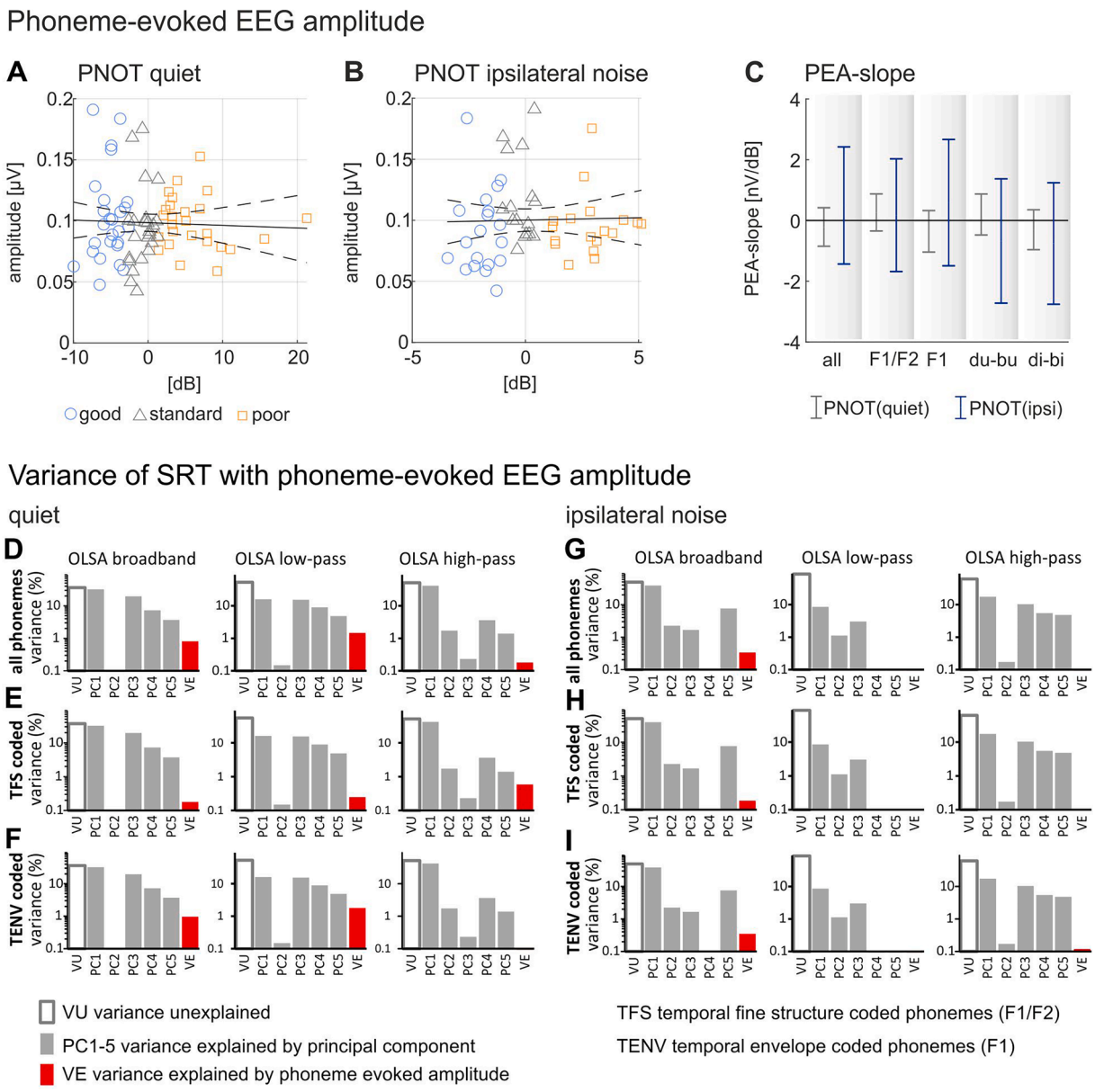

Fig. 4. No correlation between phoneme evoked amplitude (PEA) and pure-tone threshold (PTT)-normalized speech recognition threshold (PNOT) in quiet (**A**), ipsilateral noise (**B**), formant, or phoneme contrast (**C**): **A,B**: Individual speech recognition categorized into good (blue circles), standard (gray triangles), and poor (orange squares). **C**: Slopes of regression lines of PEA with PNOT for OLSA in quiet (gray lines) and OLSA in ipsilateral noise (blue lines) displayed for all phonemes (left), number of formants below PLL (F1/F2, F1), and phoneme (/du/-/bu/, /di/-/bi/). **D-F**: Variances for OLSA in quiet, **G-I:** variances for OLSA in ipsilateral speech-shaped noise. PEA was compiled for the responses to all phonemes **(D, G)**, to phonemes containing two formants and spectral contrasts below the PLL (TFS coded, /du/, /bu/, /o/) **(E, H)**, and to phonemes containing one formant below the PLL and containing spectral contrasts above the PLL (TENV coded, /di/, /bi/, /y/) **(F, I)**. Beyond the first five principal components (PC 1–5) derived from the PTT, only a non-significant small part of the variance of the speech recognition threshold in the OLSA conditions (broadband, low-pass, high-pass) was explained by the phoneme-evoked EEG amplitude (PEA).

altered neocortical activity. EEG recordings that capture subcortical auditory regions should thereby be discussed as a tool to adapt hearing aids to individual speech recognition deficits.

4.1. MGB is the putative source of phoneme-induced amplitudes and delays of EEG signals

In the present study a precise localization of the generators of speech EEG signals was not feasible at the level of dipoles, as recordings were restricted to only two channels (Picton et al., 1999) that were recorded simultaneously at both mastoids. The argument for the MGB, as the relay station between the inferior colliculus and auditory cortex, as the source of EEG recordings is based on physiological and neuroanatomical findings: (i) A fundamental frequency for phoneme stimuli of 116 Hz was, however, chosen, that is predicted to generate auditory evoked event-related potentials of subcortical origin, when modulated > 80 Hz (Herdman et al., 2002; Lu et al., 2022). (ii) In accordance with this, the recorded PEA-induced amplitudes decreased with age (see Fig. 2C), which is, however, rather inconsistent with cortical sources (Ross et al., 2020). (iii) Finally, the delay shifts of the neuroelectric responses were too large across subjects to originate from areas below the thalamus. Thus, we rather suggest that MGB is the subcortical source of the entrained neural activity measured in the present study.

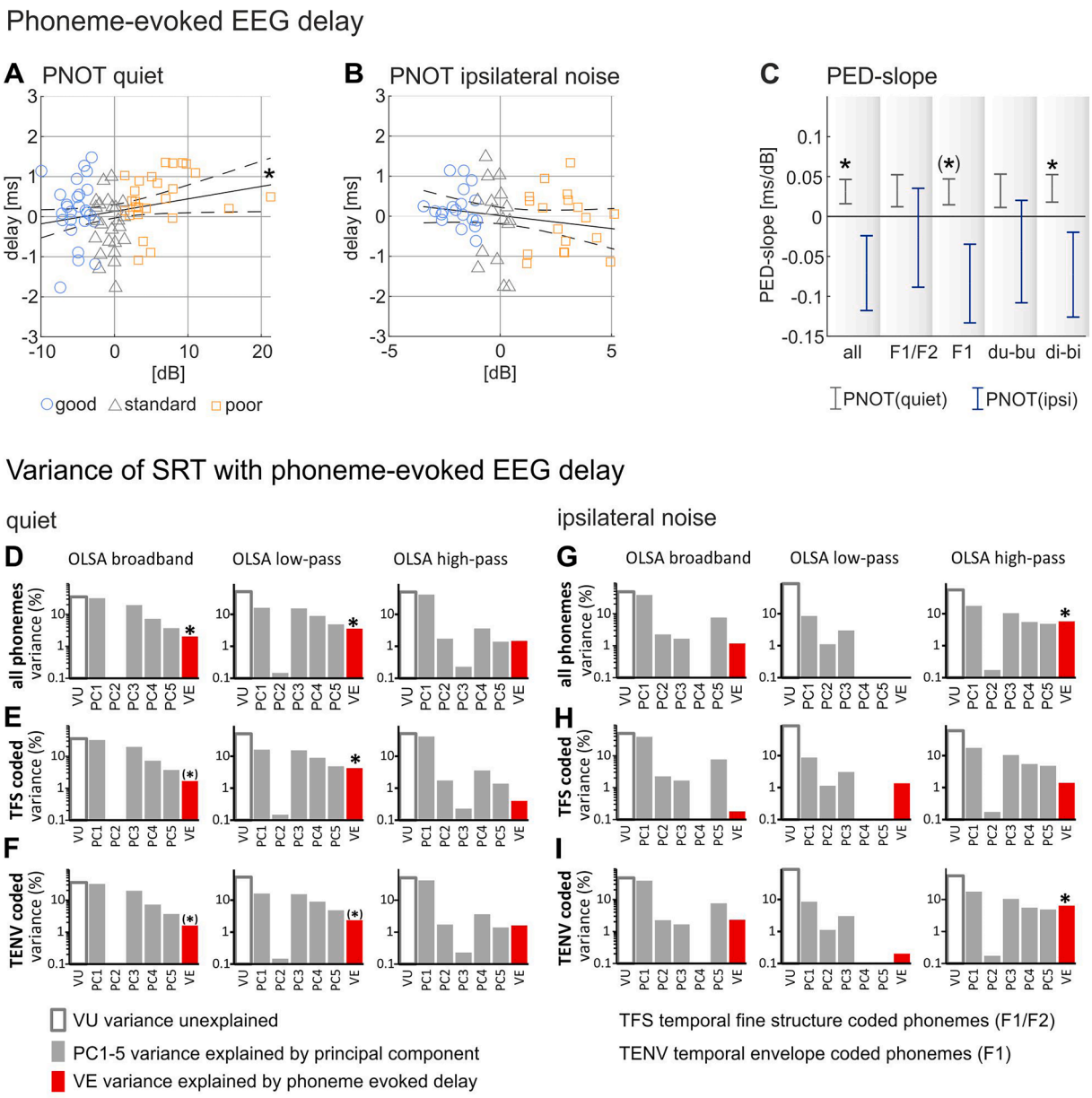

Fig. 5. Significant positive correlation of phoneme-evoked delay (PED) with pure-tone normalized speech recognition threshold (PNOT) in quiet (**A,C**) but not in ipsilateral noise (B,C), and not for phoneme stimuli with two formants below the PLL (**C**, F1/F2, /du/-/bu/). **A,B**: Individual speech recognition categorized into good (blue circles), standard (gray triangles), and poor (orange squares). **C:** Slopes of regression lines of PEA with PNOT for OLSA in quiet (gray lines) and OLSA in ipsilateral speech-shaped noise (blue lines) displayed for all phonemes (left), number of formants below PLL (F1/F2, F1), and phoneme (/du/-/bu/, /di/-/bi/). **D-F,G-I**: Beyond the first five principle components (PC 1–5), a significant part of variance of the speech recognition threshold can be explained by the EEG delay (PED) for the broadband and low-pass OLSA condition in quiet (**D-F**) mainly for phoneme stimuli with two formants and spectral contrasts below PLL (TFS coded, **D,E**) and the OLSA condition in ipsilateral noise (**G-I**), mainly for phoneme stimuli with one formant below the PLL and spectral contrasts above the PLL (TENV coded, **G, I**). * for $p<0.05$, (*) for $p<0.1$.

4.2. PEA decline related to speech in quiet, age and PTA4 rely on TENV coding

We observed a PEA decline of up to ~50 μV in subjects with worsened speech perception identified through OLSA quiet (Fig. 2A), for age (Fig. 2C) or PTA4 threshold (Fig. 2D), and less for PTA-EHF and not for OLSA in noise (Fig. 2B,C,E). Moreover, the decline of PEA only correlated with TENV coding (Fig. 3A,B). Particularly, the finding of a PEA decline correlating with TENV-based influences on OLSA quiet but not OLSA in ipsilateral noise accomplishes previous findings that (i) suggested particular monoaural processing of TENV speech information hardly to be affected by hearing loss (Moore, 2016, 2021) and that (ii) declined TENV coding is particularly relevant for speech content in noise (Carney, 2018; Elhilali et al., 2009; Johannesen et al., 2016; Le Prell, 2019; Marrufo-Perez and Lopez-Poveda, 2022; Moore, 2016, 2021; Shannon et al., 1995; Snell and Frisina, 2000). The finding suggests that TENV coding deficits for listening conditions of speech in quiet could theoretically be identified through reduced MGB EEG amplitudes. Previous findings link cochlear synaptopathy, specifically of low-spontaneous-rate **(SR),** high-threshold auditory-nerve fibers, in animals (Parthasarathy et al., 2019; Parthasarathy et al., 2014; Shaheen et al., 2015) and humans (Bharadwaj et al., 2014; Marcher-Rorsted et al., 2022; Mepani et al., 2021) with compromised TENV coding during speech in noise (Asokan et al., 2018; Bakay et al., 2018; Chambers et al.,

2016; Hesse et al., 2016; Liberman and Kujawa, 2017; Monaghan et al., 2020; Moore, 2021; Plack et al., 2014; Wu et al., 2019).

In conclusion, the findings in the present study suggest that TENV coding deficits over age are not only relevant for speech in noise but also for speech in quiet, therefore, likely require high-SR auditory-nerve fibers, as also previously suggested (Huet et al., 2022, 2018). This feature is possibly reflected in declining MGB amplitudes over age (Graphical abstract A).

4.3. PED delay related to age and PTA-EHF is linked to TFS and TENV coding

We here observed a delay of PED that correlated best with age (Fig. 2H) and PTA-EHF elevation (Fig. 2J). The delay of PED in older subjects was correlated with an up to 1.3 ms shift during aging (Fig. 2H). Particularly, the PED shifts correlating highly significantly to PTA-EHF affected TFS- and TENV-requiring speech signals (Fig. 3C,D), and were most prolonged in older participants with poor speech recognition. These participants also exhibited the highest PTA-EHF thresholds (Fig. 2, Schirmer et al., 2024). Various studies describe an influence of PTA-EHF on speech comprehension (Hunter et al., 2020; Motlagh Zadeh et al., 2019), although the mechanism remains obscure. The finding that the MGB delay leads to highly significant deficits in speech intelligibility of both TFS and TENV coding in subjects with increased PTA-EHF is basically alarming, because these extended frequency ranges are typically not tested in the clinic to diagnose speech disorders. It is, however, well established that, e.g., a deterioration of binaural TFS processing with increasing age starts relatively early in life and thereby without doubt starts with PTA-EHF threshold elevations (Hunter et al., 2020; Moore, 2016, 2021). Also, beside the role of PTTs and TENV coding for speech recognition, particularly in noise (see above), the predicted role of TFS for localization of competing sound sources (Smith et al., 2002; Verschooten et al., 2019) was shown to decrease with age (Meister et al., 2020; Moore, 2021).

How can TENV and TFS coding deficits, here reflected in MGB delays, possibly be linked to threshold elevation, in particular PTA-EHF? We should first refer to the audiometric data carried out in the same subjects as in this study. Thus, reduced peak amplitudes of supra-threshold ABR wave I of about 0.100 μV was described in the same middle-age and old subjects (Fig. 3; Schirmer et al., 2024) as used here (Figs. 2, 3). The middle-aged subjects could centrally restore ABR wave IV/VI amplitudes (Schirmer et al., 2024). However, the same group of elderly, here described to exhibit a PED of ~1.3 ms (Fig. 2H) and profound PTA-EHF elevation (Fig. 2J), was previously shown to exhibit persistently reduced low-frequency evoked ABR wave IV/VI amplitudes of ~0.050 μV and a significant delay of ABR wave III-IV of up to ~200 ms (Schirmer et al., 2024). As suprathreshold ABR waves in (Schirmer et al., 2024) were induced by acoustic signals below 8 kHz, the elevated PTA-EHF was suggested to particularly impact the amplitude of synchronous activity at lower frequencies (Schirmer et al., 2024). In line with this, elevated PTA-EHF have only previously been suggested to impact synchronized auditory responses in lower frequency ranges (Marcher-Rorsted et al., 2022). In this context, it is important to note that the low-frequency regions are those frequency components that are most crucial for high speech-sound quality (Moore, 2008). This means that the delayed PED observed here to be correlated strongly with age and PTA-EHF for TFS-dependent stimuli (Figs. 2H,J; 3C,D), may have influenced the synchronized auditory responses to lower frequencies, and possibly thereby the PED amplitude and TENV coding. In other words, we suggest that the correlation of PEA to OLSA quiet (Fig. 2A), age (Fig. 2C) and PTA-4 (Fig. 2D) for TENV-dependent stimuli, and the correlation of PED to PTA-EHF for TFS- and TENV-dependent stimuli may possibly be dependent events. Note in this context that synchronized responses required for phase locking and peak amplitudes are critically dependent on high-SR auditory-nerve fibers: Thus, peak amplitudes of ABR waves along the ascending pathway are defined through the discharge rate and synchronicity of summed auditory nerve responses (Johnson et al., 2008; Lu et al., 2022). High-SR auditory-nerve fibers are critical for the precision of phase locking (Bourien et al., 2014; Huet et al., 2022; Huet et al., 2018), they are prevalent in both high-frequency (Marcher-Rorsted et al., 2022), and low-frequency regions (Bourien et al., 2014; Gleich et al., 2016; Harris et al., 2008; He et al., 2007; Huet et al., 2022; Huet et al., 2018), and they define perception thresholds (Heil et al., 2008; Meddis, 2006; Rhode and Smith, 1986; Zohar et al., 2011). In contrast, low-SR high threshold auditory-nerve fibers hardly contribute to the synchronicity of auditory-nerve fibers (Huet et al., 2019).

We conclusively postulate that the influence of PTA-EHF on MGB delay and its correlation to TFS and TENV coding, is due less to the loss of outer hair cells than to the loss of high-SR auditory-nerve fibers. High-SR auditory-nerve fibers could, via a transformation process to low-firing auditory fibers (Suthakar and Liberman, 2021), impair the precision of synchronous response behavior and thus TFS coding in quiet (Graphical summary B). For the same reason, the reduction of the amplitude in lower frequency regions with increased PTA-EHF and likely TENV coding deficits in noise (Schirmer et al., 2024) could be explained (Graphical summary B). In line with this, recent studies, carried out in young and old subjects with near-normal thresholds, evidenced phase locking deficits and delays in subcortical regions at the level of ABR wave V to be somewhat independent of outer hair cell function (Anderson et al., 2021). The previously observed link between EHF thresholds and putative cochlear synaptopathy in subjects with poorer speech discrimination scores (Lai and Bidelman, 2022; Liberman and Dodds, 1984; Liberman et al., 2016) should be reconsidered in this context.

4.4. Poor speech comprehension in quiet or ipsilateral noise is linked to delayed or a trend of shortened MGB activity and TFS or TENV coding, respectively

We found that the PEA amplitude mainly remained unaffected when poor speech-comprehension capabilities independent of PTT in so called "poor performers" were examined in PNOT quiet or PNOT ipsilateral noise groups (Fig. 4). Neither the slope of PEA (Fig. 4C) nor the variance of OLSA speech-recognition threshold in quiet or ipsilateral noise could be linked to any change in PEA in pure-tone normalized OLSA speech-recognition threshold differences (Fig. 4D-I). Interestingly, in the identical groups neither early nor late supra-threshold ABR wave differences in poor or good speech-recognition thresholds beyond PTT were observed (Schirmer et al., 2024).

In contrast, the latency shift in the MGB observed here, as reflected through PED in groups with poor PNOT, was significantly prolonged in quiet (~1ms), and shortened in ipsilateral noise (~0.5 ms, Fig. 5). In these subjects, the prolonged PED in PNOT quiet appeared to mainly explain the speech-recognition threshold variance for low-pass filtered speech conditions (Fig. 5D), where frequencies above PLL are filtered out. Thus, this mainly reflected TFS coding, confirmed through significant speech-recognition thresholds variance for F1/F2 phoneme coding (Fig. 5F) (Graphical summary C). On the other hand, the shortened delay in PNOT in ipsilateral noise explained the speech-recognition threshold variances for performance in high-pass filtered speech conditions, where frequencies below PLL were filtered out, thus here mainly reflecting TENV (Fig. 5G) and speech-recognition threshold variance for F1 phoneme coding (Fig. 5H), respectively (Graphical summary C). Here again, in the same group in a previous study, the relationship to contributions of neural-processing differences at stimulus onset are obvious: In the poor PNOT quiet groups (Schirmer et al., 2024), where in the present study a significant PED delay (~1 ms) was noticed, previous findings observed poorer cochlear amplifier performance and delayed ABR, reaching ~0.2 ms (Schirmer et al., 2024;Graphical summary C).

Additionally, in the same poor PNOT ipsilateral noise group (Schirmer et al., 2024), in which in the present study a PED shortening of

(~0.5 ms) was observed, previous findings indicated a stronger cochlear amplifier performance, larger ASSR response, and higher uncomfortable loudness levels (Schirmer et al., 2024). The greater cochlear amplifier performance was suggested to result from altered basilar-membrane compression that would lead to a start of the growth behavior of the basilar membrane amplitude at lower levels (Schirmer et al., 2024). Under these conditions, when noise has masked high-SR auditory-nerve fibers, uncompromised modulation contrast of the speech signal would result (Johannesen et al., 2016), a phenomenon known also as recruitment (Denes and Naunton, 1950; Kubota et al., 2019). Recruitment phenomena are not only known to be reflected in DPOAE input-output functions (Rasetshwane et al., 2013), but also loudness scaling (Abdala et al., 2021), thus providing a rational for the observed elevated uncomfortable loudness levels in these poorly-performing PNOT ipsilateral noise groups (Schirmer et al., 2024). Also, the larger ASSR responses could contribute to larger uncomfortable loudness levels, and possibly thereby to compromised TENV coding (Schirmer et al., 2024). These features could, in total, contribute to the observed variance of speech-recognition thresholds that in the present study explained PED for ipsilateral noise for OLSA (high-pass filtered) and F1 (Fig. 5G,H) (Graphical summary C). For the same reasons as discussed above (4.3), the delay or shortening of the PED shift linked to PNOT quiet or ipsilateral groups, and TFS or TENV coding, respectively, indicates an association to a high-SR auditory-nerve fiber processing deficiency: Here, cochlear amplifier performance and high-SR auditory-nerve fiber processing could influence proper pre- and postsynaptic sound transmission towards fast auditory processing as suggested (Schirmer et al., 2024), and thereby influence neural timing up to MGB regions (Graphical summary C).

4.5. *Possible sources of PEA or PED changes that were observed to be dependent or independent of PTT*

We observed a maximal delay prolongation with age of about 1.3 ms (Fig. 2). This magnitude is consistent with delay shifts described on the level of the brainstem in children (Johnson et al., 2008), but may not be directly explained by the 400 μs age-induced ABR wave V shift in the inferior colliculus (Konrad-Martin et al., 2012), or the longest ABR wave V latency difference in participants with elevated PTA-EHF, reaching ~200 μs (Schirmer et al., 2024). The time delay of ~1.3 ms over age (Fig. 2 C) and ~1 ms in subjects with poor speech recognition independent of PTT (PNOT) (Fig. 5), may, however, result from a cumulative delay of auditory evoked responses. These start as early auditory response delays identified in neural responses generated as ABR wave I around ~1 ms, to ABR wave V/VI at ~6–8 ms, that were prolonged due to age by ~0.1 ms to 0.2 ms, respectively (Schirmer et al., 2024). Delay increases of that size could, in general, be caused by impairments of short-latency, high-frequency cochlear generators, but after normalizing for PTT shifts by using PNOT, the causes have to be sought further upstream.

It is thereby conceivable that a constant delay of the entire phoneme-evoked response up to auditory subcortical MGB regions is the result of a slowing of the response by a delay that accumulates between early (typically 0–10 ms), middle (10–50 ms) and late (50–400 ms) components up to the auditory cortex (Gilmore, 1995; Pfefferbaum et al., 1980). The observed delay in the present study could also theoretically contribute to the observed cortical delay of up to 50 ms – previously reported to occur in aging subjects (Price et al., 2017). This would complement previous findings that suggest that presbycusis at the level of the auditory cortex is a combination of central aging and an altered periphery (Profant et al., 2014). In further support of a specific form of cochlear synaptopathy to influence TFS based speech coding, came from observations of reduced and delayed compound auditory-nerve activity to diminish phase-locked phoneme activity during compressed speech in older subjects (Harris et al., 2021). This peripheral auditory nerve deficit was related to a decline in the density of cranial nerve VIII, as measured through high-resolution T2-weighted structural imaging (Harris et al., 2021). Both reduced and delayed compound auditory-nerve activity, and the decline in the density of the eighth nerve were discussed in the context of poorer synchrony and predicted poorer speech recognition in noise in younger and older adults (Harris et al., 2021). That study argued for a contribution of diminished low-SR auditory-nerve fiber processing to the fiber-diameter change that was linked to speech-processing deficits (Harris et al., 2021). The results discussed above (Section 4.3 and Schirmer et al., 2024), may rather speak for a contribution of high-SR low-threshold auditory-nerve fibers. Accordingly, particularly the high-SR auditory fibers, through their larger amount of myelination (Liberman, 1980; Merchan-Perez and Liberman, 1996), and higher levels of neurofilaments (Elder et al., 1998; Hoffman et al., 1987; Marszalek et al., 1996; Sun et al., 2018; Xu et al., 1996) contribute predominantly to larger fiber diameter. Furthermore, previous findings that analyzed peripheral contributors to phase-locking strength in young and old subjects with near normal thresholds using syllable induced frequency-following responses found strong evidence of phase-locking deficits evident upon ABR wave V delays that were somewhat independent of outer hair cell function (Anderson et al., 2021). The reduced phase-locking value, linked to reduced compound auditory-nerve activity peak amplitude (Anderson et al., 2021; Harris et al., 2021), and reduced auditory-nerve fiber thickness in aging subjects with speech recognition deficits in noise (Harris et al., 2021), may thus be part of similar observations of the reduced and delayed ABR wave (Schirmer et al., 2024) and PEA and PED changes (present study) both over age and independent of age.

This view may contradict previous findings of delays of auditory evoked responses in auditory cortical fields over age, and that were assumed to be mainly based on white-and grey matter deficits in cortical regions (Price et al., 2017). In those latter studies, contributions of auditory processing deficits from the periphery could, however, not be entirely excluded (Price et al., 2017).

In conclusion, our findings describe deficits in synchronous neuronal activity from the cochlear periphery to the auditory cortex. In view of this, the PEA and PED changes could also be part of the reduced amplitudes and delayed cortical responses observed over age in numerous studies, mostly described as changes in evoked potentials, that reflect likely intracortical network activity over age (Bertoli et al., 2002; Blackwood and Muir, 1990; Puttabasappa et al., 2017; Shukla et al., 2000; Zhao et al., 2024). We suggest reconsidering this previous information on age-dependent evoked potential changes in the context of a possibly subcortical source, as suggested in the current study. We cannot exclude corticothalamic feedback (see for a review: Antunes and Malmierca, 2021) to have influenced MGB activity, as corticofugal feedback can shape speech representations up to the brainstem level (Lai and Bidelman, 2022). We however here suggest that altered phoneme-induced auditory fine-structure responses (Schirmer et al., 2024), and phoneme-induced subcortical, likely MGB activity changes, here analyzed through EEG recording, may be related events. A relation of a predicted inner hair cell synapse weakness (Schirmer et al., 2024) that through cumulative changes in the ascending circuits leads to the observed amplitude reduction and delay in the MGB (present study) could be further strengthened through analyzing cortical (e.g. MEG) responses in the same subject, a process that is ongoing.

Outlook for future research: The findings urgently emphasize the need for improved routine clinical techniques to diagnose sound processing at the stimulus onset, in combination with, e.g., EEG recording and measurements of PTA-EHF. Here, specific TFS- and TENV-based stimuli using combined techniques that include speech comprehension analysis and stimulus onset detection techniques (Schirmer et al., 2024) with subcortical (EEG, present study) or cortical imaging analysis (e.g. MEG, in process) could provide new challenges for improved hearing aid/CI adaptation. This is particularly valuable as, e.g., the amplitude of speech EEG responses to syllables, appearing as neural responses at middle and longer latencies (> 6 ms), are a good predictor of speech

comprehension in hearing-aid users (Shetty and Puttabasappa, 2017). On the basis of the present findings, in addition delays of speech EEG responses to syllables should be urgently reconsidered for hearing-aid studies.

Role of the funding source

This work was supported by the Deutsche Forschungsgemeinschaft DFG KN 316/13–1, DFG RU 713/6–1, KL 1093/12–1; ERA-NET NEURON JTC 2020: BMBF 01EW2102 CoSySpeech and FWO G0H6420N; IZKF Promotionskolleg of the Faculty of Medicine, University Hospital of Tübingen. VICI Grant (Grant No. 918–17–603), Netherlands Organization for Scientific Research (NWO) and the Netherlands Organization for Health Research and Development (ZonMw). The funding source(s) are not involved in study design; data collection, analysis and interpretation, or writing and submitting the manuscript.

CRediT authorship contribution statement

Konrad Dapper: Writing – review & editing, Software, Investigation, Formal analysis, Data curation. **Stephan M. Wolpert:** Writing – review & editing, Supervision, Methodology. **Jakob Schirmer:** Formal analysis, Data curation. **Stefan Fink:** Writing – review & editing, Formal analysis. **Etienne Gaudrain:** Software, Methodology. **Deniz Başkent:** Software, Methodology. **Wibke Singer:** Writing – review & editing, Visualization, Formal analysis. **Sarah Verhulst:** Software, Methodology. **Christoph Braun:** Writing – review & editing, Supervision, Methodology. **Ernst Dalhoff:** Methodology. **Lukas Rüttiger:** Writing – review & editing, Writing – original draft, Supervision, Software, Project administration, Methodology. **Matthias H.J. Munk:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology. **Marlies Knipper:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors have no conflicts of interest to declare. All co-authors have seen and agreed with the manuscript's contents, and there is no financial interest to report.

Acknowledgments

We thank the audiometrists of the HNO-Klinik Tübingen for their expertise in recording audiograms and stels-ol (desmosa@gmx.de) for English language services.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2024.120958.

Data availability

The raw data supporting the conclusions of this article will be made available by the authors on request.

References

Abdala, C., Ortmann, A.J., Guardia, Y.C., 2021. Weakened cochlear nonlinearity during human aging and perceptual correlates. Ear Hear. 42, 832–845.
Allison, T., Hume, A.L., Wood, C.C., Goff, W.R., 1984. Developmental and aging changes in somatosensory, auditory and visual evoked potentials. Electroencephalogr. Clin. Neurophysiol. 58, 14–24.
Anderson, S., Bieber, R., Schloss, A., 2021. Peripheral deficits and phase-locking declines in aging adults. Hear Res. 403, 108188.
Antunes, F.M., Malmierca, M.S., 2021. Corticothalamic pathways in auditory processing: recent advances and insights from other sensory systems. Front. Neural Circ. 15, 721186.
Apoux, F., Bacon, S.P., 2004. Relative importance of temporal information in various frequency regions for consonant identification in quiet and in noise. J. Acoust. Soc. Am. 116, 1671–1680.
Asokan, M.M., Williamson, R.S., Hancock, K.E., Polley, D.B., 2018. Publisher correction: sensory overamplification in layer 5 auditory corticofugal projection neurons following cochlear nerve synaptic damage. Nat. Commun. 9, 3158.
Bajin, M.D., Dahm, V., Lin, V.Y.W., 2022. Hidden hearing loss: current concepts. Curr. Opin. Otolaryngol. Head. Neck. Surg. 30, 321–325.
Bakay, W.M.H., Anderson, L.A., Garcia-Lazaro, J.A., McAlpine, D., Schaette, R., 2018. Hidden hearing loss selectively impairs neural adaptation to loud sound environments. Nat. Commun. 9, 4298.
Bertoli, S., Smurzynski, J., Probst, R., 2002. Temporal resolution in young and elderly subjects as measured by mismatch negativity and a psychoacoustic gap detection task. Clin. Neurophysiol. 113, 396–406.
Bharadwaj, H.M., Verhulst, S., Shaheen, L., Liberman, M.C., Shinn-Cunningham, B.G., 2014. Cochlear neuropathy and the coding of supra-threshold sound. Front. Syst. Neurosci. 8, 26.
Bidelman, G.M., Carter, J.A., 2023. Continuous dynamics in behavior reveal interactions between perceptual warping in categorization and speech-in-noise perception. Front. Neurosci. 17, 1032369.
Blackwood, D.H., Muir, W.J., 1990. Cognitive brain potentials and their application. Br. J. Psychiatry Suppl. 9, 96–101.
Bourien, J., Tang, Y., Batrel, C., Huet, A., Lenoir, M., Ladrech, S., Desmadryl, G., Nouvian, R., Puel, J.L., Wang, J., 2014. Contribution of auditory nerve fibers to compound action potential of the auditory nerve. J. Neurophysiol. 112, 1025–1039.
Brand, T., Kollmeier, B., 2002. Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. J. Acoust. Soc. Am. 111, 2801–2810.
Carcagno, S., Plack, C.J., 2022. Relations between speech-reception, psychophysical temporal processing, and subcortical electrophysiological measures of auditory function in humans. Hear Res. 417, 108456.
Carney, L.H., 2018. Supra-threshold hearing and fluctuation profiles: implications for sensorineural and hidden hearing loss. J. Assoc. Res. Otolaryngol. 19, 331–352.
Chambers, A.R., Resnik, J., Yuan, Y., Whitton, J.P., Edge, A.S., Liberman, M.C., Polley, D. B., 2016. Central gain restores auditory processing following near-complete cochlear denervation. Neuron 89, 867–879.
den Besten, C.A., Monksfield, P., Bosman, A., Skarzynski, P.H., Green, K., Runge, C., Wigren, S., Blechert, J.I., Flynn, M.C., Mylanus, E.A.M., Hol, M.K.S., 2019. Audiological and clinical outcomes of a transcutaneous bone conduction hearing implant: six-month results from a multicentre study. Clin. Otolaryngol. 44, 144–157.
Denes, P., Naunton, R.F., 1950. The clinical detection of auditory recruitment. J. Laryngol. Otol. 64, 375–398.
Elder, G.A., Friedrich Jr, V.L., Kang, C., Bosco, P., Gourov, A., Tu, P.H., Zhang, B., Lee, V. M., Lazzarini, R.A., 1998. Requirement of heavy neurofilament subunit in the development of axons with large calibers. J. Cell Biol. 143, 195–205.
Elhilali, M., Ma, L., Micheyl, C., Oxenham, A.J., Shamma, S.A., 2009. Temporal coherence in the perceptual organization and cortical representation of auditory scenes. Neuron 61, 317–329.
Erdfelder, E., Faul, F., Buchner, A., 1996. GPOWER: a general power analysis program. Behav. Res. Methods Instrum. Comput. 28, 1–11.
Frisina, R.D., 2009. Age-related hearing loss: ear and brain mechanisms. Ann. N. Y. Acad. Sci. 1170, 708–717.
Fullgrabe, C., Moore, B.C., Stone, M.A., 2014. Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. Front. Aging Neurosci. 6, 347.
Garrett, M., Vasilkov, V., Mauermann, M., Wilson, J.L., Henry, K.S., Verhulst, S., 2024. Speech-in-noise intelligibility difficulties with age: the role of cochlear synaptopathy. bioRxiv., 2020.2006.2009.142950.
Gaudrain, E., D., B., Verhulst, S., 2024. CoSySpeech: Material for vowel and consonant contrasts targeting specific peripheral coding mechanisms. Zenodo.
Gilmore, R., 1995. Evoked potentials in the elderly. J. Clin. Neurophysiol. 12, 132–138.
Gleich, O., Semmler, P., Strutz, J., 2016. Behavioral auditory thresholds and loss of ribbon synapses at inner hair cells in aged gerbils. Exp. Gerontol. 84, 61–70.
Golub, J.S., Brickman, A.M., Ciarleglio, A.J., Schupf, N., Luchsinger, J.A., 2020. Association of subclinical hearing loss with cognitive performance. JAMa Otolaryngol. Head. Neck. Surg. 146, 57–67.
Gregory, S., Billings, J., Wilson, D., Livingston, G., Schilder, A.G., Costafreda, S.G., 2020. Experiences of hearing aid use among patients with mild cognitive impairment and Alzheimer's disease dementia: a qualitative study. SAGe Open. Med. 8, 2050312120904572.
Guest, H., Munro, K.J., Prendergast, G., Millman, R.E., Plack, C.J., 2018. Impaired speech perception in noise with a normal audiogram: no evidence for cochlear synaptopathy and no relation to lifetime noise exposure. Hear Res. 364, 142–151.
Guilleminot, P., Graef, C., Butters, E., Reichenbach, T., 2023. Audiotactile stimulation can improve syllable discrimination through multisensory integration in the theta frequency band. J. Cogn. Neurosci. 35, 1760–1772.
Harris, K.C., Ahlstrom, J.B., Dias, J.W., Kerouac, L.B., McClaskey, C.M., Dubno, J.R., Eckert, M.A., 2021. Neural presbyacusis in humans inferred from age-related differences in auditory nerve function and structure. J. Neurosci. 41, 10293–10304.
Harris, K.C., Mills, J.H., He, N.J., Dubno, J.R., 2008. Age-related differences in sensitivity to small changes in frequency assessed with cortical evoked potentials. Hear Res. 243, 47–56.

He, N.J., Mills, J.H., Dubno, J.R., 2007. Frequency modulation detection: effects of age, psychophysical method, and modulation waveform. J. Acoust. Soc. Am. 122, 467–477.
Heil, P., Neubauer, H., Brown, M., Irvine, D.R., 2008. Towards a unifying basis of auditory thresholds: distributions of the first-spike latencies of auditory-nerve fibers. Hear Res. 238, 25–38.
Herdman, A.T., Lins, O., Van Roon, P., Stapells, D.R., Scherg, M., Picton, T.W., 2002. Intracerebral sources of human auditory steady-state responses. Brain Topogr. 15, 69–86.
Hesse, L.L., Bakay, W., Ong, H.C., Anderson, L., Ashmore, J., McAlpine, D., Linden, J., Schaette, R., 2016. Non-monotonic relation between noise exposure severity and neuronal hyperactivity in the auditory midbrain. Front. Neurol. 7, 133.
Hoffman, D.W., Whitworth, C.A., Jones, K.L., Rybak, L.P., 1987. Nutritional status, glutathione levels, and ototoxicity of loop diuretics and aminoglycoside antibiotics. Hear Res. 31, 217–222.
Huet, A., Batrel, C., Dubernard, X., Kleiber, J.C., Desmadryl, G., Venail, F., Liberman, M. C., Nouvian, R., Puel, J.L., Bourien, J., 2022. Peristimulus time responses predict adaptation and spontaneous firing of auditory-nerve fibers: from rodents data to humans. J. Neurosci. 42, 2253–2267.
Huet, A., Batrel, C., Wang, J., Desmadryl, G., Nouvian, R., Puel, J.L., Bourien, J., 2019. Sound coding in the auditory nerve: from single fiber activity to cochlear mass potentials in gerbils. Neuroscience 407, 83–92.
Huet, A., Desmadryl, G., Justal, T., Nouvian, R., Puel, J.L., Bourien, J., 2018. The interplay between spike-time and spike-rate modes in the auditory nerve encodes tone-in-noise threshold. J. Neurosci. 38, 5727–5738.
Humes, L.E., Busey, T.A., Craig, J., Kewley-Port, D., 2013. Are age-related changes in cognitive function driven by age-related changes in sensory processing? Atten. Percept. Psychophys. 75, 508–524.
Hunter, L.L., Monson, B.B., Moore, D.R., Dhar, S., Wright, B.A., Munro, K.J., Zadeh, L.M., Blankenship, C.M., Stiepan, S.M., Siegel, J.H., 2020. Extended high frequency hearing and speech perception implications in adults and children. Hear Res. 397, 107922.
Johannesen, P.T., Perez-Gonzalez, P., Kalluri, S., Blanco, J.L., Lopez-Poveda, E.A., 2016. The influence of cochlear mechanical dysfunction, temporal processing deficits, and age on the intelligibility of audible speech in noise for hearing-impaired listeners. Trends. Hear. 20. https://doi.org/10.1177/2331216516641055.
Johnson, K.L., Nicol, T., Zecker, S.G., Bradlow, A.R., Skoe, E., Kraus, N., 2008. Brainstem encoding of voiced consonant–vowel stop syllables. Clin. Neurophysiol. 119, 2623–2635.
Keithley, E.M., 2020. Pathology and mechanisms of cochlear aging. J. Neurosci. Res. 98, 1674–1684.
Konrad-Martin, D., Dille, M.F., McMillan, G., Griest, S., McDermott, D., Fausti, S.A., Austin, D.F., 2012. Age-related changes in the auditory brainstem response. J. Am. Acad. Audiol. 23, 18–35 quiz 74-15.
Korczak, P., Smart, J., Delgado, R., Strobel, T.M., Bradford, C., 2012. Auditory steady-state responses. J. Am. Acad. Audiol. 23, 146–170.
Kubota, T., Ito, T., Abe, Y., Chiba, H., Suzuki, Y., Kakehata, S., Aoyagi, M., 2019. Detecting the recruitment phenomenon in adults using 80-Hz auditory steady-state response. Auris Nasus Larynx 46, 696–702.
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.
Kuwada, S., Anderson, J.S., Batra, R., Fitzpatrick, D.C., Teissier, N., D'Angelo, W.R., 2002. Sources of the scalp-recorded amplitude-modulation following response. J. Am. Acad. Audiol. 13, 188–204.
Lai, J., Bidelman, G.M., 2022. Relative changes in the cochlear summating potentials to paired-clicks predict speech-in-noise perception and subjective hearing acuity. JASa Express. Lett. 2, 102001.
Le Prell, C.G., 2019. Effects of noise exposure on auditory brainstem response and speech-in-noise tasks: a review of the literature. Int. J. Audiol. 58, S3–S32.
Le Prell, C.G., Spankovich, C., Lobarinas, E., Griffiths, S.K., 2013. Extended high-frequency thresholds in college students: effects of music player use and other recreational noise. J. Am. Acad. Audiol. 24, 725–739.
Levy, S.C., Freed, D.J., Nilsson, M., Moore, B.C., Puria, S., 2015. Extended high-frequency bandwidth improves speech reception in the presence of spatially separated masking speech. Ear Hear. 36, e214–e224.
Liberman, M.C., 1980. Morphological differences among radial afferent fibers in the cat cochlea: an electron-microscopic study of serial sections. Hear Res. 3, 45–63.
Liberman, M.C., Dodds, L.W., 1984. Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. Hear Res. 16, 55–74.
Liberman, M.C., Epstein, M.J., Cleveland, S.S., Wang, H., Maison, S.F., 2016. Toward a differential diagnosis of hidden hearing loss in humans. PLoS One 11, e0162726.
Liberman, M.C., Kujawa, S.G., 2017. Cochlear synaptopathy in acquired sensorineural hearing loss: manifestations and mechanisms. Hear Res. 349, 138–147.
Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S.G., Huntley, J., Ames, D., Ballard, C., Banerjee, S., Burns, A., Cohen-Mansfield, J., Cooper, C., Fox, N., Gitlin, L. N., Howard, R., Kales, H.C., Larson, E.B., Ritchie, K., Rockwood, K., Sampson, E.L., Samus, Q., Schneider, L.S., Selbaek, G., Teri, L., Mukadam, N., 2017. Dementia prevention, intervention, and care. Lancet 390, 2673–2734.
Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., Moore, B.C., 2006. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. Proc. Natl. Acad. Sci. U. S. A. 103, 18866–18869.
Lu, H., Mehta, A.H., Oxenham, A.J., 2022. Methodological considerations when measuring and analyzing auditory steady-state responses with multi-channel EEG. Curr. Res. Neurobiol. 3, 100061.
Marcher-Rorsted, J., Encina-Llamas, G., Dau, T., Liberman, M.C., Wu, P.Z., Hjortkjaer, J., 2022. Age-related reduction in frequency-following responses as a potential marker of cochlear neural degeneration. Hear Res. 414, 108411.
Marrufo-Perez, M.I., Lopez-Poveda, E.A., 2022. Adaptation to noise in normal and impaired hearing. J. Acoust. Soc. Am. 151, 1741.
Marszalek, J.R., Williamson, T.L., Lee, M.K., Xu, Z., Hoffman, P.N., Becher, M.W., Crawford, T.O., Cleveland, D.W., 1996. Neurofilament subunit NF-H modulates axonal diameter by selectively slowing neurofilament transport. J. Cell Biol. 135, 711–724.
Meddis, R., 2006. Auditory-nerve first-spike latency and auditory absolute threshold: a computer model. J. Acoust. Soc. Am. 119, 406–417.
Meister, H., Wenzel, F., Gehlen, A.K., Kessler, J., Walger, M., 2020. Static and dynamic cocktail party listening in younger and older adults. Hear Res. 395, 108020.
Mepani, A.M., Verhulst, S., Hancock, K.E., Garrett, M., Vasilkov, V., Bennett, K., de Gruttola, V., Liberman, M.C., Maison, S.F., 2021. Envelope following responses predict speech-in-noise performance in normal-hearing listeners. J. Neurophysiol. 125, 1213–1222.
Merchan-Perez, A., Liberman, M.C., 1996. Ultrastructural differences among afferent synapses on cochlear hair cells: correlations with spontaneous discharge rate. J. Comp. Neurol. 371, 208–221.
Monaghan, J.J.M., Garcia-Lazaro, J.A., McAlpine, D., Schaette, R., 2020. Hidden hearing loss impacts the neural representation of speech in background noise. Curr. Biol. 30, 4710–4721 e4714.
Monson, B.B., Hunter, E.J., Lotto, A.J., Story, B.H., 2014. The perceptual significance of high-frequency energy in the human voice. Front. Psychol. 5, 587.
Moore, B.C., 2008. Basic auditory processes involved in the analysis of speech sounds. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 947–963.
Moore, B.C.J., 2016. Effects of age and hearing loss on the processing of auditory temporal fine structure. Adv. Exp. Med. Biol. 894, 1–8.
Moore, B.C.J., 2021. Effects of hearing loss and age on the binaural processing of temporal envelope and temporal fine structure information. Hear Res. 402, 107991.
Moore, D., Hunter, L., Munro, K., 2017. Benefits of extended high-frequency audiometry for everyone. Hear J. 70, 50–55.
Morise, M., Yokomori, F., Ozawa, K., 2016. WORLD: a vocoder-based high-quality speech synthesis system for real-time applications. IEICe Trans. Inf. Syst. E99, 1877–1884.
Motlagh Zadeh, L., Silbert, N.H., Sternasty, K., Swanepoel, W., Hunter, L.L., Moore, D.R., 2019. Extended high-frequency hearing enhances speech perception in noise. Proc. Natl. Acad. Sci. U. S. A. 116, 23753–23759.
Oxenham, A.J., 2018. How we hear: the perception and neural coding of sound. Annu. Rev. Psychol. 69, 27–50.
Parthasarathy, A., Bartlett, E.L., Kujawa, S.G., 2019. Age-related changes in neural coding of envelope cues: peripheral declines and central compensation. Neuroscience 407, 21–31.
Parthasarathy, A., Datta, J., Torres, J.A., Hopkins, C., Bartlett, E.L., 2014. Age-related changes in the relationship between auditory brainstem responses and envelope-following responses. J. Assoc. Res. Otolaryngol. 15, 649–661.
Peelle, J.E., Wingfield, A., 2016. The neural consequences of age-related hearing loss. Trends Neurosci. 39, 486–497.
Pfefferbaum, A., Ford, J.M., Roth, W.T., Kopell, B.S., 1980. Age differences in P3-reaction time associations. Electroencephalogr. Clin. Neurophysiol. 49, 257–265.
Pichora-Fuller, M.K., Schneider, B.A., Macdonald, E., Pass, H.E., Brown, S., 2007. Temporal jitter disrupts speech intelligibility: a simulation of auditory aging. Hear Res. 223, 114–121.
Picton, T., Alain, C., Woods, D., John, M., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., Trujillo, N., 1999. Intracerebral sources of human auditory-evoked potentials. Audiol. Neurotol. 4, 64–79.
Plack, C.J., Barker, D., Prendergast, G., 2014. Perceptual consequences of "hidden" hearing loss. Trends. Hear. 18. https://doi.org/10.1177/2331216514550621.
Plourde, G., Garcia-Asensi, A., Backman, S., Deschamps, A., Chartrand, D., Fiset, P., Picton, T.W., 2008. Attenuation of the 40-Hz auditory steady state response by propofol involves the cortical and subcortical generators. Anesthesiology 108, 233–242.
Presacco, A., Simon, J.Z., Anderson, S., 2016a. Effect of informational content of noise on speech representation in the aging midbrain and cortex. J. Neurophysiol. 116, 2356–2367.
Presacco, A., Simon, J.Z., Anderson, S., 2016b. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. J. Neurophysiol. 116, 2346–2355.
Price, D., Tyler, L.K., Neto Henriques, R., Campbell, K.L., Williams, N., Treder, M.S., Taylor, J.R., Brayne, C., Bullmore, E.T., Calder, A.C., Cusack, R., Dalgleish, T., Duncan, J., Matthews, F.E., Marslen-Wilson, W.D., Rowe, J.B., Shafto, M.A., Cheung, T., Davis, S., Geerligs, L., Kievit, R., McCarrey, A., Mustafa, A., Samu, D., Tsvetanov, K.A., van Belle, J., Bates, L., Emery, T., Erzinglioglu, S., Gadie, A., Gerbase, S., Georgieva, S., Hanley, C., Parkin, B., Troy, D., Auer, T., Correia, M., Gao, L., Green, E., Allen, J., Amery, G., Amunts, L., Barcroft, A., Castle, A., Dias, C., Dowrick, J., Fair, M., Fisher, H., Goulding, A., Grewal, A., Hale, G., Hilton, A., Johnson, F., Johnston, P., Kavanagh-Williamson, T., Kwasniewska, M., McMinn, A., Norman, K., Penrose, J., Roby, F., Rowland, D., Sargeant, J., Squire, M., Stevens, B., Stoddart, A., Stone, C., Thompson, T., Yazlik, O., Barnes, D., Dixon, M., Hillman, J., Mitchell, J., Villis, L., Henson, R.N.A., Cam, C.A.N., 2017. Age-related delay in visual and auditory evoked responses is mediated by white- and grey-matter differences. Nat. Commun. 8, 15671.
Profant, O., Skoch, A., Balogova, Z., Tintera, J., Hlinka, J., Syka, J., 2014. Diffusion tensor imaging and MR morphometry of the central auditory pathway and auditory cortex in aging. Neuroscience 260, 87–97.
Puttabasappa, M., Rajanna, M., Jaisinghani, P., Shukla, S., 2017. Auditory P300 in typical individuals: age and gender effect. Int. J. Health Sci. Res. 7, 247–257.

Rance, G., 2008. The Auditory Steady-State Response: Generation, Recording, and Clinical Application. Plural Publishing.
Rasetshwane, D.M., Argenyi, M., Neely, S.T., Kopun, J.G., Gorga, M.P., 2013. Latency of tone-burst-evoked auditory brain stem responses and otoacoustic emissions: level, frequency, and rise-time effects. J. Acoust. Soc. Am. 133, 2803–2817.
Rhode, W.S., Smith, P.H., 1986. Encoding timing and intensity in the ventral cochlear nucleus of the cat. J. Neurophysiol. 56, 261–286.
Ross, B., Tremblay, K.L., Alain, C., 2020. Simultaneous EEG and MEG recordings reveal vocal pitch elicited cortical gamma oscillations in young and older adults. Neuroimage 204, 116253.
Schirmer, J., Wolpert, S., Dapper, K., Rühle, M., Wertz, J., Wouters, M., Eldh, T., Bader, K., Singer, W., Gaudrain, E., Başkent, D., Verhulst, S., Braun, C., Rüttiger, L., Munk, M.H.J., Dalhoff, E., Knipper, M., 2024. Neural adaptation at stimulus onset and speed of neural processing as critical contributors to speech comprehension independent of hearing threshold or age. J. Clin. Med. 13, 2725.
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.
Shaheen, L.A., Valero, M.D., Liberman, M.C., 2015. Towards a diagnosis of cochlear neuropathy with envelope following responses. J. Assoc. Res. Otolaryngol. 16, 727–745.
Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. Science (1979) 270, 303–304.
Shetty, H.N., Puttabasappa, M., 2017. Encoding of speech sounds at auditory brainstem level in good and poor hearing aid performers. Braz. J. Otorhinolaryngol. 83, 512–522.
Shukla, R., Trivedi, J.K., Singh, R., Singh, Y., Chakravorty, P., 2000. P300 event related potential in normal healthy controls of different age groups. Indian J. Psychiatry 42, 397–401.
Sliwinska-Kowalska, M., 2015. Hearing. Handb. Clin. Neurol. 131, 341–363.
Smith, R.L., Zwislocki, J.J., 1975. Short-term adaptation and incremental responses of single auditory-nerve fibers. Biol. Cybern. 17, 169–182.
Smith, Z.M., Delgutte, B., Oxenham, A.J., 2002. Chimaeric sounds reveal dichotomies in auditory perception. Nature 416, 87–90.
Snell, K.B., Frisina, D.R., 2000. Relationships among age-related differences in gap detection and word recognition. J. Acoust. Soc. Am. 107, 1615–1626.
Song, M., Wang, D., Li, J., Chen, G., Zhang, X., Wang, H., Wang, Q., 2022. Sudden sensorineural hearing loss as the initial symptom in patients with acoustic neuroma. Front. Neurol. 13, 953265.
Sun, S., Babola, T., Pregernig, G., So, K.S., Nguyen, M., Su, S.M., Palermo, A.T., Bergles, D.E., Burns, J.C., Muller, U., 2018. Hair cell mechanotransduction regulates spontaneous activity and spiral ganglion subtype specification in the auditory system. Cell 174, 1247–1263 e1215.
Suthakar, K., Liberman, M.C., 2021. Auditory-nerve responses in mice with noise-induced cochlear synaptopathy. J. Neurophysiol. 126, 2027–2038.
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.
Verschooten, E., Shamma, S., Oxenham, A.J., Moore, B.C.J., Joris, P.X., Heinz, M.G., Plack, C.J., 2019. The upper frequency limit for the use of phase locking to code temporal fine structure in humans: a compilation of viewpoints. Hear Res. 377, 109–121.
Vitela, A.D., Monson, B.B., Lotto, A.J., 2015. Phoneme categorization relying solely on high-frequency energy. J. Acoust. Soc. Am. 137, EL65–EL70.
von Gablenz, P., Holube, I., 2017. [Hearing loss and speech recognition in the elderly]. Laryngorhinootologie 96, 759–764.
Wagener, K., Brand, T., Kollmeier, B., 1999. Development and evaluation of a German sentence test part III: evaluation of the Olden-burg sentence test. Z. Audiol. 38, 86–95.
Weiss, T.F., Rose, C., 1988. A comparison of synchronization filters in different auditory receptor organs. Hear Res. 33, 175–179.
Wu, P.Z., Liberman, L.D., Bennett, K., de Gruttola, V., O'Malley, J.T., Liberman, M.C., 2019. Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear. Neuroscience 407, 8–20.
Xu, Z., Marszalek, J.R., Lee, M.K., Wong, P.C., Folmer, J., Crawford, T.O., Hsieh, S.T., Griffin, J.W., Cleveland, D.W., 1996. Subunit composition of neurofilaments specifies axonal diameter. J. Cell Biol. 133, 1061–1069.
Zelle, D., Dalhoff, E., Gummer, A.W., 2016. Objektive Hördiagnostik mit DPOAE. HNO 64, 822–830.
Zhao, R., Yue, T., Xu, Z., Zhang, Y., Wu, Y., Bai, Y., Ni, G., Ming, D., 2024. Electroencephalogram-based objective assessment of cognitive function level associated with age-related hearing loss. Geroscience 46, 431–446.
Zohar, O., Shackleton, T.M., Nelken, I., Palmer, A.R., Shamir, M., 2011. First spike latency code for interaural phase difference discrimination in the guinea pig inferior colliculus. J. Neurosci. 31, 9192–9204.